Technical Physics, Vol. 46, No. 9, 2001, pp. 1069–1075. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 1–8.

THEORETICAL AND MATHEMATICAL PHYSICS

Removal of Asphalt–Paraffin Deposits in Oil Pipelines by a Moving Source of High-Frequency Electromagnetic Radiation

V. A. Balakirev, G. V. Sotnikov, Yu. V. Tkach, and T. Yu. Yatsenko

Institute of Electromagnetic Investigations, Kharkov, 310022 Ukraine

e-mail: sotnikov@kipt.kharkov.ua

Received October 11, 2000; in final form, March 6, 2001

Abstract—A new method for removing asphalt–paraffin and gasohydrate plugs in oil pipelines with a movable source of electromagnetic radiation, electromagnetic pig, is suggested. The pig melts the plug when the latter absorbs intense electromagnetic radiation and heats up. Effective melting of the dielectric plug is achieved with the source moving along the pipeline as the solid–liquid interface propagates. The time of paraffin plug removal and the dependence of this time on the radiation frequency are found with the model suggested. The efficiency of the method is estimated. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Today, the transportation of natural hydrocarbons through oil and gas pipelines is becoming more and more popular. However, the conditions of transportation often favor the formation of thick asphalt–paraffin or gas hydrate plugs. The reliable and cost-efficient exploitation of such pipelines under these conditions requires advanced plug removal methods to be devised. Techniques to prevent plug formation are of no less importance.

The use of one or another currently available plug removal method depends on the composition of the plug, its structure, properties, etc. Asphalt–paraffin plugs in oil pipelines are usually destroyed by heating (with steam or hot water), chemically, or mechanically. All these approaches are expensive and technically difficult; moreover, they (e.g., chemical methods) sometimes entail adverse side effects.

Because of this, the use of high-power electromagnetic radiation to thermally destroy asphalt–paraffin deposits in oil wells, compressor and pumping plants, and oil pipelines is of indubitable interest.

Early (and basic) research on the use of electromagnetic radiation in the oil industry has been concentrated on increasing the oil production capacity by applying high-frequency electromagnetic (HF EM) radiation to oil pools [1–5]. When absorbed, the electromagnetic field heats oil, making it less viscous. The idea of using HF EM radiation to remove paraffin plugs from oil wells has been put forward in [6, 7]. In [8], the numerical simulation of heating a paraffin plug that fills part of an oil well was performed. In all these works, the frequency of EM radiation was relatively small (less than 100 MHz). This is dictated by the value of the volume heat release density, $Q \sim \alpha\exp(-\alpha z)$, in the plug due to the absorption of EM radiation (α is the absorption coefficient, z is the distance to the radiation source). At high frequencies, the absorption coefficient varies as the radiation frequency (if the dielectric loss tangent is weakly dependent on frequency). Therefore, if the frequency and, accordingly, the attenuation of the HF power in the plug are too low, the heat release is insignificant and the plug is not heated to the desired level. Asphalt–paraffin plugs are transparent to low-frequency radiation. If, however, the frequency is too high, the radiation is absorbed in the region nearest the source and the oil severely overheats. Therefore, the frequency is selected such that the absorption coefficient is on the order of the reciprocal paraffin plug length. This condition is readily met upon heating oil pools or upon removing paraffin plugs in a well, which can be viewed as coaxial transmission lines in terms of electrodynamics. In an actual oil pipeline, however, this condition is impossible to satisfy. An oil pipeline is a cylindrical waveguide that transmits only waves with higher-than-cutoff frequencies. For example, a pipeline of radius 72 cm filled with paraffin with the permittivity reported in [7–9] has the cutoff frequency $f = 1.048 \times 10^8$ Hz for the E_{01} wave. At an operating frequency $f = 1.4 \times 10^8$ Hz, the power attenuation coefficient $\alpha = 0.08$ m^{-1}. If the paraffin length is 100 m, the EM power is attenuated e^8 times. Obviously, the paraffin plug will never be melted in this case.

In this work, we suggest removing paraffin plugs in real oil pipelines using a movable EM radiation source. Its velocity depends on the propagation rate of the liquid–solid interface during the melting of asphalt–paraffin deposits by HF EM radiation. The device suggested will be called an electromagnetic pig analogous to a mechanical pig (scraper) used in the oil and gas indus-

try for the mechanical cleaning of paraffin and gasohydrate plugs from the pipelines.

BASIC EQUATIONS

The heating and melting of a paraffin plug will be described in terms of the thermal conduction equation with a given external heat source. As was noted, a paraffin plug in the pipeline can be considered as a dielectric-filled cylindrical metallic waveguide. It is assumed that the problem is axisymmetric and the waveguide is filled completely. The thermal conduction equation will be solved numerically with the through-calculation method [10]. Because of this, the equation for thermal conduction will be written in the general form without separating the phases in the explicit form:

$$\rho c_T \frac{\partial T}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(\lambda r \frac{\partial T}{\partial r}\right) + \frac{\partial}{\partial z}\left(\lambda \frac{\partial T}{\partial z}\right) + Q(r, z, t), \quad (1)$$

where ρ is the density of high-paraffin oil, c_T is its specific heat, and λ is the thermal conductivity.

The density and the thermal conductivity are assumed to be temperature-independent, while the specific heat at the phase transition temperature T_S has δ-like singularity:

$$c_T = c_0 + L\delta(T - T_S), \quad (2)$$

where L is the latent heat of phase transition and $\delta(T - T_S)$ is the delta function.

The volume heat release density is expressed as

$$Q = Q_0\Theta(z - z_0(t))\exp(-\alpha(z - z_0(t))). \quad (3)$$

This formula implies that the EM radiation source moves following the law $z = z_0(t)$. The explicit form of Q_0 is given below [see formula (7)]. In (3),

$$\Theta(z - z_0) = \begin{cases} 1, & z \geq z_0 \\ 0, & z < z_0, \end{cases}$$

and α is the coefficient of HF power attenuation.

The source is situated in the plane $z = 0$ and starts generating at the time instant $t = 0$. The EM wave decays because of volume losses in the dielectric plug and surface losses in the metallic walls of the cylindrical waveguide (the walls of the waveguide have a finite impedance). In view of the axial symmetry of the problem, it is easy to derive an equation from which the attenuation coefficient $\alpha \equiv \alpha_V + \alpha_S = 2k_z''$ (α_V and α_S are the attenuation coefficients due to volume and surface losses, and k_z'' is the imaginary part of the longitudinal wave number $k_z = k_z' + ik_z''$)

for the wave E_{0n} can be determined:

$$k_\perp \frac{J_0(k_\perp R)}{J_1(k_\perp R)} = \frac{1 + i}{2} k\varepsilon_0 \sqrt{\frac{\omega}{2\pi\sigma}}. \quad (4)$$

In (4), $k_\perp^2 = k^2\varepsilon_0 - k_z^2$, $k = \omega/c$, ω is the circular frequency, c is the speed of light, J_0 and J_1 are the Bessel functions of the zeroth and the first order, ε_0 is the complex permittivity of paraffin, R is the radius of the waveguide, and σ is the conductivity of the pipeline metal. Assuming that the imaginary part ε_0'' of the permittivity is much smaller than the real part ε_0', we obtain, from (3), the approximate solution

$$\alpha_V = \frac{\omega^2}{c^2}\frac{\varepsilon_0''}{k_{z_0}'}, \quad \alpha_S = \frac{\omega\varepsilon_0'}{cRk_{z_0}'}\sqrt{\frac{\omega}{2\pi\sigma}},$$
$$k_{z_0}' = \sqrt{\frac{\omega^2}{c^2}\varepsilon_0' - \frac{\mu_n^2}{R^2}}, \quad (5)$$

where μ_n is the nth zero of the Bessel function J_0.

For a stationary source, the volume heat release density can be determined from the HF losses when the EM wave passes through an absorbing medium:

$$Q_0 = \frac{\omega}{8\pi}|\mathbf{E}|^2\varepsilon_0''. \quad (6)$$

For the case of E_{0n}, the other-than-zero components of the electric field are E_z and E_r, the expressions for which are well known in the case of a dielectric-filled cylindrical waveguide [11]. After necessary transformations, we obtain

$$Q_0 = \frac{\varepsilon_0''}{\pi \mathrm{Re}(k_z\varepsilon_0)}\frac{\mu_n^2}{R^4}$$
$$\times P\left[J_0^2\left(\mu_n\frac{r}{R}\right) + \frac{|k_z|^2R^2}{\mu_n^2}J_1^2\left(\mu_n\frac{r}{R}\right)\right], \quad (7)$$

where P is the HF power radiated from the source.

Thermal conductivity equation (1) must be complemented by boundary conditions. At the near end face of the plug, $z = 0$, we specify a boundary condition in the form of convective heat exchange obeying Newton's law:

$$\lambda\left.\frac{\partial T}{\partial z}\right|_{z=0} = \kappa_1[T(r, 0, t) - T_0], \quad (8)$$

where T_0 is the ambient temperature and the initial temperature of the paraffin plug and κ_1 is the heat exchange coefficient.

At the far end of the plug, $z = H$, heat exchange is absent:

$$\lambda\left.\frac{\partial T}{\partial z}\right|_{z=H} = 0. \quad (9)$$

At the side surface of the cylinder, $r = R$, we also specify convective heat exchange but with a different κ. In addition, we take into account heat release due to the

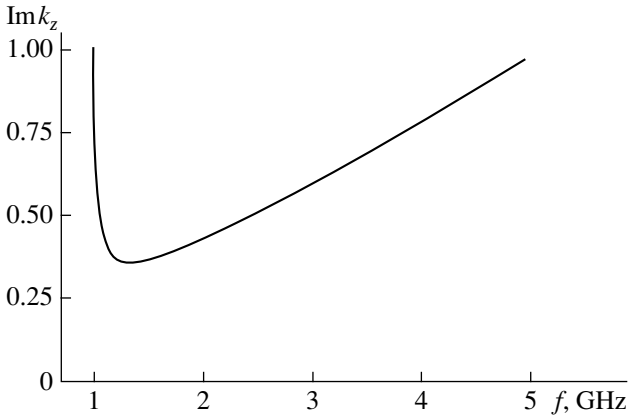


Fig. 1. Frequency dependence of the imaginary part of the longitudinal wave number for a paraffin-filled metallic cylindrical waveguide. The relative permittivity $\varepsilon = 2.3 + i0.0276$ [12, 13].

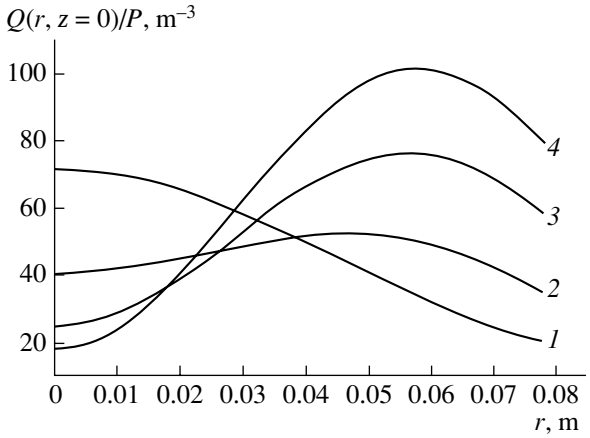


Fig. 2. Transverse distribution of the heat loss power density normalized to the source power in the paraffin-filled cylindrical waveguide. f = (*1*) 1.4, (*2*) 2, (*3*) 3, and (*4*) 4×10^9 Hz.

absorption of the EM wave at the surface of the metallic waveguide:

$$-\lambda \frac{\partial T}{\partial r}\bigg|_{r=R} = \kappa[T(R, z, t) - T_0] - q(z, t), \qquad (10)$$

where $\kappa = \mathrm{Nu}\lambda/R$ is the coefficient of heat exchange with the environment and Nu is the Nusselt number.

The power absorbed in the metallic walls of the pipeline can be found by solving the rigorously posed electromagnetic problem in the metal and in the dielectric. Having determined the electric field components and applying a formula like (6) to the metallic area with a appropriate permittivity, we find the power absorbed by the metal volume. Actually, the skin depth is much less than the pipeline wall thickness; therefore, we can assume that the HF power is absorbed at the inner surface of the pipeline. With the skin depth tending to zero, the expression for the surface heat release density is given by

$$q = \frac{\omega|\varepsilon_0|^2}{c\,\mathrm{Re}(k_z\varepsilon_0)}\sqrt{\frac{\omega}{2\pi\sigma}}\frac{P}{2\pi R^2}\exp(-\alpha(z - z_0(t))) \times \Theta(z - z_0(t)). \qquad (11)$$

Here, as in (3), we take into account the fact that the source radiates only in a forward direction and moves in the pipeline longitudinally, following the law $z = z_0(t)$.

To close the set of equations that characterizes the removal of a paraffin plug by a moving EM radiation source, it is necessary to specify the law of its motion $z_0(t)$. We noted in the introduction that $z_0(t)$ is defined by the motion of the solid–liquid interface. However, the motion of the interface is governed also by differential equation (1). In the explicit form, an equation for interface velocity can be written only for the one-dimensional case [10]. Therefore, when numerically solving Eq. (1) with boundary conditions (8)–(10) and a complex spatial dependence of $Q(r, z)$, we will assume that, during motion, the spacing between the source and the interface remains constant.

NUMERICAL RESULTS

For the numerical analysis of asphalt–paraffin plug removal, we took advantage of the physical parameters of high-paraffin oil [8]: $\rho = 950$ kg/m^3, $c_0 = 3$ kJ/(kg K), $T_S = 50$°C, $L = 300$ kJ/kg, and $\lambda = 0.125$ W/(m K). The ambient temperature was set equal to $T_0 = 20$°C. The real and the imaginary parts of the dielectric constant ε_0 were taken for dehydrated oil, since they are weakly dependent on frequency over a wide range [12, 13]. In the calculations, we used $\varepsilon_0' \approx 2.3$ and $\tan\delta = \varepsilon_0''/\varepsilon_0' \approx 1.2 \times 10^{-2}$. The heat exchange coefficients were $\kappa_1 = 0.2$ W/(m^2 K) and $\kappa = 1.613$ W/(m^2 K). The value of κ corresponds to the Nusselt number Nu = 1 (the pipeline in dry soil). To numerically solve Eq. (1) with boundary conditions (8)–(10), we applied an explicit difference scheme on a uniform rectangular mesh. The singularity at the point $r = 0$ in the Laplacian was bypassed in the standard way [14]: the δ function in Eq. (2) for the specific heat was approximated by a step with a half-width of 0.4°C.

To check the accuracy with which the difference scheme approximates the set of equations, we used the equation for energy balance that is directly derived from Eq. (1) and relationships (7)–(10):

$$\rho\frac{\partial}{\partial t}\int dV\int_{T_0}^{T} c(T')dT' = P[1 - \exp(-\alpha(H - z_0(t)))]$$

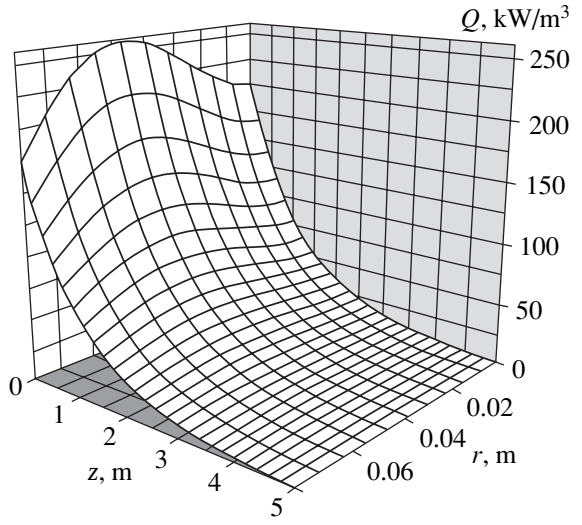


Fig. 3. Distribution of the heat loss power density over the volume of the paraffin plug.

$$-2\pi\kappa_1\int_0^R drr[T(r, z = 0) - T_0] \qquad (12)$$

$$-2\pi\kappa R\int_0^H dz[T(r = R, z) - T_0].$$

Energy balance (12) coincides with that in [15] if the EM radiation source is stationary [$z_0(t) = 0$].

The results given below were obtained for a model pipeline of radius $R = 0.0775$ m. It was assumed to be made of steel with a conductivity $\sigma = 0.37 \times 10^{17}$ s^{-1}. The paraffin plug length H was set equal to 5 m.

Figure 1 plots the imaginary part k_z'' of the longitudinal wave number against frequency for the above parameters of the plugged cylindrical pipeline. Hereafter, we will consider only the wave E_{01}. The cutoff frequency for E_{01} for this pipeline is $f_0 \simeq 0.97 \times 10^9$ Hz.

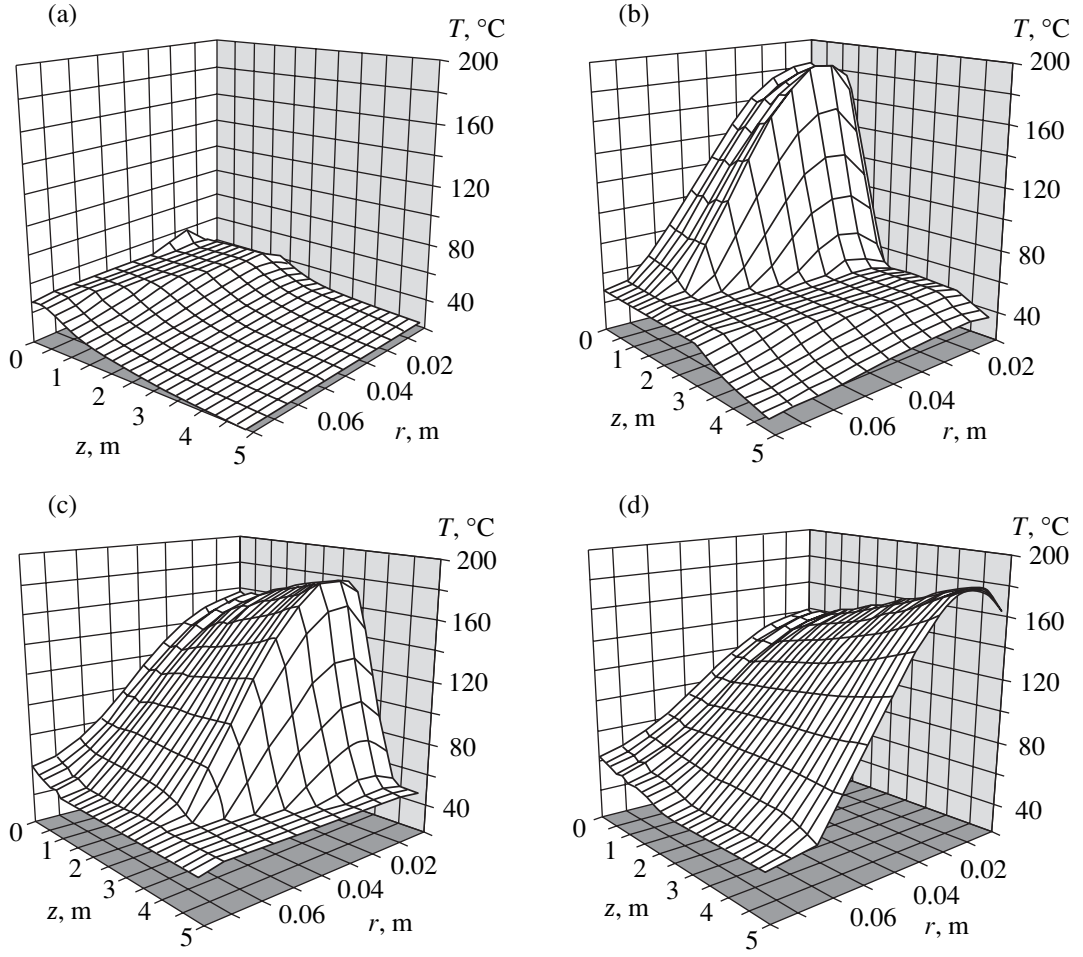


Fig. 4. Temperature profiles along the pipeline length $H = 5$ m at different time instants during the removal of the asphalt–paraffin plug by the electromagnetic pig. $R = 0.0775$ m, $f = 1.4 \times 10^9$ Hz. t = (a) 15, (b) 75, (c) 135, and (d) 180 min.

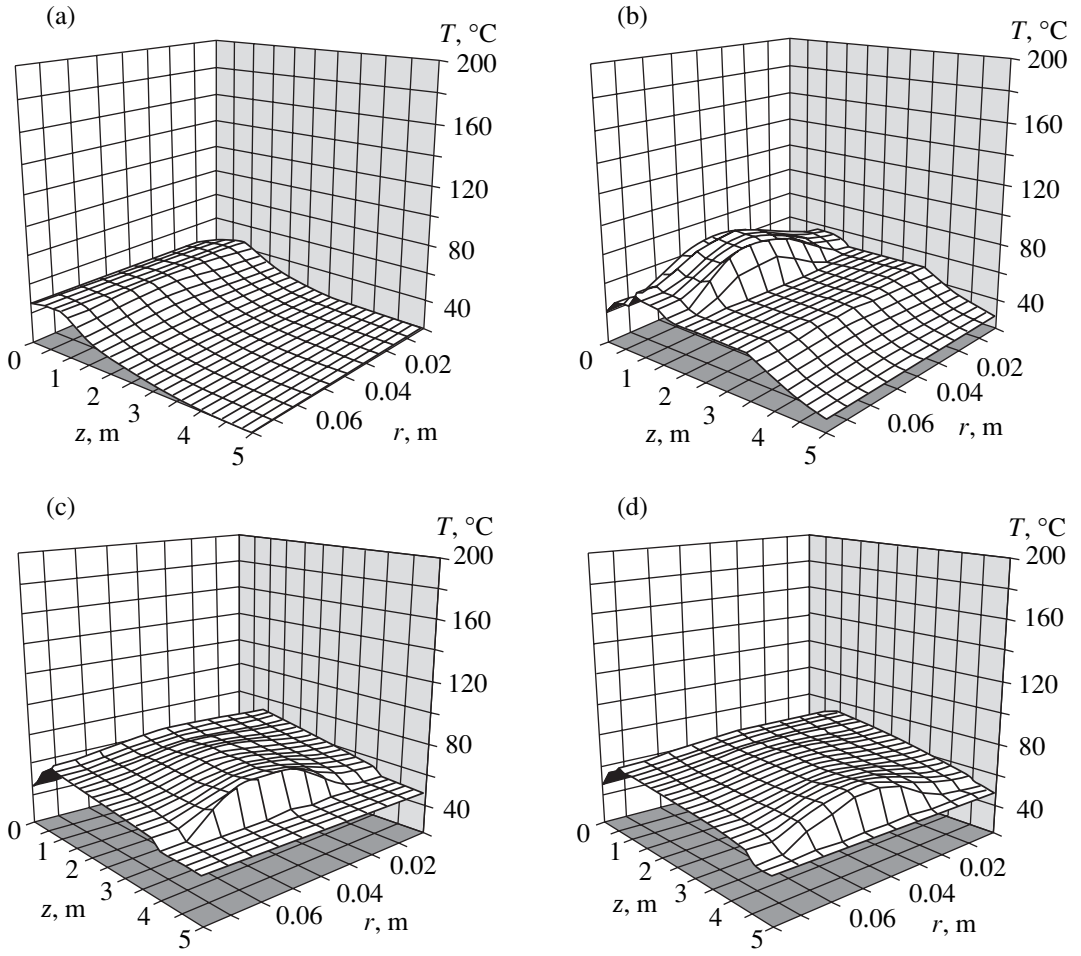


Fig. 5. The same as in Fig. 4 for $f = 2 \times 10^9$ Hz (in panel "d", $t = 150$ min).

The value of $k_z'' \approx 0.37$ m^{-1} has a minimum at $f \simeq 1.38 \times 10^9$ Hz and then grows with frequency. Although the depth of HF power penetration into the plug is of minor importance in our method of plug removal, it should be taken as long as possible. A penetration depth that is too short would make the on-line control of source movement difficult.

The phase transition substantially depends on the distribution of the volume heat release density $Q(r, z)$. In Fig. 2, the distributions $Q(r, z = 0)$ in the waveguide cross section that are normalized to the power of the stationary source are shown for various frequencies. At the minimum of the attenuation coefficient (and smaller frequencies), the longitudinal component of the electric field E_{01} with a maximum at the axis of the cylindrical waveguide (Fig. 2, curve *1*) plays a dominant part in the distribution of the heat release density Q. As the frequency increases, so does the transverse component of the electric field. At $f = 2 \times 10^9$ Hz, the transverse and the longitudinal components are comparable in magnitude (Fig. 2, curve *2*). At higher frequencies, the heat release density distribution is governed by the transverse component of the electric field E_{01}, whose maximum is closer to the waveguide walls (Fig. 2; curves *3*, *4*). The wave with a frequency $f = 2 \times 10^9$ Hz provides the most uniform distribution of the heat density over the cross section. In this case, the thermal energy density in the peak $Q(r, z = 0)$ is smaller than in the other cases in Fig. 2; however, it is the E_{01} wave with a frequency of 2×10^9 Hz that is the most preferable for removing a paraffin plug with radius $R = 7.75$ cm. In this case, heating is uniform; hence, the paraffin uniformly melts over the cross section of the plug. The overheating of separate oil layers, which reduces the efficiency of the method and causes some engineering troubles, is prevented. In the longitudinal direction, the heat release density drops exponentially. As the frequency rises, the density $Q(r, z)$ in the longitudinal direction drops faster, as demonstrated in Fig. 1. Figure 3 exemplifies $Q(r, z)$ for a stationary source $z_0(t) = 0$ and $P = 5$ kW.

The results of the numerical simulation of paraffin plug removal with the electromagnetic pig are shown in Figs. 4–6 for $f = 1.4 \times 10^9$ Hz, 2×10^9 Hz, and 3×10^9,

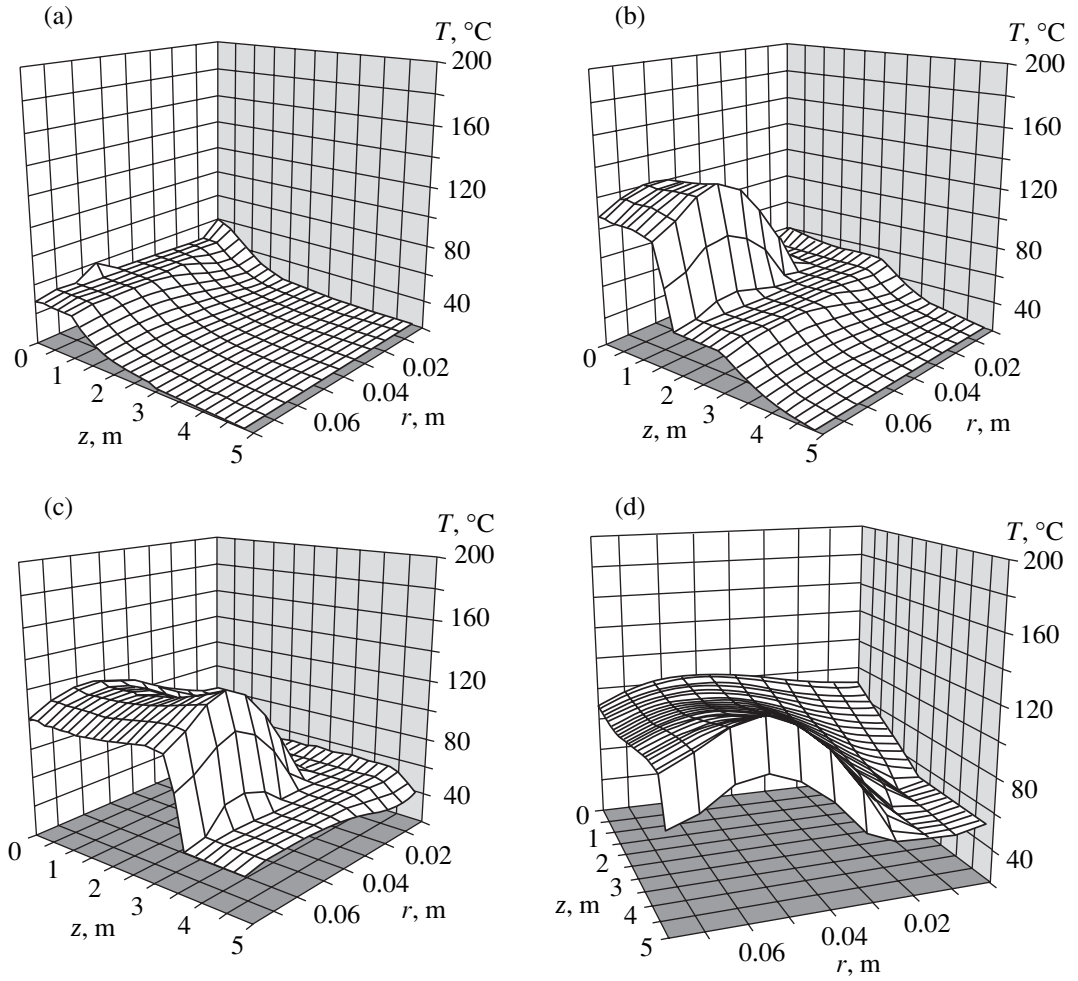


Fig. 6. The same as in Fig. 4 for $f = 3 \times 10^9$ Hz (in panel "d", t = 195 min).

respectively, at different time instants. In all cases, the source power is P = 5 kW. As was expected, the melting process strongly depends on the $Q(r, z)$ distribution (Figs. 2, 3). Initially, the temperature surface has a plateau near the source. This plateau is the widest at $f = 2 \times 10^9$ Hz, because the distribution of the thermal energy density over the plug cross section is the most uniform in this case. The region where melting starts strictly correlates with the peak of the $Q(r, z)$ curve. At a low frequency $f = 1.4 \times 10^9$ Hz, melting begins at the center of the dielectric plug. At high frequencies, $f \geq 3 \times 10^9$ Hz, the process starts closer to the periphery of the cylinder. At $f = 2 \times 10^9$ Hz, the process starts propagating from the middle of the cylinder radius. In the last case, the distribution of the oil and paraffin temperatures over the cross section turns out to be the most uniform. For this frequency, the maximal temperature of the melted paraffin and oil does not exceed 75°C, while at $f = 1.4 \times 10^9$ Hz, it approaches 175°C and at $f = 3 \times 10^9$ Hz, the maximal temperature is ≈120°C. The time of plug removal for H = 5 m strongly depends on the frequency in the range 1–5 GHz under study. At $f = 1.4 \times 10^9$ Hz, this time is 3 h; at $f = 2 \times 10^9$ Hz, it is the least, 2.5 h; and at $f = 3 \times 10^9$ Hz, it equals 3.25 h.

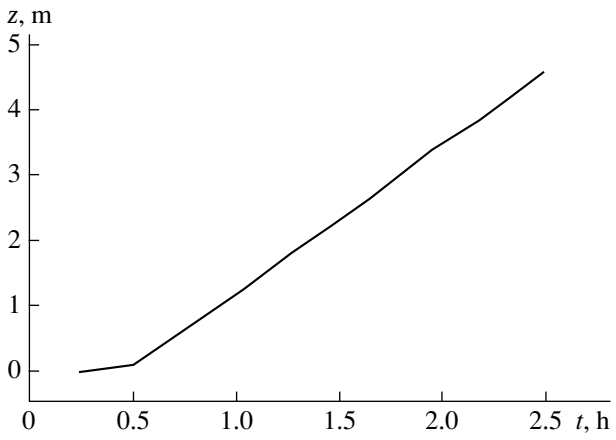


Fig. 7. Position of the electromagnetic pig relative to its initial location in the pipeline at different time instants.

With further increase in the radiation frequency, the time of plug removal grows, because the penetration depth of the HF power in the direction of pig motion decreases and the radial distribution of the heat release density becomes less uniform. A progressively larger part of the useful power is spent on local overheating rather than on melting the whole plug, and is dissipated through the side surface.

Figure 7 depicts the source position along the pipeline relative to its initial location. The source power is $P = 5$ kW, and the EM radiation frequency is $f = 1.4 \times 10^9$ Hz. Within $t < 0.5$ h, the electromagnetic pig remains stationary until the early layer is melted in its cross section. Then, the pig velocity increases and becomes constant in 0.5 h.

CONCLUSION

Let us estimate the efficiency of our method. We define the efficiency of the pig as the useful-to-total work ratio. By useful work, we mean the work spent on the complete melting of the paraffin. The total work is the energy emitted by the source into the oil pipeline. For $f = 2 \times 10^9$ Hz and the plug parameters listed above, the efficiency is ≈70%.

Thus, our investigations have shown that the removal of extended plugs in pipelines is possible if the velocity of the EM source is self-consistently derived from the law of motion of the liquid–solid interface.

REFERENCES

1. F. L. Sayakhov, S. I. Chistyakov, G. A. Babalyan, and B. I. Fedorov, Izv. Vyssh. Uchebn. Zaved., Neft' Gaz, No. 2, 47 (1972).
2. F. L. Sayakhov, G. A. Babalyan, and A. N. Al'mesh'ev, Neft. Khoz., No. 12, 33 (1975).
3. F. L. Sayakhov, M. Kh. Fatykhov, and O. L. Kuznetsov, Izv. Vyssh. Uchebn. Zaved., Neft' Gaz, No. 3, 36 (1981).
4. Ngok Haĭ Zyong, A. G. Katushev, and R. N. Nigmatulin, Prikl. Mat. Mekh. **51**, 29 (1987).
5. Ngok Haĭ Zyong, N. D. Musaev, and R. I. Nigmatulin, Prikl. Mat. Mekh. **51**, 973 (1987).
6. F. L. Sayakhov, M. A. Fatykhov, and N. Sh. Imashev, USSR Inventor's Certificate No. 1314756, Byull. Izobret., No. 20 (1986).
7. A. T. Akhmetov, A. I. D'yachuk, A. A. Kislitsyn, *et al.*, RF Inventor's Certificate No. 1707190, Byull. Izobret., No. 3 (1992).
8. A. A. Kislitsyn, Prikl. Mekh. Tekh. Fiz. **37** (3), 75 (1996).
9. A. A. Kislitsyn, Prikl. Mekh. Tekh. Fiz. **34** (3), 97 (1993).
10. A. A. Samarskiĭ and B. D. Moiseenko, Zh. Vychisl. Mat. Mat. Fiz. **5**, 816 (1965).
11. L. A. Vaĭshteĭn, *Electromagnetic Waves* (Radio i Svyaz', Moscow, 1988).
12. S. I. Chistyakov, F. L. Sayakhov, and L. N. Bondarenko, Neft. Khoz., No. 11, 51 (1969).
13. S. I. Chistyakov, N. F. Denisova, and F. L. Sayakhov, Izv. Vyssh. Uchebn. Zaved., Neft' Gaz, No. 5, 53 (1972).
14. A. A. Samarskiĭ and A. V. Gulin, *Numerical Methods* (Nauka, Moscow, 1989).
15. V. A. Balakirev, G. V. Sotnikov, Yu. V. Tkach, and T. Yu. Yatsenko, Elektromagn. Yavleniya **1**, 552 (1998).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1076–1081. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 9–14.

THEORETICAL AND MATHEMATICAL PHYSICS

The Effect of Clustering on the Melting of III–V Semiconductors

S. V. Bulyarskiĭ, P. E. L'vov, and V. V. Svetukhin

Ul'yanovsk State University, Ul'yanovsk, 432700 Russia

e-mail: pavel@sv.uven.ru

Received December 14, 2000

Abstract—A thermodynamic model of cluster formation in the melts of binary semiconductor compounds is proposed. Cluster formation in the gallium arsenide melt is studied, and expressions for equilibrium concentrations of clusters of different size, as well as a liquidus equation, are obtained. In the limiting case of small cluster concentration, the latter equation is consistent with the results of the theory of quasi-chemical interaction. The system of equations derived enables the satisfactory description of experimental data on aftermelting and liquidus curves in gallium arsenide. The calculated values of the melting enthalpy and entropy agree well with data for the energy of dissociation. *© 2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Melting and crystallization of solids remain some of the least-studied problems in solid-state physics. As the temperature increases, the atomic arrangement in a solid becomes disordered and the atoms pass into liquid (melt). The long-range order breaks down in a certain temperature range and some intermediate state exists between the solid and the melt. This state is characterized by the presence of small-sized ordered regions, or atomic clusters. The existence of the clusters at temperatures above melting points (the aftermelt phenomenon) has been proved experimentally [1, 2]. In many materials, the percentage of particles forming clusters at melting points may reach 10%.

A liquidus curve, or the dependence of the phase transition temperature on the melt composition, is of great importance for semiconductor technology. Liquidus curves are most often derived in terms of the theory of quasi-chemical interaction. This theory relies on the equality of the chemical potentials upon the formation of interatomic bonds in solution components and ignores clustering, thus deeling with the total concentrations of atoms. However, such an approach is valid only at temperatures far exceeding the melting point.

In this paper, a model of interparticle interaction in the liquid phase is proposed. It includes cluster formation and enables the description of both liquidus and aftermelt curves.

MELT FREE ENERGY WITH RESPECT TO CLUSTERING

Let us consider clustering in the melt of a III–V semiconductor. Results obtained here can readily be generalized to any other binary compound.

We represent the Gibbs energy of melt as the sum of the energy of free particles and the energy of particles aggregated into clusters:

$$G = G^L + G^S. \quad (1)$$

(1) The free energy of melt G^L can be written in the form [3]

$$G^L = N_A^L \sum_i g_A^i p_i + N_B^L \sum_i g_B^i p_i - kT \ln\left[\frac{N^L!}{N_A^L! N_B^L!}\right], \quad (2)$$

where N_A^L and N_B^L are the numbers of particles A and B, respectively, in the melt; g_A^i and g_B^i are the energies of particles A and B that have i bonds with particles of the same sort; and $N^L = N_A^L + N_B^L$ is the total number of sites in the melt.

Formula (2) involves the probabilities p_i, which are given by

$$p_i = \frac{z!}{i!(z-i)!} \frac{(N_A^L)^{z-i} (N_B^L)^i}{(N_A^L + N_B^L)^z}, \quad (3)$$

where z is the number of nearest neighbors.

Let g_{AA}^L and g_{BB}^L be the energies of pair interaction between like particles and g_{AB}^L be the energy of pair interaction between particles A and B. In this notation, the energies g_A^i and g_B^i are expressed in the form

$$\begin{aligned} g_A^i &= \frac{1}{2} g_{AA}^L i + \frac{1}{2} g_{AB}^L (z-i) = \frac{1}{2}(g_{AA}^L - g_{AB}^L) i + \frac{z}{2} g_{AB}^L, \\ g_B^i &= \frac{1}{2} g_{BB}^L i + \frac{1}{2} g_{AB}^L (z-i) = \frac{1}{2}(g_{BB}^L - g_{AB}^L) i + \frac{z}{2} g_{AB}^L. \end{aligned} \quad (4)$$

The substitution of (3) and (4) into (2) makes it possible to sum over i and considerably simplify further

1063-7842/01/4609-1076$21.00

calculations:

$$G^L = N_A^L\left(\frac{z}{2}g_{AB}^L + \frac{z}{2}(g_{AA}^L - g_{AB}^L)\frac{N_B^L}{N_A^L + N_B^L}\right)$$

$$+ N_B^L\left(\frac{z}{2}g_{AB}^L + \frac{z}{2}(g_{BB}^L - g_{AB}^L)\frac{N_A^L}{N_A^L + N_B^L}\right)$$

$$- kT\ln\left[\frac{(N_A^L + N_B^L)!}{N_A^L!\,N_B^L!}\right] = \frac{z}{2}g_A^L(N_A^L + N_B^L) \qquad (5)$$

$$+ \frac{z}{2}(g_{AA} + g_{BB} - 2g_{AB})\frac{N_A^L N_B^L}{N_A^L + N_B^L}$$

$$- kT\ln\left[\frac{(N_A^L + N_B^L)!}{N_A^L!\,N_B^L!}\right].$$

Now the melt energy per particle of either sort can be obtained [3]:

$$g_A^L = \frac{\partial G^L}{\partial N_A^L} = \frac{z}{2}g_{AB}^L + \Omega(x_B^L)^2 + kT\ln[x_A^L],$$
$$g_B^L = \frac{\partial G^L}{\partial N_B^L} = \frac{z}{2}g_{AB}^L + \Omega(x_A^L)^2 + kT\ln[x_B^L], \qquad (6)$$

where x_A^L and x_B^L are the relative concentrations of the particles in the melt and

$$\Omega = \frac{z}{2}(g_{AA}^L + g_{BB}^L - 2g_{AB}^L)$$

is the parameter of quasi-chemical interaction.

The energies g_{AA}^L, g_{BB}^L, and g_{AB}^L are expressed in terms of the entropy and enthalpy of pair interaction; therefore, the interaction parameter Ω should linearly depend on temperature

$$\Omega = a - bT, \qquad (7)$$

where a and b are constants.

Equation (6) is independent of the absolute number of particles in the melt; thus, Eq. (5) can be written in the form

$$G^L = g_A^L N_A^L + g_B^L N_B^L. \qquad (8)$$

It is important to note that g_A^L and g_B^L are not the chemical potentials of particles A and B, because the chemical potential is defined as the derivative of total energy (1) with respect to the total number of particles of a given sort,

$$\mu_A = \frac{\partial G}{\partial N_A} \neq \frac{\partial G^L}{\partial N_A^L}, \qquad (9)$$

and necessarily depends on the energy of clustered particles.

(2) The free energy of clusters in a system is given by [4]

$$G^S = \sum_m g_m N_m \qquad (10)$$

$$- kT\ln\left[\frac{N!}{N^L!\prod_m N_m!(m_A^A!\,m_B^B!\,m_A^B!\,m_B^A!)^{N_m}}\prod_m (NP_m)^{N_m}\right].$$

Here, g_m is the energy of a cluster of m particles, P_m is the formation probability of a cluster of m particles, the factor N takes into account that a cluster may be arbitrarily located within the system, and m_α^β is the number of particles of sort $\alpha = \{A, B\}$ at the site $\beta = \{A, B\}$ in a cluster of size m [5].

Suppose that the clusters are stoichiometric and do not contain antisite defects. Then, for each cluster of m particles, the following relationships are true:

$$m_A = m_A^A = m/2, \quad m_B = m_B^B = m/2,$$
$$m_B^A = m_A^B = 0. \qquad (11)$$

This approximation can be used, because the homogeneity region for III–V semiconductors is very narrow; i.e., the variation of the component ratio in the melt practically does not change the composition of the solid phase.

The formation probability of a cluster having m particles of one sort is given by

$$P_m = \left(\frac{1}{N}\right)^m\left(1 - \frac{1}{N}\right)^{N' - m}. \qquad (12)$$

Since $N' \gg m$ and $N \approx N'$ (this follows from the experimentally observed incompressibility of liquids), the second factor in (12) can be omitted as having the order of unity.

In the case of clusters consisting of two unlike particles occupying sites of a certain type (in III–V semiconductors, these are sites in the two sublattices), the probability will have the form

$$P_m = \left(\frac{1}{2N}\right)^m. \qquad (13)$$

Using (11) and (13), we obtain the thermodynamic probability W appearing in (10):

$$W = \frac{N!}{\prod_m N_m!((m/2)!)^{2N_m}}\prod_m\left(N\frac{1}{(2N)^m}\right)^{N_m}. \qquad (14)$$

With the Stirling formula, formula (14) can be rear-

ranged to the form

$$W = \frac{N!}{\prod_m N_m!(m!)^{N_m}} \prod_m \left(N\frac{1}{(N)^m}\right)^{N_m}. \quad (15)$$

The conservation equations for the number of particles of both sorts are

$$\varphi_A = N_A - N_A^L - \frac{1}{2}\sum_m mN_m = 0,$$
$$\varphi_B = N_B - N_B^L - \frac{1}{2}\sum_m mN_m = 0. \quad (16)$$

Using the Lagrange method of multipliers, we construct a functional that will take into account Eqs. (16):

$$\Phi = \sum_m g_m N_m - kT\ln\left[\frac{N!}{\prod_m N_m!(m!)^{N_m}}\prod_m\left(\frac{1}{N^{m-1}}\right)^{N_m}\right] + \lambda_A\left(N_A - N_A^L - \frac{1}{2}\sum_m mN_m\right) + \lambda_B\left(N_B - N_B^L - \frac{1}{2}\sum_m mN_m\right). \quad (17)$$

Minimizing functional (17) yields the following expressions for the number of clusters and free particles of both sorts:

$$N_A^L = \exp\left[-\frac{\frac{z}{2}g_{AB}^L + \Omega(x_B^L)^2 + \lambda_A}{kT}\right], \quad (18)$$

$$N_B^L = \exp\left[-\frac{\frac{z}{2}g_{AB}^L + \Omega(x_A^L)^2 + \lambda_B}{kT}\right], \quad (19)$$

$$N_m = \frac{1}{m!N^{m-1}}\exp\left\{-\frac{g_m + \frac{m}{2}\lambda_A + \frac{m}{2}\lambda_B}{kT}\right\}, \quad (20)$$

where parameters λ_A and λ_B are determined from the solution of Eqs. (16).

LIQUIDUS CURVES

It is evident from Eqs. (16) that the total number of particles in the clusters N^S is

$$N^S = \sum_{m=2}^{\infty} mN_m. \quad (21)$$

To calculate the sum in (21), several assumptions on the free energy of m particle clusters should be made. We will assume that the clusters are strictly stoichiometric and their energy is proportional to the number of particles

$$g_m = \frac{z}{2}g_{AB}^S m. \quad (22)$$

Also, the clusters are assumed to be large enough that the surface energy in Eq. (22) is much smaller than the binding energy inside the clusters and is neglected. Substituting (20) and (22) into (21) yields

$$N^S = \exp\left\{-\frac{\frac{z}{2}g_{AB}^S + \lambda_A/2 + \lambda_B/2}{kT}\right\} \times\left[\exp\left(\frac{\exp\left\{-\frac{\frac{z}{2}g_{AB}^S + \lambda_A/2 + \lambda_B/2}{kT}\right\}}{N}\right) - 1\right]. \quad (23)$$

To construct the liquidus curve, it is necessary to specify the melting point, since the melt–crystal phase transition actually occurs in a certain temperature range that may be as wide as several tens of degrees.

We will assume that an [*AB*] crystal melts if the energy of interparticle interaction upon clustering changes by a certain value. It follows from (23) that the melt crystallization starts at a certain point where the expression $g_{AB}^S + \lambda_A/2 + \lambda_B/2 - kT\ln N$ takes a particular value. Thus, the equation

$$\frac{z}{2}g_{AB}^S + \lambda_A/2 + \lambda_B/2 - kT\ln N = \text{const} \quad (24)$$

is the liquidus equation for III–V compound semiconductors. This equation can be rewritten in a more familiar form. To do this, we express the parameters λ_A and λ_B from (18) and (19) as

$$\lambda_A = -\frac{z}{2}g_{AB}^L - \Omega(x_B^L)^2 + kT\ln N_A^L,$$
$$\lambda_B = -\frac{z}{2}g_{AB}^L - \Omega(x_A^L)^2 + kT\ln N_B^L. \quad (25)$$

The substitution of expressions (25) into (24) results in the most general liquidus equation

$$\frac{z}{2}(g_{AB}^S - g_{AB}^L) - \frac{1}{2}\Omega((x_A^L)^2 + (x_B^L)^2) - \frac{1}{2}kT\ln[X_A^L X_B^L] = \text{const}, \quad (26)$$

where $X_A^L = N_A^L/N$ and $X_B^L = N_B^L/N$ are the relative concentrations of free particles in the melts.

Let us demonstrate that, from (26), the liquidus equation, well known from the theory of quasi-chemical interaction [6, 7], can be derived. In the general case, due to the presence of the clusters, $x_A^L = X_A^L/(X_A^L + X_B^L) \neq X_A^L$, $x_B^L = X_B^L/(X_A^L + X_B^L) \neq X_B^L$, and $X_A^L + X_B^L \neq 1$. The first approximation stems from the fact that the fraction of clusters in the melt is small, which is true for many semiconductor melts [2]. If the fraction of clusters is small, $x_A^L \approx X_A^L$ and $x_B^L \approx X_B^L$; therefore, the liquidus equation takes the form

$$\Delta g_{AB}^F + \frac{1}{2}\Omega((x_A^L)^2 + (x_B^L)^2) + \frac{1}{2}kT\ln[x_A^L x_B^L] = \text{const}, \quad (27)$$

where $\Delta g_{AB}^F = -z(g_{AB}^S - g_{AB}^L)/2$ is the change in the energy of interaction between unlike particles upon the solid–liquid phase transition.

This change takes into account only the *A*–*B* interaction; however, in the melt, the *A*–*A* and *B*–*B* interactions, which depend on the parameter Ω, also take place. Hence, Δg_{AB}^F is not equal to the melting energy of the crystal [*AB*]: $\Delta g_{AB}^F \neq \Delta G_{AB}^{\text{melt}}$.

Equation (26) defines a family of curves according to the constant on its right. However, such an uncertainty can be avoided by using experimental data on melting semiconductor crystals, i.e., on melting points and stoichiometric compositions. The stoichiometric composition of III–V semiconductors is $x_A^L = x_B^L = 0.5$. The substitution of the melting temperature $T = T^F$ and the corresponding composition $x_A^L = 0.5$ into Eq. (27) makes it possible to find the constant on the right-hand side of (26).

Now initial liquidus equation (24) can be rewritten in the form

$$kT\ln[4x_A^L(1-x_A^L)] + (\Delta S^F + k\ln 4)(T^F - T) + \Omega[2(x_A^L)^2 - 2x_A^L + 1] - \Omega(T^F)/2 = 0, \quad (28)$$

which exactly coincides with the well-known liquidus equation for III–V semiconductors [7]. Thus, a family (27) of liquidus curves transforms into the equation emerging from the theory of quasi-chemical interaction if the cluster concentration is small.

SIMULATION OF LIQUIDUS AND AFTERMELTING CURVES IN GALLIUM ARSENIDE

Using the above-proposed model, we simulated cluster formation in the gallium arsenide melt. Both experimental liquidus curves [8] and data on the aftermelting phenomenon [2] were taken into account.

It is evident from Eqs. (7), (23), and (27) that the specific position of the aftermelting and liquidus curves depends on the values of *a*, *b*, ΔH^F, and ΔS^F. By varying these parameters, both experimental dependencies can be described satisfactorily.

Substituting Eqs. (25) into (23) yields the equation for the aftermelt curves

$$X^S = \frac{N^S}{N} = \exp\left\{-\frac{\Delta H^F - T\Delta S^F + \frac{a-bT}{2}((x_A^L)^2 + (x_B^L)^2) + \frac{kT}{2}\ln[X_A^L X_B^L]}{kT}\right\} \times \left[\exp\left(\exp\left\{-\frac{\Delta H^F - T\Delta S^F + \frac{a-bT}{2}((x_A^L)^2 + (x_B^L)^2) + \frac{kT}{2}\ln[X_A^L X_B^L]}{kT}\right\}\right) - 1\right]. \quad (29)$$

Equation (29) should be solved jointly with Eqs. (16), which are conveniently represented in terms of the relative values

$$X_A^{\text{total}} = X_A^L + \frac{1}{2}X^S, \qquad X_B^{\text{total}} = X_B^L + \frac{1}{2}X^S, \quad (30)$$

where X_A^{total} and X_B^{total} are the full particle concentrations in the melt: $X_A^{\text{total}} + X_B^{\text{total}} = 1$.

The simulation of the aftermelting curves can be simplified if it is remembered that gallium arsenide melts congruently. This means that the relative concentrations of the particles in the melt are equal, $x_A^L = x_B^L =$

Experimental and theoretical values of the enthalpy and entropy of melting for gallium arsenide

h, eV	*s*, k	T^F, K	T_{melt}, K [7]	E_{dis}, eV [9]	$\Delta H_{AB}^{\text{melt}}$, eV [1]
3.36	25.7	1504	1511	3.26	1.16

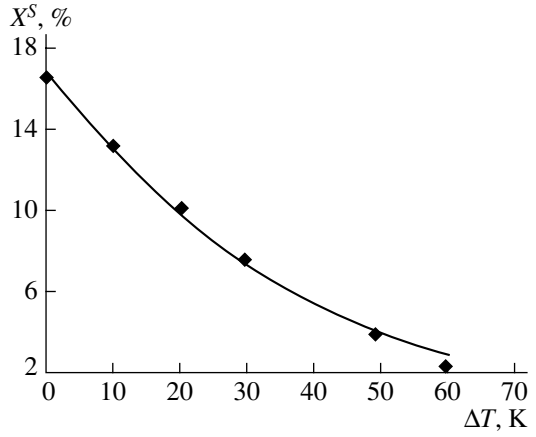


Fig. 1. Percentage of clusters in gallium arsenide X^S vs. $\Delta T = T - T^F$ (continuous curve). Data points are taken from [2].

0.5, and the total numbers of the particles are also equal, $X_A^{\text{total}} = X_B^{\text{total}} = 0.5$. Then, Eq. (29) can be transformed to

$$X^S = \frac{1}{2}(1 - X^S)$$

$$\times \exp\left\{-\frac{\left(\Delta H^F + \frac{a}{4}\right) - T\left(\Delta S^F + \frac{b}{4}\right)}{kT}\right\}\left[\exp\left(\frac{1}{2}(1 - X^S)\right.\right.$$
$$\left.\left.\times \exp\left\{-\frac{\left(\Delta H^F + \frac{a}{4}\right) - T\left(\Delta S^F + \frac{b}{4}\right)}{kT}\right\}\right) - 1\right]. \quad (31)$$

It follows from (31) that cluster formation during melting of the stoichiometric crystal can be characterized by some values of enthalpy h and entropy s:

$$h = \Delta H^F + a/4, \quad s = \Delta S^F + b/4. \quad (32)$$

Thus, to describe the aftermelting curves, it is necessary to vary only two parameters: h and s. For h = 3.36 eV and s = 25.7 k, experimental aftermelting curves are adequately described by Eq. (31) (Fig. 1).[1]

The simulation of the liquidus curves with regard for clustering was carried out using the expression resulting from relationships (27) and (30):

$$T(X_A^{\text{total}}) = \frac{a/2((1 - X_A^{\text{total}} - X^S(T))^2 + (X_A^{\text{total}} - X^S(T))^2) - (\text{const} - \Delta H^F)(1 - X^S(T))^2}{\Delta S^F - k\ln[(1 - X_A^{\text{total}} - X^S(T))(X_A^{\text{total}} - X^S(T))] + b/2((1 - X_A^{\text{total}} - X^S(T))^2 + (X_A^{\text{total}} - X^S(T))^2)},$$

$$X^S(T) = \exp\left\{-\frac{\text{const}}{kT}\right\} \times\left(\exp\left\{\exp\left\{-\frac{\text{const}}{kT}\right\}\right\} - 1\right). \quad (33)$$

The constant in Eq. (33) was determined from formula (27) and the stoichiometric crystal melting conditions $x_A^L = x_B^L = 0.5$ and $T = T^F$ with X^S obtained from Eq. (31).

The eventual simulation of the liquidus curve was carried out by fitting parameters a and b. In (32), the values of ΔH^F and ΔS^F were varied in such a way as to satisfy the values of h and s obtained. The parameters a, b, ΔH^F, and ΔS^F for gallium arsenide are a = –13.76 eV, $b = -8.32 \times 10^{-3}$ eV/K, ΔH^F = 6.8 eV, and ΔS^F = 99.4271 k. With such fitting, the aftermelt curve is described automatically at any values of the parameters a and b. The final liquidus curve for gallium arsenide is shown in Fig. 2.

HEAT OF MELTING

By definition, the heat of melting is the energy needed for a material to pass from the solid phase to the melt. Hence, the heat of melting per particle is written in the form

$$\Delta C_{AB}^{\text{melt}} = \frac{1}{2}(g_A^L + g_B^L - 2g_{AB}^S). \quad (34)$$

In our case, by the solid phase, we mean a set of clusters. Using (6) and (22), (34) can be recast as

$$\Delta C_{AB}^{\text{melt}} = (\Delta H^F - T\Delta S^F) + \frac{a - bT}{2}((x_A^L)^2 + (x_B^L)^2) + \frac{kT}{2}\ln[X_A^L X_B^L]. \quad (35)$$

From this expression, it follows that the heat of melting of a compound semiconductor depends not only on the binding energy between its particles but on the temperature and composition of the melt as well.

Since gallium arsenide melts congruently, expression (35) can be transformed into the form

$$\Delta C_{AB}^{\text{melt}} = h - Ts + kT\ln[1 - X^S], \quad (36)$$

[1] Both parameters are given per particle.

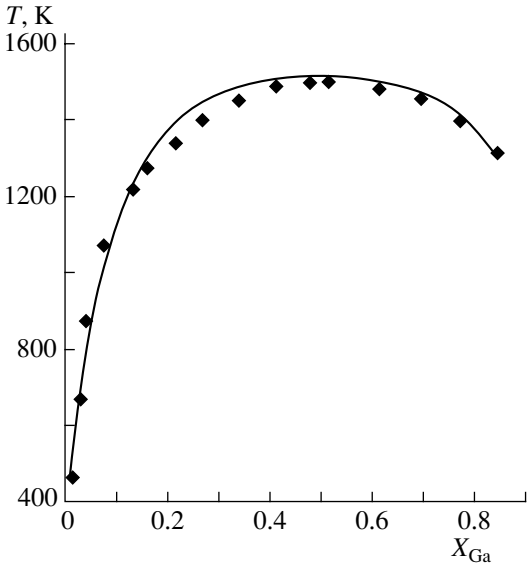


Fig. 2. Simulated liquidus curve for gallium arsenide (continuous curve). Data points are taken from [8].

where the values of h and s are determined by relationships (32).

Experiments on aftermelting in gallium arsenide [2] show that the relative concentration of the particles forming clusters at temperatures exceeding the melting point is less than 16%; therefore, when evaluating the entropy and the enthalpy of melting, one can neglect the logarithmic term. Thus, the parameters h and s obtained from the simulation of the aftermelting curves are nothing else than the enthalpy and the entropy of melting, respectively, of gallium arsenide.

The enthalpy and the entropy of melting of gallium arsenide obtained from simulating the liquidus and aftermelting curves are presented in the table. For comparison, the energy of dissociation E_{dis} and the heat of melting ΔH_{AB}^{melt} of gallium arsenide, as well as the melting point T^F calculated from the relationship between the enthalpy and entropy $h = T^F s$, are also presented. The enthalpy h calculated is almost three times higher than the experimental enthalpy value and approaches the energy of dissociation. The melting point obtained agrees well with the experimental data.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grants nos. 00-01-00209 and 00-01-00283).

REFERENCES

1. A. A. Aĭvazov, V. M. Glazov, and A. R. Regel', *Entropy of Melting of Metals and Semiconductors* (TsNII Élektronika, Moscow, 1978).
2. A. R. Regel' and V. M. Glazov, Élektron. Tekh., No. 9 (194), 7 (1984).
3. S. V. Bulyarskiĭ and V. V. Prikhod'ko, Pis'ma Zh. Tekh. Fiz. **25** (7), 33 (1999) [Tech. Phys. Lett. **25**, 263 (1999)].
4. P. E. L'vov and V. V. Svetukhin, Khim. Fiz. **18** (2), 93 (1999).
5. S. V. Bulyarskiĭ, V. V. Svetukhin, and P. E. L'vov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 385 (2000) [Semiconductors **34**, 371 (2000)].
6. I. Prigogine, *The Molecular Theory of Solutions* (North-Holland, Amsterdam, 1957; Metallurgiya, Moscow, 1990).
7. N. C. Casey, Jr. and M. B. Panish, *Heterostructure Lasers* (Academic, New York, 1978; Mir, Moscow, 1981), Vol. 2.
8. C. Hilsum and A. C. Rose-Innes, *Semiconducting III–V Compounds* (Pergamon, Oxford, 1961; Inostrannaya Literatura, Moscow, 1963).
9. W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* (Freeman, San Francisco, 1980; Mir, Moscow, 1983), Vol. 1.

Translated by M. Lebedev

Technical Physics, Vol. 46, No. 9, 2001, pp. 1082–1087. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 15–20.

GASES AND LIQUIDS

An Interferometric Technique for Measuring Gas Flow Velocity at Large Peclet Numbers

A. A. Pikulev

All-Russia Research Institute of Experimental Physics (VNIIEF), Russian Federal Nuclear Center, ul. Mira 37, Sarov, Nizhni Novgorod oblast, 607190 Russia

Received November 29, 2000

Abstract—A technique for determining the gas flow velocity at large Peclet numbers using an interferometric approach has been developed. It is shown that if a heat source is introduced in a uniform gas flow, areas under curves describing shifts of interference maxima are approximately equal to each other and tend to a certain limiting value. The limiting area is inversely proportional to the gas flow velocity and does not depend on the heat conduction and the type of flow. A theoretical expression has been derived that gives the relative error of velocity determinations by the method developed. *© 2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

At present, interferometric methods are widely used for investigations of thermophysical and gasdynamic parameters of gas flows. For example, in experiments [1, 2], the authors investigated a laser model using Michelson and Mach–Zehnder interferometers [3]. The laser model was a channel consisting of two sections with heat exchangers in between. The gas flowing through the channel was infused with uranium fission fragments. The interferometers were employed to measure the gas temperature and to investigate the thermal boundary layer contiguous to the heated channel wall and characteristics of the heat exchanger wake. In this last experiment, a thin Nichrom wire heated by an electric current was placed across the flow downstream of the heat exchanger. A schematic of the section with the installed wire is shown in Fig. 1. Study of the thermal wake of the heated wire makes it possible to determine the heat conduction in the gas as a function of the distance from the heat exchanger and, consequently, to identify the type of the flow and the size of convective cells forming in the heat exchanger wake [4, 5].

To process the results of such experiments, it is necessary to know the average gas flow velocity. In experiments [1, 2], the velocity was determined from the pressure differential. Such an approach demands calibration of the measuring instrument, and the measurement accuracy is not high. Therefore, it is desirable to carry out independent measurements of the velocity by other methods. Measurements of the velocity with Venturi and Pitot tubes [6] are difficult in gas-tight channels. Use of a pulsed heat-loss anemometer [7] specially designed for this purpose has two disadvantages: (1) it must be calibrated; and (2) the measurement results depend on the heat conduction of the gas, that is, on the type of flow.

In this work, it has been shown that at large Peclet numbers, the gas velocity can be determined on the basis of interferometric experiments with a heated wire placed across the stream. The proposed method has several advantages, namely (1) the measurement results do not depend on the heat conduction and the type of flow; (2) additional experiments are not necessary (the same interference pattern can be used that is taken in the study of the heat exchanger wake); and (3) calibration is unnecessary.

BASIC THERMODYNAMIC EQUATIONS

Let us consider the thermodynamic aspects of the gas flow. In the region considered, the flow is assumed to be isobaric and the stream lines are assumed to be parallel to the *x* axis. The equation of stationary transfer of the energy under constant pressure, ignoring the dissipative terms, has the form [8]

$$c_p\rho(\mathbf{u}, \nabla T) = \mathrm{div}(\mu\nabla T) + q, \qquad (1)$$

where ρ, $\mathbf{u}$, and T are the density, the velocity, and temperature of the gas, respectively; μ is the thermal con-

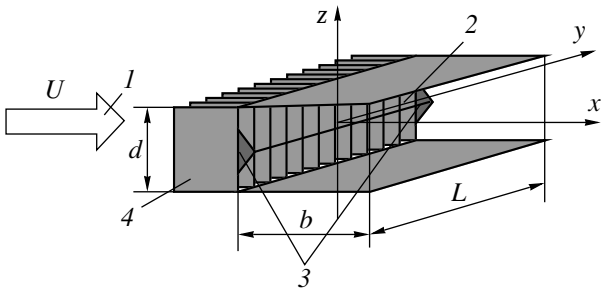


Fig. 1. Schematic of the gas section (the second heat exchanger is not shown): (*1*) gas flow direction, (*2*) wire, (*3*) wire holders, and (*4*) heat exchanger.

1063-7842/01/4609-1082$21.00 © 2001 MAIK "Nauka/Interperiodica"

ductivity; c_p is the specific heat of the gas at constant pressure; and q is the power of the volume heat source.

At the channel inlet, the gas has the velocity U_0, density ρ_0, and temperature T_0. Let U denote the only nonzero component of the gas velocity u_x. From the continuity equation [8] it follows that $\rho U = \rho_0 U_0$. For convenience we will use dimensionless variables. To this end, let us introduce a characteristic linear dimension, the distance h along which the temperature distribution will be studied. As the characteristic temperature and velocity, let us use T_0 and U_0, respectively, and represent the temperature in the form $T = T_0 + \Delta T$. Then we have $x = h\bar{x}$, $y = h\bar{y}$, $z = h\bar{z}$, and $\Delta T = T_0\overline{\Delta T}$. We assume the temperature variation to be small, that is, $|\Delta T| \ll T_0$. The energy equation in dimensionless form is

$$\frac{\partial\overline{\Delta T}}{\partial\bar{x}} = \frac{\overline{\nabla^2}\overline{\Delta T}}{\mathrm{Pe}} + \bar{q}, \quad \bar{q} = \frac{hq}{c_p\rho_0 U_0 T_0}, \qquad \mathrm{Pe} = \frac{c_p\rho_0 U_0 h}{\mu}, \tag{2}$$

where Pe is the Peclet number.

Below, unless otherwise indicated, we shall use the dimensionless variables and omit the upper bars.

Let us assume that the convective heat transfer along the x axis far exceeds the diffusive heat transfer, that is, $\mathrm{Pe} \gg 1$. In this case, the energy equation can be simplified:

$$\frac{\partial\Delta T}{\partial x} = \frac{1}{\mathrm{Pe}}\left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\Delta T + q. \tag{3}$$

Let a localized heat source of power W be located in the vicinity of the origin of the coordinates. Let us integrate Eq. (3) over a volume of $V = \{x \in [-a, a]; y, z \in (-\infty, +\infty)\}$ of the rectangular parallelepiped including the source region. We obtain

$$\int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} \Delta T dz = W. \tag{4}$$

Consequently, the flux of energy transferred by the gas through any plane $x = \mathrm{const} > 0$ is equal to the power of the heat source. It is obvious that the energy flux through some surface S that deviates negligibly from the plane $x = \mathrm{const} > 0$ is also approximately equal to W,

$$\iint_S \Delta T(\mathbf{i}, d\mathbf{s}) \cong W, \tag{5}$$

where $\mathbf{i}$ is the unit vector of the x axis and $d\mathbf{s}$ is the unit area of the surface S, whose orientation coincides with the positive direction of the normal.

Let us consider the following problem: a point heat source located at the origin of the coordinates is placed in a uniform gas flow. The power of the heat source is equal to W. The flow disturbance caused by the heat source is assumed to be negligible. Let us represent the thermal diffusivity of the gas in the form $\chi = \chi_l + \chi_t$, where χ_l and χ_t are the laminar and turbulent thermal diffusivities, respectively [9], and $\chi_t = \chi_t(x)$. This makes it possible to take into account the effect of small-scale convective mixing, for instance, of the heat exchanger wake. In a laminar flow, $\chi_t = 0$. Taking into account the definition of the Peclet number (2), we obtain

$$\frac{1}{\mathrm{Pe}} = \frac{1}{\mathrm{Pe}_l} + \frac{1}{\mathrm{Pe}_t(x)}, \tag{6}$$

where Pe_l and Pe_t are the laminar and turbulent Peclet numbers.

Perturbation of the temperature field ΔT satisfies two boundary conditions: (1) the condition before the heat source and (2) the condition at infinity $\Delta T(x < 0 \cup y, z \longrightarrow \pm\infty) = 0$. Note that in addition to these conditions, ΔT also satisfies the integral condition (5). The solution of Eq. (3) has the form of the function of a point source [10]:

$$\Delta T = \frac{p^2 W}{\pi}\sigma(x)e^{-p^2(y^2+z^2)}, \tag{7}$$

$$\frac{1}{p^2} = 4\int_0^x \frac{dx}{\mathrm{Pe}}, \tag{8}$$

where $\sigma(x)$ is the Heaviside function (11).

Let us also give a solution to Eq. (3) for a linear heat source with a uniform energy release along its length, when the source coincides with the y axis. In this case, the temperature does not depend on the y coordinate and the solution is

$$\Delta T = \frac{pw}{\sqrt{\pi}}\sigma(x)e^{-p^2z^2}, \tag{9}$$

where w is the power of the linear heat source per unit length.

MEASUREMENT OF THE GAS FLOW VELOCITY

Changes in the temperature of a medium, as a rule, are accompanied by changes in its refractive index, which can be detected using the interferometric technique. For an ideal gas under constant pressure, the refractive index is related to the temperature by a known formula [12] (at $|\Delta T| \ll 1$)

$$\Delta n = -(n_0 - 1)\Delta T. \tag{10}$$

Phase incursion of a ray of light in the test channel is

$$\varphi = k\int_0^L \Delta n dy = -k(n_0 - 1)\int_0^L \Delta T dy, \tag{11}$$

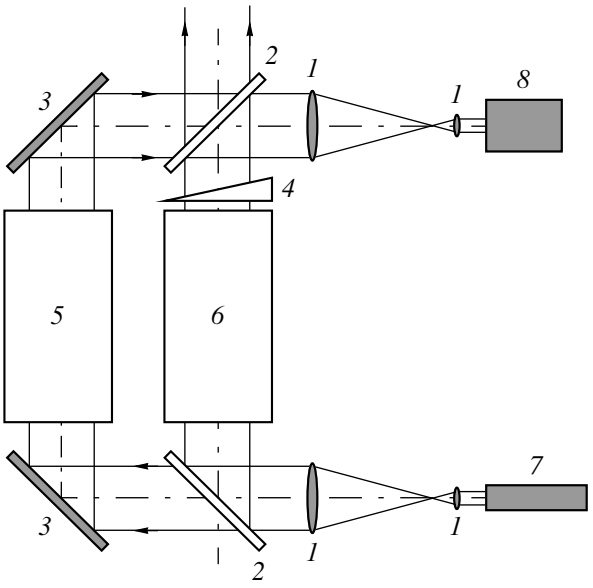


Fig. 2. Diagram of the nonequal-branch Mach–Zehnder interferometer: (*1*) focusing lenses, (*2*) semitransparent mirrors, (*3*) mirrors-reflectors, (*4*) optical wedge, (*5*) test channel, (*6*) reference channel, (*7*) probing laser, and (*8*) recording camera.

where k is the wave number and L is the length of the test channel.

The scheme of the experiment with the use of a Mach–Zehnder interferometer is shown in Fig. 2. The plane-parallel light beam of the probing laser is split in two; one beam passes through the test channel and the other, through the reference channel. For convenience, an optical wedge is inserted in the reference beam in the direction of gas flow (x axis). It is desirable to equalize the intensities of the light beams to achieve maximum contrast of the interference pattern. Positions of the interference maxima are defined by the following equation [3]:

$$\varphi(x_m(z), z) + 2\pi N x_m(z) + \varphi_0 = 2\pi m, \qquad (12)$$

where N is the number of fringes per unit length, φ_0 is the initial phase difference, and $x_m(z)$ is the position of the mth maximum.

When the gas is not heated, the interference maxima are evenly spaced straight lines parallel to the z axis. Their positions are given by the formula

$$2\pi N\Delta_m + \varphi_0 = 2\pi m, \quad \varphi_0 = 2\pi N\Delta_0, \qquad (13)$$

where Δ_m is the undisturbed position of the mth maximum.

At $N = 0$, we have a so-called infinite-width fringe. The problem associated with the use of an infinite-width fringe is the evaluation of the initial phase incursion φ_0, because it is hard to suggest an effective approach for determining it. Let $\xi_m(z)$ denote the shifts of the interference maxima from their undisturbed positions Δ_m; then, $\xi_m(z) = x_m(z) - \Delta_m$. An equation for the shifts of the maxima of fringes of finite width has the form

$$\varphi(x_m(z), z) = -2\pi N\xi_m(z). \qquad (14)$$

The main quantity characterizing the shifts of the interference fringes from their undisturbed positions is the area under the curve ξ_m. This area is

$$S_m = \int_{-\infty}^{+\infty} \xi_m dz. \qquad (15)$$

Integrating expression (14) with respect to the z coordinate between infinite limits and using formulas (11) and (15), we get

$$\int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} \Delta T(x_m(z), y, z)dz = \frac{2\pi NS_m}{k(n_0 - 1)}. \qquad (16)$$

The integration in (16) is performed over a surface of cylindrical shape. Comparing with (15), we find

$$2\pi NS_m \cong 2\pi NS = k(n_0 - 1)W. \qquad (17)$$

Now we have the following result: all areas S_m are approximately equal to each other and tend to S. The farther downstream of the heat source the interference maximum is located, the more precise expression (17) is. This is associated with an increase in the local Peclet number downstream and weakening of the role of heat conduction in heat transfer. Hence, the law of equality of the areas S_m is an interferometric expression for the law of conservation of the heat energy at large Peclet numbers. This expression does not depend on the shape of the heat source and the type of flow. The gas velocity is determined by the formula (in dimensional variables)

$$U_0 \cong \frac{k(n_0 - 1)W}{2\pi c_p \rho_0 NST_0}. \qquad (18)$$

Let us determine the phase incursion caused by a point heat source. We assume that the laser beam probes the entire region of the thermal wake of the heat source along the y and z axes. From formulas (8) and (11), it follows that

$$\varphi = -\frac{(n_0 - 1)kW}{\sqrt{\pi}} p(x)\exp(-p^2 z^2). \qquad (19)$$

The shifts of interference maxima of finite-width fringes from their undisturbed positions are calculated using formula (14), from which, using (19), we get

$$\xi_m = \frac{p(x_m)S}{\sqrt{\pi}}\exp(-p(x_m)^2 z^2). \qquad (20)$$

Formulas (19) and (20) are valid not only for a point source but for any linear source located along the y axis with an arbitrary distribution of the heat release along the source length. In this case, as before, the entire

wake of the source must be probed by the laser beam in the planes normal to the x axis, and W is the total power of the heat source.

Equation (20) is transcendental; that is, its solution cannot be represented in the form of elementary functions. For simplicity, let us assume that the relation $\xi_m \ll \Delta_m$ is satisfied and seek a solution to (20) by the successive approximations method. As an initial approximation, we use the following expression:

$$\xi_m^{(0)} = \frac{p(\Delta_m)S}{\sqrt{\pi}}\exp(-p(\Delta_m)^2 z^2). \quad (21)$$

Integrating (21) with respect to the z coordinate gives

$$S_m^{(0)} = \int_{-\infty}^{+\infty} \xi_m^{(0)} dz = S. \quad (22)$$

Consequently, all areas $S_m^{(0)}$ are equal to each other and do not depend on the thermal conductivity, which coincides with the above result expressed by (17). Analysis of Eq. 20 shows that the areas S_m satisfy the following conditions:

$$0 < S_1 < S_2 < \ldots < S_m < \ldots < S, \quad \lim_{m\to\infty} S_m = S. \quad (23)$$

From inequality (23), it is seen that the use of formula (18) gives a somewhat overestimated gas velocity.

To clarify the issue of the accuracy of formula (18), let us expand the right-hand side of expression (20) into a Taylor series with respect to ξ_m, retaining the terms up to the second order of smallness inclusively. After integrating and collecting like terms, we obtain

$$S_m^{(2)} = S\left\{1 + \frac{p_x}{2\sqrt{2\pi}}S + \frac{pp_{xx} - p_x^2}{3\pi\sqrt{3}}S^2\right\}_{x=\Delta_m}, \quad \frac{p(\Delta_m)S}{\Delta_m\sqrt{\pi}} \ll 1. \quad (24)$$

In the case when the Peclet number is constant along the gas stream, (24) takes the form

$$S_m^{(2)} = S\left\{1 - \frac{\alpha_m}{4\sqrt{2}} + \frac{\alpha_m^2}{6\sqrt{3}}\right\}, \quad \alpha_m = \frac{S}{2\Delta_m}\sqrt{\frac{\mathrm{Pe}}{\pi\Delta_m}} \ll 1. \quad (25)$$

Thus, the accuracy of the law of areas is governed by the second terms in expansions (24) or (25). The accuracy increases rapidly with distance downstream from the heat source; therefore, practically under any experimental conditions it is possible to find fringes for which the law of conservation of areas is valid to a high degree of accuracy.

The results obtained above are illustrated in Figs. 3–6. The interferograms were obtained under the following conditions: the working gas was air at atmospheric conditions, the flow velocity was $U_0 = 5$ m/s, the initial phase was $\varphi_0 = 0$, the source power was $W = 6$ W, and the laser wavelength was $\lambda = 632$ nm. For a characteristic length $h = 1$ cm under these conditions, the Peclet number is Pe = 2704; that is, the restriction imposed on the magnitude of Peclet number is satisfied. As shown in [4, 5], the turbulent Peclet number in the heat exchanger wake may be several tens of times less than the laminar Peclet number. This fact should be taken into account when employing the above model of the thermal wake of a linear source. Figure 3 displays an infinite-fringe interferogram of the flow with a point (linear) heat source at $\chi_t = 0$. The effect of the convective wake of the heat exchanger under the same tuning of the interferometer is illustrated in Fig. 4. Intensification of the heat exchange in the wake was simulated by inserting immediately behind the heat exchanger a region of length x_t, where $\chi_t > 0$. At $x > x_t$ the turbulent thermal conductivity was assumed to be zero. The lengths x_t and magnitudes of χ_t in the wake of the plate heat exchanger under different conditions can be found in [4, 5]. In this study, we accepted $x_t = 1$ cm and $\chi_t = 5\chi_l$. Comparison between Figs. 3 and 4 shows that the pres-

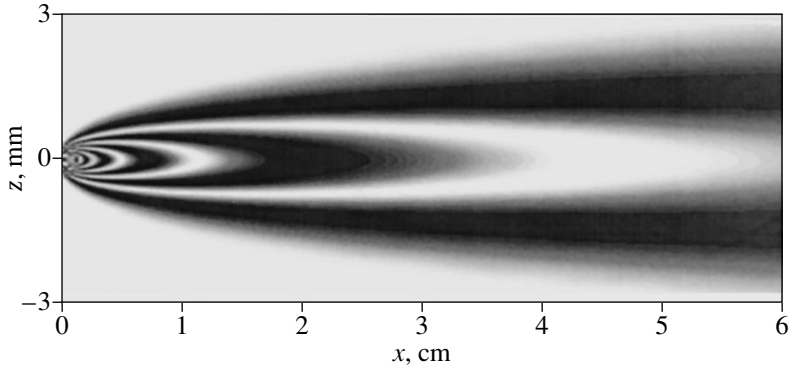


Fig. 3. An infinite-fringe interferogram of the gas flow with a linear heat source at $\chi_t = 0$.

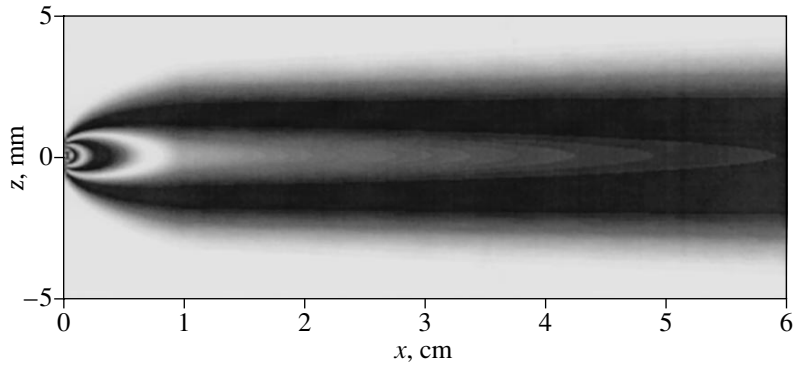


Fig. 4. Same as in Fig. 3 at $\chi_t = 5\chi_l$ and $x_t = 1$ cm.

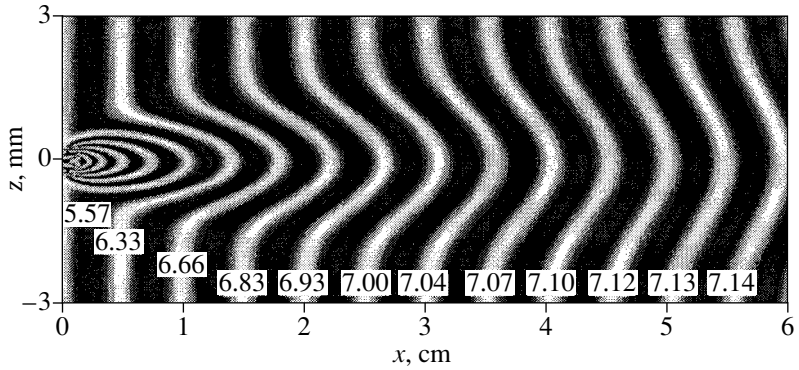


Fig. 5. A finite-fringe interferogram of the gas flow with a linear heat source at $\chi_t = 0$.

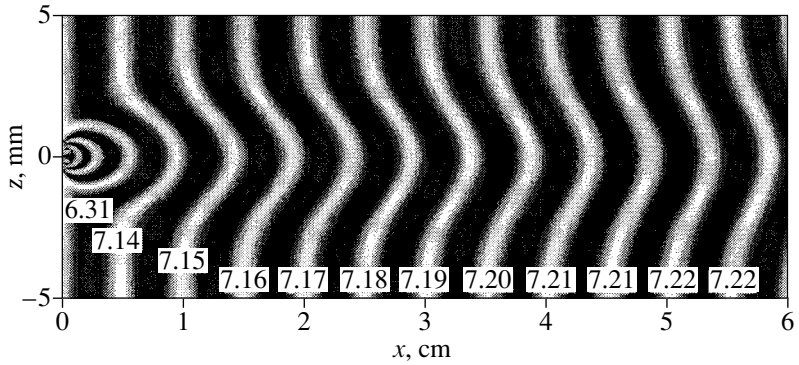


Fig. 6. Same as in Fig. 5 at $\chi_t = 5\chi_l$ and $x_t = 1$ cm.

ence of the heat exchanger causes faster expansion of the thermal wake of the source in the zone of convective mixing.

Finite-fringe interferograms in Figs. 5 and 6 were obtained for $N = 2$ cm^{-1}, other experimental conditions being the same as those in Figs. 3 and 4. Beside each interference maximum the areas S_m are given in units of mm^2. The areas were obtained by numerical integration of Eq. 20 using the successive approximation procedure. The limiting values of the areas in the both case are $S = 7.3$ mm^2. Comparison between Figs. 5 and 6 shows that the presence of the zone of intensive mixing behind the heat exchanger (Fig. 6) makes the approach of the areas S_m to the limiting value S faster than in the case of its absence (Fig. 5). At $m > 4$ the relative error $\varepsilon_m = (S - S_m)/S$ does not exceed 4% and is quite adequately described by the function

$$0 < \varepsilon_m < \varepsilon_m^{(1)} = \frac{S - S_m^{(1)}}{S} = \frac{\alpha_m}{4\sqrt{2}}. \qquad (26)$$

The magnitude of error $\varepsilon_m^{(1)}$ found under the

assumption of laminarity of the flow at Pe = 0 turned out to be maximal; therefore, it can be used as an estimation of the accuracy of velocity determinations for any type of flow.

It was assumed above that nonuniformity of the refractive index in the test channel only causes phase changes in the probing laser beam and does not influence its rectilinear propagation. However, if gas density variations are large enough, the light rays are deflected from their initial direction toward a denser medium. This leads to an increase of S_m areas for fringes located immediately downstream of the wire and violation of the law expressed by (23). This fact should be taken into account when carrying out experiments.

To summarize, let us describe the procedure of velocity measurement.

(1) In the course of an interferometric experiment, the power of a heat source W and the gas temperature T_0 in front of the wire are measured.

(2) In the interferogram, the quantity of undisturbed fringes per unit length N is determined and the areas S_m under the curves describing shifts of the interference maxima are calculated.

(3) Values of the velocities U_m corresponding to maxima m are calculated using formula (18).

(4) From the set of U_m, the magnitude of the velocity is selected which corresponds to the minimal sum of the relative error of measurement of the areas S_m and $\varepsilon_m^{(1)}$ (see (26)).

ACKNOWLEDGMENTS

The author expresses his gratitude to L.V. L'vov for helpful discussions.

REFERENCES

1. V. V. Borovkov, B. V. Lazhintsev, S. P. Mel'nikov, *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **54**, 2009 (1990).
2. V. V. Borovkov, B. V. Lazhintsev, V. A. Nor-Arevyan, *et al.*, Kvantovaya Élektron. (Moscow) **22**, 1187 (1995).
3. A. Sommerfeld, *Vorlesungen über theoretische Physik*, Bd. 4: *Optik* (Dieterich, Wiesbaden, 1950; Inostrannaya Literatura, Moscow, 1953).
4. F. S. Milos and A. Acrivos, Phys. Fluids **29**, 1353 (1986).
5. F. S. Milos, A. Acrivos, and J. Kim, Phys. Fluids **30**, 7 (1987).
6. G. K. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge Univ. Press, Cambridge, 1967; Mir, Moscow, 1973).
7. V. M. Kul'gavchuk and V. Yu. Mat'ev, in *Proceedings of the 2nd International Conference on Physics of Nuclearly Excited Plasma and Problems of Nuclear-Pumped Lasers, Arzamas-16, 1995*, Vol. 2, p. 93.
8. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1988; Pergamon, New York, 1987).
9. L. G. Loĭtsyanskiĭ, *Fluid and Gas Mechanics* (Nauka, Moscow, 1973).
10. S. L. Sobolev, *Equations of Mathematical Physics* (Nauka, Moscow, 1992).
11. A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis* (Nauka, Moscow, 1989).
12. *Tables of Physical Quantities: A Handbook*, Ed. by I. K. Kikoin (Atomizdat, Moscow, 1976).

Translated by N. Mende

Technical Physics, Vol. 46, No. 9, 2001, pp. 1088–1092. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 21–25.

GAS DISCHARGES, PLASMA

A Spherical Probe with a Time-Varying Potential in a Stationary Collisional Plasma

A. V. Kashevarov

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

e-mail: sasha@kura.aerocentr.msk.su

Received March 31, 2000; in final form, October 10, 2000

Abstract—The problem of a spherical probe in a stationary, weakly ionized, collisional plasma is studied analytically under the assumption of a slight harmonic temporal modulation of the constant probe potential, which is equal to the potential of the environment. The solution is obtained in two limiting cases of a thick and a thin space-charge sheath. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The diagnostic data from probe measurements are fairly difficult to interpret when the plasma is collision-dominated, i.e., when the characteristic mean free path of the plasma particles is small compared to the characteristic probe size, $\lambda \ll R$, so that the plasma can be regarded as a continuous medium. For a long time, research effort has been focused on the interpretation of data from probes operating in steady modes in both stationary and flowing plasmas (see, e.g., [1–4]), whereas a more complicated theory of unsteady operating modes of the probes is still far from being completely elaborated.

The probe operation in unsteady modes was investigated only in a few papers, which can be divided into two groups. The first group involves papers aimed at studying the behavior of the probe current during the evolution of the probe from one steady state to another because of rapid change in the applied potential. For example, in my recent paper [5], an analytic solution to the problem of the evolution of the current to a spherical probe in a stationary, weakly ionized, high-density plasma was obtained in the limiting case of a thick space-charge sheath when a small potential difference between the probe and the plasma drops to zero in a jumplike manner.

The papers of the second group [6–8] were intended to determine the impedance of the probes with modulated potentials. A slight sinusoidal temporal modulation of the constant potential of a probe substantially extends the possibilities of the probe diagnostics. In principle, the data from probes with modulated potentials make it possible not only to measure the electron temperature and density and the plasma space potential (which are all traditionally obtained from the steady-state current–voltage characteristic of the probe), but also to estimate other important plasma parameters.

Thus, Baksht [6] arrived at the conclusion that the ion diffusion coefficient can be determined from the measured impedance of a spherical probe in a stationary plasma. In that paper, an analysis was carried out for a thin collisionless space-charge sheath around the probe and the Debye screening radius was assumed to satisfy the condition $\lambda_D \ll \lambda \ll R$.

The objective of the present paper is to analytically investigate the impedance of a spherical probe in two of the other possible characteristic operating modes in a continuous medium, specifically, in the limiting cases of a thin ($\lambda \ll \lambda_D \ll R$) and thick ($\lambda \ll R \ll \lambda_D$) collisional space-charge sheath.

FORMULATION OF THE PROBLEM

We treat the problem with the following simple formulation. We consider a spherical probe in a stationary, weakly ionized, high-density plasma consisting of neutral particles, electrons, and singly charged positive ions and neglect chemical reactions between the plasma components. Such a plasma is thermally equilibrium and remains unchanged in transport properties. Under these conditions, the basic set of equations for the probe [1] can be generalized to describe unsteady operating modes as follows:

$$\beta^{-1}\partial n_+/\partial t - \nabla(\partial n_+ + n_+\nabla\psi) = 0, \quad (1)$$

$$\partial n_-/\partial t - \nabla(\nabla n_- - n_-\nabla\psi) = 0, \quad (2)$$

$$\alpha^2\nabla\psi = n_- - n_+. \quad (3)$$

Equations (1)–(3) are written in dimensionless form and describe non-steady-state diffusion of charged particles in an electric field. Here, the ratio of the diffusion coefficient of the ions to that of the electrons, $\beta = D_+/D_-$, is assumed to be much smaller than unity, $\beta \ll 1$; n_+ and n_- are the ion and electron number densities normalized to the charged-particle density at infinity, N_∞; and the dimensionless electric potential ψ is related to the dimensional potential φ by $\psi = e\varphi/kT_-$, where e is the

charge of an electron, k is Boltzmann constant, T_- is the electron temperature, and φ is the difference between the probe and plasma potentials. The radial coordinate r is in units of the probe radius R, and the time t is in units of the ratio R^2/D_-. The dimensionless parameter $\alpha = \lambda_D/R$ characterizes the thickness of the space-charge sheath around the probe.

Far from the probe (at $r \longrightarrow \infty$), Eqs. (1)–(3) are supplemented with the boundary conditions

$$n_+(\infty) = n_-(\infty) = 1, \quad \psi(\infty) = 0. \tag{4}$$

At the probe surface ($r = 1$), the charged particles are assumed to be completely absorbed,

$$n_+(1) = n_-(1) = 0. \tag{5}$$

and the probe potential is specified as

$$\psi(1) = \psi_p + a\sin\omega t, \tag{6}$$

where ψ_p and a are constants ($a \ll 1$).

Let the modulation frequency ω be low enough to guarantee the satisfaction of the underlying assumptions of our problem, in particular, the validity of the Einstein relationship between the diffusion coefficients and the mobilities of charged particles and the assumption that the plasma is in thermal equilibrium (the latter allows us to eliminate the electron energy equation from the mathematical formulation of the problem).

The task now is to calculate the electron (I_-) and ion (I_+) currents to the probe as functions of the modulation frequency ω. The dimensionless current $I_\mp$ is normalized to $4\pi eN_\infty RD_\mp$ and is defined as

$$I_\mp = dn_\mp/dr|_{r=1}. \tag{7}$$

With allowance for the fact that the currents I_- and I_+ are oppositely directed, the total dimensionless current I_Σ to the probe is equal to

$$I_\Sigma = I_- - \beta I_+.$$

The time-dependent problem (1)–(7) can be solved analytically for $\psi_p = 0$, i.e., for the case in which the probe potential oscillates with a small amplitude about the plasma potential. Regardless of the value of α, the steady solution ($\omega = 0$) has the form $\psi_0(r) = 0$, which corresponds to

$$n_{0\mp}(r) = 1 - r^{-1}. \tag{8}$$

Under these conditions, no charge separation occurs. Consequently, the cases of thick and thin space-charge sheaths are merely formal limits, because, in reality, no sheath forms around a probe operating in a steady mode. However, for a probe with a modulated potential, we may again speak of an unsteady space-charge sheath (either thick or thin).

SOLUTION FOR $\alpha \longrightarrow \infty$

For $\alpha \longrightarrow \infty$, Eq. (3) degenerates into the Laplace equation $\Delta\psi = 0$. The solution to the Laplace equation that satisfies the boundary conditions (4) and (6) is $\psi = ar^{-1}\sin\omega t$, or, in complex form,

$$\psi = ar^{-1}\exp(i\omega t). \tag{9}$$

In this limit, Eqs. (1) and (2) are independent of one another, so we can consider only one of them, e.g., Eq. (2). We represent the electron number density $n_-(r, t)$ as

$$n_-(r,t) = n_0(r) + \tilde{n}_-(r,t) = n_0(r) + \sum_{m=1}^{\infty} n_m(r)\exp(im\omega t). \tag{10}$$

Substituting expressions (8)–(10) into Eq. (2) and singling out the coefficients of the fundamental harmonic of ω, we arrive at the following ordinary differential equation for $n_1(r)$:

$$n_1'' + 2r^{-1}n_1' - i\omega n_1 = -ar^{-4}. \tag{11}$$

The boundary conditions on Eq. (11) follow from conditions (4) and (5): $n_1(1) = n_1(\infty) = 0$. The solution to Eq. (11) can be found in accordance with [9]:

$$\begin{aligned} n_1(r) = {} & -2^{-1}ar^{-2} + 2^{-1}ar^{-1}\exp[-\sqrt{i\omega}(r-1)] \\ & \times\{1 + 2^{-1}\sqrt{i\omega}[\exp(\sqrt{i\omega})E_1(\sqrt{i\omega}) \\ & - \exp(-\sqrt{i\omega})E_1(-\sqrt{i\omega})]\} - 4^{-1}ar^{-1}\sqrt{i\omega} \\ & \times[\exp(r\sqrt{i\omega})E_1(r\sqrt{i\omega}) - \exp(-r\sqrt{i\omega})E_1(-r\sqrt{i\omega})], \end{aligned} \tag{12}$$

where

$$E_1(z) = \int_\infty^z \frac{e^{-u}}{u}du$$

is the integral exponent [10]. Differentiating solution (12) with respect to r at $r = 1$ yields the complex amplitude of the electron current associated with the fundamental harmonic of the electron density oscillations:

$$I_1 = 2^{-1}a - 2^{-1}ai\omega\exp(i\omega)E_1(\sqrt{i\omega}). \tag{13}$$

Clearly, the amplitude I_2 of the electron current associated with the second harmonic of the electron density oscillations is on the order of $\sim a^2$, so that for $a \ll 1$, this harmonic, as well as all higher harmonics, can be neglected.

It is convenient to represent the complex quantity I_1 as

$$I_1 = aM(\omega)\exp[i\Phi(\omega)]. \tag{14}$$

The absolute value and the phase shift of the time-varying current I_1 are related to the impedance $Z(\omega)$ of

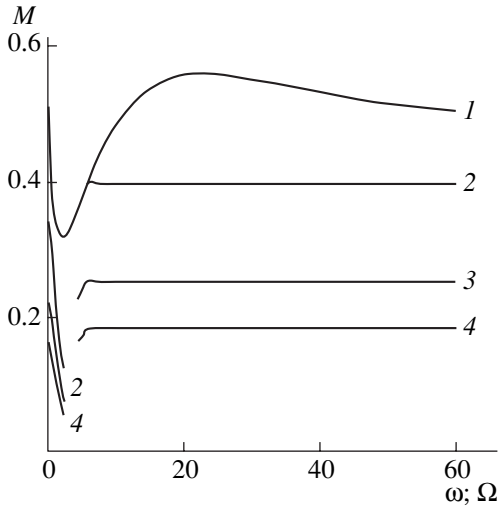


Fig. 1. Frequency dependence of the amplitude of the time-varying electron current for (*1*) $\alpha \longrightarrow \infty$, (*2*) $\alpha = 0.1$, (*3*) $\alpha = 0.01$, and (*4*) $\alpha = 0.001$. Curve *1* represents the $M(\omega)$ dependence, and curves *2–4* are for the $M(\Omega)$ dependence.

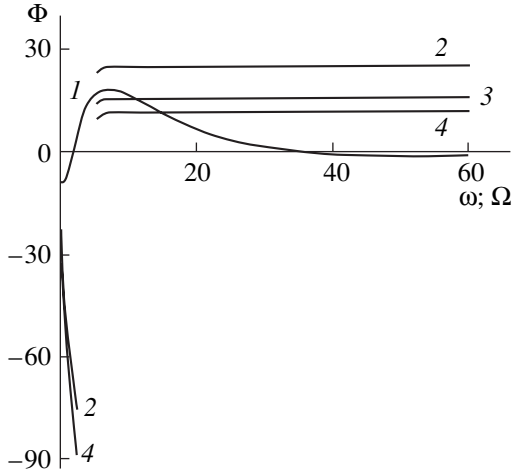


Fig. 2. Frequency dependence of the phase shift of the time-varying electron current. Curves *1–4* indicate the same dependences as in Fig. 1.

the probe by

$$Z = M^{-1}\exp(-i\Phi).$$

We thus arrive at the final expression for the sought-for electron current to the probe:

$$\begin{aligned} I_- &= I_0 + \mathrm{Im}[I_1\exp(i\omega t)] \\ &= 1 + aM(\omega)\sin[\omega t + \Phi(\omega)]. \end{aligned} \tag{15}$$

Figures 1 and 2 illustrate the time evolutions $M(\omega)$ and $\Phi(\omega)$, which were determined from formula (13) with the help of the tables presented in [10].

For $z \ll 1$, we have $E_1(z) \cong -\gamma - \ln z$, where $\gamma = 0.5772\ldots$ is Euler's constant. For $z \longrightarrow \infty$, the corresponding asymptotic expression is $E_1(z) \sim z^{-1}\exp(-z)$ [10]. Taking into account the identity $\sqrt{i} = (1 + i)/\sqrt{2}$ and using formula (13), we can obtain explicit expressions for $M(\omega)$ and $\Phi(\omega)$. For $\omega \ll 1$, we have

$$M(\omega) = (1 - \omega\pi/4)/2, \quad \Phi(\omega) = \omega(\gamma + 2^{-1}\ln\omega).$$

For $\omega \longrightarrow \infty$, we have

$$M(\omega) = 2^{-1}[1 - \sqrt{\omega}\exp(-\sqrt{\omega/2})\sin(\sqrt{\omega/2} + \pi/4)],$$

$$\Phi(\omega) = \sqrt{\omega}\exp(-\sqrt{\omega/2})\sin(\sqrt{\omega/2} + \pi/4).$$

Clearly, the ion current I_+ is also described by expression (15), but with $-a$ in place of a and $M(\omega)$ and $\Phi(\omega)$ in place of $M(\omega/\beta)$ and $\Phi(\omega/\beta)$. Since the dimensionless currents I_+ and I_- are of the same order of magnitude, the total current is approximately equal to $I_\Sigma \approx I_-$, provided that $\beta \ll 1$.

SOLUTION FOR $\alpha \longrightarrow 0$

In this limit, we must solve the full set of Eqs. (1)–(3). However, for $\beta \longrightarrow 0$, Eq. (1) degenerates into the equation $\partial n_+/\partial t = 0$, which indicates that the ion distribution in space is steady-state and thus can be described by the time-independent expression (8). For this reason, we can only consider Eqs. (2) and (3).

We again represent the electron number density $n_-(r, t)$ by formula (10) and substitute it into Eqs. (2) and (3). In this way, the potential ψ should also be represented in a form similar to expression (10):

$$\psi(r, t) = \sum_{m=1}^{\infty} \psi_m(r)\exp(im\omega t).$$

As a result, we arrive at the equations

$$\xi^4 n_1'' - \xi^4 E_1 + (1 - \xi)\xi^4 E_1' = i\omega n_1, \tag{16}$$

$$\alpha^2\xi^4 E_1' = -n_1, \tag{17}$$

where $\xi = 1/r$ and $E_1 = -d\psi_1/dr$. Equation (17) implies that for $\alpha \longrightarrow 0$, we may set $n_1 \cong 0$ over almost the entire space except for a thin sheath near the probe surface ($\xi \cong 1$). According to Eq. (16), the quasineutral plasma region is described by the equation

$$(1 - \xi)E_1' - E_1 = 0,$$

whose solution is

$$E_1 = C/(1 - \xi), \tag{18}$$

where C is an integration constant. Taking into account the boundary conditions (4), we can obtain from solution (18) the amplitude ψ_1 of the fundamental harmonic

of the probe potential:

$$\psi_1 = C\ln(1-\xi). \tag{19}$$

We can see that $\psi_1 \longrightarrow \infty$ as $\xi \longrightarrow 1$, which contradicts boundary condition (6). This indicates that, near the probe surface, the product $\alpha^2 E_1'$ remains finite as $\alpha \longrightarrow 0$; in other words, in a thin space-charge sheath around the probe, we have $n_1 \neq 0$. An unsteady space-charge sheath can be analyzed in essentially the same way as for a steady sheath in the classical paper by Cohen [11].

We substitute Eq. (17) into Eq. (16) and transform to new variables in the resulting equation:

$$\zeta = \alpha^{-2/3}(1-\xi), \quad E_1(\xi) = \alpha^{-2/3}F_1(\zeta). \tag{20}$$

For $\alpha \longrightarrow 0$, we neglect terms on the order of $O(\alpha^{2/3})$ to obtain

$$F_1''' = (\zeta + i\Omega)F_1' + F_1. \tag{21}$$

Here, we keep the term with $\Omega = \alpha^{4/3}\omega$, because, for $\alpha \longrightarrow 0$, we have $\Omega = O(1)$ as $\omega \longrightarrow \infty$. For $\zeta \longrightarrow \infty$, the solution to Eq. (21) should approach solution (18) for the quasineutral plasma region. In other words, in terms of variables (20), we have

$$F_1(\zeta) \longrightarrow C/\zeta, \quad \zeta \longrightarrow \infty. \tag{22}$$

Integration of Eq. (21) gives [9]

$$F_1'' - (\zeta + i\Omega)F_1 = -C. \tag{23}$$

The solution to Eq. (23) that satisfies condition (22) has the form

$$F_1 = A\mathrm{Ai}(\zeta + i\Omega) + \pi C\mathrm{Gi}(\zeta + i\Omega). \tag{24}$$

Here, A is an integration constant and Ai and Gi are Airy functions [12]; moreover, the function Gi(z) satisfies the equation $y'' - zy = -\pi^{-1}$. The constant A can be expressed in terms of C with the help of condition (5), which takes the form $n_1|_{\xi=1} = 0$, or with the help of Eq. (17) with transformation (20) such that $F_1'(0) = 0$. As a result, from solution (24), we obtain

$$A = -\pi C\mathrm{Gi}'(i\Omega)/\mathrm{Ai}'(i\Omega). \tag{25}$$

For $\Omega = 0$ (i.e., for low modulation frequencies $\omega \sim 1$), we have $\mathrm{Gi}'(0) = -3^{-1/2}\mathrm{Ai}'(0)$ and $A = 3^{-1/2}\pi C$.

At the probe surface, we have $\psi_1|_{\xi=1} = a$; on the other hand, $\psi_1(1) = \int_0^1 E_1(\xi)d\xi$. Applying transformation (20) to the integral in the latter expression gives

$$a = C\ln\alpha^{2/3} + C\ln\zeta_* + A\int_{\zeta_*}^{0}\mathrm{Ai}(\zeta + i\Omega)d\zeta + \pi C\int_{\zeta_*}^{0}\mathrm{Gi}(\zeta + i\Omega)d\zeta, \tag{26}$$

where ζ_* is the coordinate of the conditional boundary of the space-charge sheath. The first two terms on the right-hand side of formula (26) were obtained by applying transformation (20) to expression (19); they determine the potential at the boundary of the quasineutral plasma region. The last two terms describe the voltage drop across the space-charge sheath near the probe surface.

For $\zeta_* \longrightarrow \infty$, we can use the familiar asymptotic formula [13]

$$\int_0^{\zeta_*}\mathrm{Gi}(\zeta)d\zeta \sim \frac{2\gamma + \ln 3}{3\pi} + \frac{1}{\pi}\ln\zeta_* \tag{27}$$

and the relationship [12]

$$\int_0^{\zeta_*}\mathrm{Ai}(\zeta)d\zeta \longrightarrow 1/3. \tag{28}$$

We can show that formulas (27) and (28) are also valid for the integrals in expression (26) (see [14]). We insert formulas (25), (27), and (28) into expression (26) to find the constant C:

$$C = 3a/[2\ln\alpha + \pi\mathrm{Gi}'(i\Omega)/\mathrm{Ai}'i\Omega - 2\gamma - \ln 3].$$

As a result, the complex amplitude $I_1 = dn_1(\zeta)/d\zeta|_{\zeta=0} = F_1''|_{\zeta=0}$ of the electron current associated with the fundamental harmonic of the electron density oscillations is equal to

$$I_1(\Omega) = -C + i\Omega A\mathrm{Ai}(i\Omega) + i\Omega\pi C\mathrm{Gi}(i\Omega). \tag{29}$$

For $\Omega = 0$, we have $I_1 = -C$; moreover, we can see that the amplitude I_1 is real:

$$I_1(0) = 3a/(\ln\alpha^{-2} + 3^{-1/2}\pi + 2\gamma + \ln 3). \tag{30}$$

Note that the ratio $I_1(0)/a$ is the slope of the steady-state current–voltage (I–V) characteristic of the probe for $\psi_p = 0$. The question of the slope of the I–V characteristic of a probe whose potential is equal to the potential of the environment was studied by Benilov and Tirskiĭ [15] in connection with the possibility of determining the electron temperature. In particular, for an isothermal plasma, they obtained the relationship

$$dI_-(\psi)/d\psi|_{\psi=0} = (1.587 + 3^{-1}\ln\alpha^{-2})^{-1}. \tag{31}$$

According to the analytic solution (30), the first term in parentheses in expression (31) is approximately equal to $(2\gamma + \ln 3 + 3^{-1/2}\pi)/3 \approx 1.356$.

For $\Omega \neq 0$, it is difficult to calculate the amplitude $I_1(\Omega)$ from formula (29) because the tables of the Airy functions in the range of complex arguments are incomplete or because the numerical algorithms for calculating these functions are lacking. The calculations were carried out with the help of the four-digit tables of the Airy functions Ai(z) and Bi(z) and their derivatives Ai'(z) and Bi'(z) [16]. The Airy function Gi(z) and its derivative Gi'(z) were calculated from the formulas [12]

$$\mathrm{Gi}(z) = \frac{1}{3}\mathrm{Bi}(z) + \mathrm{Ai}(z)\int_0^z \mathrm{Bi}(u)du - \mathrm{Bi}(z)\int_0^z \mathrm{Ai}(u)du,$$

$$\mathrm{Gi}'(z) = \frac{1}{3}\mathrm{Bi}'(z) + \mathrm{Ai}'(z)\int_0^z \mathrm{Bi}(u)du - \mathrm{Bi}'(z)\int_0^z \mathrm{Ai}(u)du.$$

Unfortunately, the tables in [16] do not contain the values of the functions Ai($i\Omega$) and Bi($i\Omega$)and their derivatives in the range $\Omega > 2.4$. For this reason, the dependence $I_1(\Omega)$ for $\Omega \gg 1$ was calculated from the large argument asymptotic expansions of the Airy functions. Note that Gi($i\Omega$) ~ Bi($i\Omega$) for $\Omega \gg 1$ [14]. The asymptotic formulas for the functions Ai($i\Omega$) and Bi($i\Omega$) and their derivatives are presented in [12]. As a result, for $\Omega \gg 1$, we have

$$I_1(\Omega) = -C\{1 - 2\sqrt{\pi}\Omega^{3/4}$$
$$\times \exp[-\sqrt{2}\Omega^{3/2}(1-i)/3 + 3\pi i/8]\},$$
$$C = 3a/\{2\ln\alpha - 2\gamma - \ln 3 + i\pi - 2\pi$$
$$\times \exp[-2^{3/2}\Omega^{3/2}(1+i)/3]\}.$$

Recall that it is convenient to represent the complex quantity $I_1(\Omega)$ by formula (14). The functions $M(\Omega)$ and $\Phi(\Omega)$ are shown Figs. 1 and 2 for several values of $\alpha \ll 1$.

DISCUSSION

From Fig. 1, we can see that curve *1* ($\alpha \longrightarrow \infty$) is very similar to curves *2–4* ($\alpha \ll 1$). For $\alpha \longrightarrow 0$, we have

$$M(\omega) \longrightarrow 3/\sqrt{(2\ln\alpha - 2\gamma - \ln 3)^2 + \pi^2}$$

as $\omega \longrightarrow \infty$. For $\alpha \longrightarrow \infty$, we have $M(\omega) \longrightarrow 1/2$ as $\omega \longrightarrow \infty$. In other words, for high modulation frequencies ω, the current amplitude increases with α. On the other hand, at small and large values of α, the time evolutions $\Phi(\Omega)$ are significantly different (Fig. 2). For low modulation frequencies, curves *1–4* are similar to each other: the current lags behind the voltage in phase and the reactive part of the impedance of the probe is inductive. For high frequencies ω, the phase shift approaches zero, $\Phi(\omega) \longrightarrow 0$, as $\alpha \longrightarrow \infty$ (curve *1*); this indicates that the reactive part of the impedance vanishes. For $\alpha \longrightarrow 0$, we have

$$\Phi(\omega) \longrightarrow \arctan[\pi/(\ln\alpha^{-2} + 2\gamma + \ln 3)]$$

as $\omega \longrightarrow \infty$. In this case, the reactive part of the impedance is capacitive; moreover, the phase difference between the current and voltage increases with α. Consequently, for high modulation frequencies, the dependence of the phase shift on the ratio α is nonmonotonic.

Since the analytic treatment presented here has been performed only for two particular limiting cases, the results obtained are rather difficult to apply in practice. However, it would hardly be possible to carry out a complete analytic study of the general case of arbitrary values of the constant component of the probe potential. To do this, it might be necessary to apply numerical methods. The above analytic solutions may be useful in verifying the computational results.

REFERENCES

1. P. Chung, L. Talbot, and K. Touryan, *Electric Probes in Stationary and Flowing Plasmas* (Springer-Verlag, Berlin, 1975; Mir, Moscow, 1978).
2. M. S. Benilov, Teplofiz. Vys. Temp. **26**, 993 (1988).
3. M. S. Benilov and B. V. Rogov, J. Appl. Phys. **70**, 6726 (1991).
4. A. V. Kashevarov, Teplofiz. Vys. Temp. **36**, 700 (1998).
5. A. V. Kashevarov, Teplofiz. Vys. Temp. **38**, 147 (2000).
6. F. G. Baksht, Zh. Tekh. Fiz. **48**, 2019 (1978).
7. E. F. Prozorov and K. N. Ul'yanov, Teplofiz. Vys. Temp. **21**, 538 (1983).
8. E. F. Prozorov and K. N. Ul'yanov, Teplofiz. Vys. Temp. **21**, 1179 (1983).
9. E. Kamke, *Differentialgleichungen*, Bd. I: *Gewöhnliche Differentialgleichungen* (Geest and Portig, Leipzig, 1964; Nauka, Moscow, 1976).
10. *Tables of Integral Exponential Functions in Complex Domain. Library of Mathematical Tables*, Ed. by K. A. Karpov (Vychisl. Tsentr Akad. Nauk SSSR, Moscow, 1965), Vol. 31.
11. I. M. Cohen, Phys. Fluids **6**, 1492 (1963).
12. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
13. M. Rothman, Q. J. Mech. Appl. Math. **7**, 379 (1954).
14. F. W. J. Olver, *Asymptotics and Special Functions* (Academic, New York, 1974; Nauka, Moscow, 1990).
15. M. S. Benilov and G. A. Tirskiĭ, Dokl. Akad. Nauk SSSR **240**, 1324 (1978) [Sov. Phys. Dokl. **23**, 395 (1978)].
16. P. M. Woodward and A. M. Woodward, Philos. Mag., Ser. 7 **37** (267), 236 (1946).

Translated by O. Khadin

Technical Physics, Vol. 46, No. 9, 2001, pp. 1093–1100. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 26–32.

GAS DISCHARGES, PLASMA

Porous Piezoelectric Composites with Extremely High Reception Parameters

V. Yu. Topolov and A. V. Turik

Rostov State University, pr. Stachki 194, Rostov-on-Don, 344004 Russia

e-mail: topolov@phys.rnd.runnet.ru

Received November 13, 2000

Abstract—The reception parameters Q_{33}^* and Q_h^* of modified fiber composites based on high-$e_{33}^{FC}/c_{33}^{FC,E}$ ferroelectric piezoceramics (e_{33}^{FC} is the piezoelectric constant, $c_{33}^{FC,E}$ is the elastic modulus) are considered as functions of the electromechanical properties and the porosity of the components. Pore configurations at which the values $(Q_{33}^*)^2 \approx (Q_h^*)^2 \approx 7 \times 10^{-9}$ Pa^{-1} for the polymer matrices of 1–0–3 composites are much higher than for conventional ternary composites are analyzed. © *2001 MAIK "Nauka/Interperiodica".*

Composites based on ferroelectric piezoelectric ceramic (FPC) materials offer many physical properties of practical value. Of binary piezoactive composites, fiber materials (extended FPC filaments in a polymer matrix that have 1–3 connectivity) are most commonly used. In α_1–α_2 connectivity, the integers α_1, α_2 = 0–3 define the number of axes in an $(X_1X_2X_3)$ Cartesian system along which the first, α, and the second, α_2, components are distributed continuously [1, 2]. Numerous experiments and theoretical studies suggest that FPC-based 1–3 composites exhibit a nonmonotonic concentration dependence of all four piezoelectric constants d_{ij}^* [3], e_{ij}^* [3–6], g_{ij}^* [7], and h_{ij}^* [4]; planar, k_p^*, and thickness, k_t^*, electromechanical coupling coefficients [3–6, 8]; anisotropy coefficients of the piezoelectric constants $\zeta_e^* = e_{33}^*/e_{31}^* = h_{33}^*/h_{31}^*$ and $\zeta_d^* = d_{33}^*/d_{31}^* = g_{33}^*/g_{31}^*$ [3, 6]; reception parameters Q_{33}^* and Q_h^* [9–13]; and other effective constants. For the composites polarized along the OX_3 axis (∞mm symmetry), the reception parameters squared are given by

$$(Q_{33}^*)^2 = d_{33}^* g_{33}^* = (d_{33}^*)^2/\varepsilon_{33}^{*\sigma},$$
$$(Q_h^*)^2 = d_h^* g_h^* = (d_h^*)^2/\varepsilon_{33}^{*\sigma}, \quad (1)$$

where $d_h^* = d_{33}^* + 2d_{31}^*$ and $g_h^* = g_{33}^* + 2g_{31}^*$ are the hydrostatic piezoelectric constants and $\varepsilon_{33}^{*\sigma}$ is the permittivity of the mechanically unloaded material.

The parameters Q_{33}^* and Q_h^* in (1) are related to the power densities and the signal-to-noise ratios of piezoelectric elements, and d_h^* and g_h^* characterize the hydrostatic stability of the elements and piezoelectric devices. It has been found that the peaks of the concentration dependences of $(Q_{33}^*)^2(m)$ and $(Q_h^*)^2(1)$ for 1–3 composites "Pb(Zr, Ti)O_3 (volume concentration m)–polymer matrix (volume concentration $1 - m$)" are several tens of times higher than those for $(Q_{33}^*)^2(1)$ and $(Q_h^*)^2(1)$; for various 1–3 composites, $\max(Q_h^*)_{1-3}^2(m) =$ 2000–9500 and $\max(Q_{33}^*)_{1-3}^2(m)$ = 20000–50000 (in units of 10^{-15} Pa^{-1}) [9, 11, 13, 14].[1] The calculated values of $|\zeta_e^*|, |\zeta_d^*|$ [6], and g_h^* [9, 11, 13] of the 1–3 composites are frequently much greater than experimentally found $e_{33}^{FC}/|e_{31}^{FC}|$, $d_{33}^{FC}/|d_{31}^{FC}|$, and $g_{33}^{FC} + 2g_{31}^{FC}$ of FPC composites. Moreover, the complication of the composite geometry may raise g_h^* and g_{33}^* by more than one order of magnitude above $g_{33}^{FC} + 2g_{31}^{FC}$ and g_{33}^{FC}, respectively [15] (hereafter, the superscript *FC* is assigned to the electromechanical constants of ferroelectric ceramics used as the components of the composites). The above effective properties of binary fiber composites and like materials are finding wide application in piezoelectric transducers, as well as in hydroacoustic, medical, and other devices.

The possibility of producing ternary composites involving FPC filaments has been discussed in [16–18]. In particular, it has been shown [18] that $(Q_h^*)_{1-0-3}^2$ may exceed 10 000 (PPa)$^{-1}$ for 1–0–3 connectivity and $(Q_h^*)_{1-2-2}^2$ may attain 50 500 (PPa)$^{-1}$ for 1–2–2 connectivity (FPC filaments in a 2–2 polymer matrix) [2]. The

[1] According to [2, 10, 11, 14], $(Q_{33}^*)^2$ and $(Q_h^*)^2$ will hereafter be expressed in (PPa)$^{-1}$.

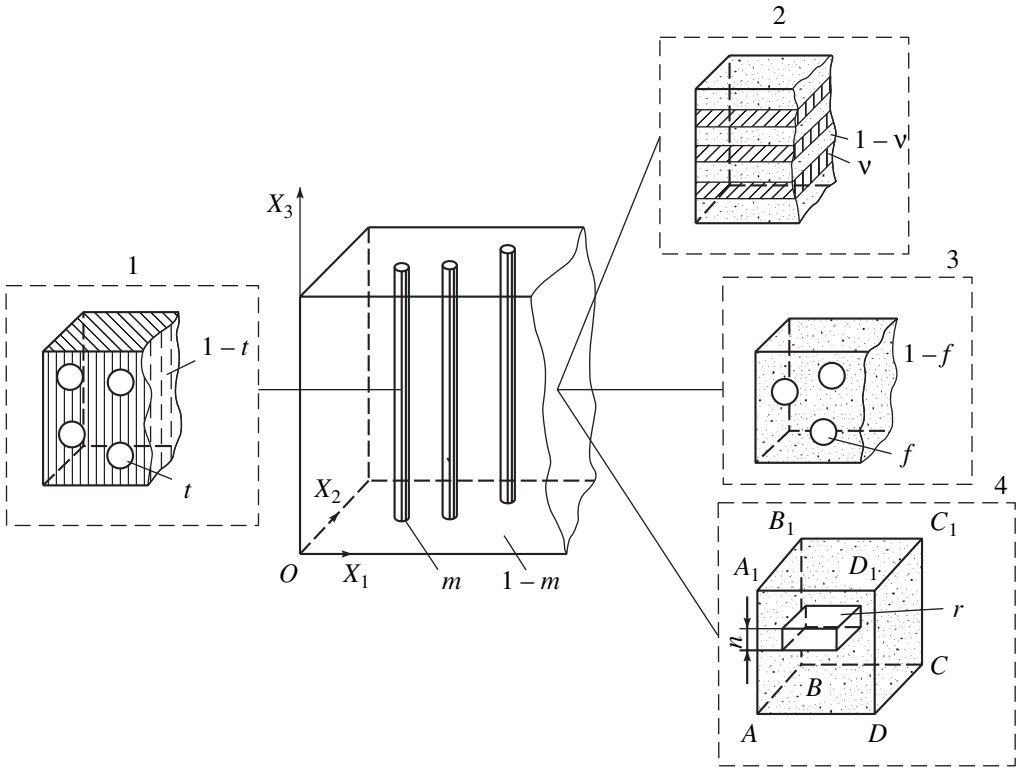


Fig. 1. Schematic diagram of the composites with various connectivities (insets 1–3). The volume concentration of spherical pores in the filaments is *t* and in the surrounding matrix it is *f* (spherical pores) or *nr* (parallelepiped-like pores arranged within the Banno cubic cell [20, 21] $ABCDA_1B_1C_1D_1$ as shown in inset 4), where *n* and *r* are concentration parameters: *n* equals the ratio of the parallelepiped height to the cube edge $|AA_1|$; and *r* is the ratio of the parallelepiped base area to the cube face area $|AB| \cdot |BC|$. In inset 2, ν and 1 – ν are the volume concentrations of the layers of the first and second types, respectively.

latter value exceeds the value $(Q_h^*)^2_{2\text{–}2\text{–}0} = 50\,000$ (PPa)$^{-1}$ (for 2–2–0 connectivity), which until recently has been the highest for ternary composites [2]. Of much interest is also the effect of pore configuration on the effective properties of FPC [19–22] and polymer [12] materials. Although the averaging methods applied to ferroelectric piezoceramics (FPCs) and porous polymers are approximate and the elastic and dielectric properties of the solid components (FPCs and polymers) and air substantially differ (at least by 10^3 times and by 10–10^3 times, respectively), the results sometimes agree well with experimental data. Agreement is observed for the porous media that can be considered as composite with connectivity 0–3 (i.e., the pores are closed) [19, 20] or 3–3 (the pores are open) [21]. It has been shown [12] that the reception parameters of the fiber composites can be much improved compared with $d_{33}^{FC} g_{33}^{FC}$ and $d_h^{FC} g_h^{FC}$ by producing a complex configuration of pores in the polymer matrix. The pores provide a negative Poisson's ratio σ_{13}^M for the matrix and are responsible for the nonuniform fields of internal mechanical stresses. The values of $\max(Q_h^*)^2_{GT} = 2.200 \times 10^6$ (PPa)$^{-1}$ for $\sigma_{13}^M > 0$ and $\max(Q_h^*)^2_{GT} = 5.655 \times 10^6$ (PPa)$^{-1}$ for $\sigma_{13}^M < 0$ [12], which were obtained for the 1–0–3 composite "PZT-5A filaments–polyurethane matrix with conical pores," are two or three orders of magnitude higher than the values of $(Q_h^*)^2$ for $Pb(Zr,Ti)O_3$-based composites used in practice [2, 14]. The above-mentioned jump from $(Q_h^*)^2_{1\text{–}3} = (10^3\text{–}10^4)$ (PPa)$^{-1}$ to $(Q_h^*)^2_{1\text{–}2\text{–}2} \approx 5 \times 10^4$ (PPa)$^{-1}$, which is associated with modification of the fiber structure without allowance for porosity [18], as well as unique data [12] for $(Q_h^*)^2$ for 1–0–3 porous composites with $\operatorname{sgn}\sigma_{13}^M = \pm 1$, are strong impetuses to extend works on the prediction of the effective properties of the ternary composites. In this article, we elaborate upon and generalize theoretical concepts [18] and analyze factors that can improve the reception parameters and the hydrostatic sensitivity of the porous composites with 1–3 connectivity.

In the model we put forward, the composite consists of extended cylindrical piezoactive filaments running parallel to the polarization axis OX_3 and surrounded by a nonpiezoelectric matrix (Fig. 1). Two cases are considered: porous FPC filaments of connectivity 0–3 sur-

Table 1. Electromechanical constants of ferroelectric and polymer components measured at room temperature

FPC material	$c_{11}^{FC,E}$, 10^{10} Pa	$c_{12}^{FC,E}$, 10^{10} Pa	$c_{13}^{FC,E}$, 10^{10} Pa	$c_{33}^{FC,E}$, 10^{10} Pa	$c_{44}^{FC,E}$, 10^{10} Pa	e_{31}^{FC}, C/m^2	e_{33}^{FC}, C/m^2	e_{15}^{FC}, C/m^2	$\varepsilon_{11}^{FC,\xi}/\varepsilon_0$	$\varepsilon_{33}^{FC,\xi}/\varepsilon_0$	$d_{33}^{FC}g_{33}^{FC}$, (PPa)$^{-1}$	$d_h^{FC}g_h^{FC}$, (PPa)$^{-1}$
PKR-7M [25]	13.3	9.2	9.1	12.5	2.28	–9.5	31.1	20.0	1980	1810	13000	84
PZT-5A [12, 27]	12.1	7.54	7.52	11.1	2.11	–5.4	15.8	12.3	916	830	9300	67

Polymer	c_{11}^{p}, 10^{10} Pa	c_{12}^{p}, 10^{10} Pa	$\varepsilon_{kk}^{p}/\varepsilon_0$, k = 1; 2; 3
Elastomer [9]	0.114	0.0931	5.0
Araldit [26]	0.78	0.44	4.0
Polyurethane	0.442	0.260	3.5

rounded by a homogeneous polymer matrix (Fig. 1, inset 1) and FPC filaments surrounded by a heterogeneous polymer matrix. The heterogeneous matrix may contain layers of two dissimilar polymers (connectivity 2–2, Fig. 1, inset 2), consist of a polymer with air-filled pores in the form of a sphere or a parallelepiped (cellular matrix of connectivity 0–3; Fig. 1, insets 3 and 4), or be laminated and porous. In the last case, it is assumed that the layers of porous and homogeneous polymers of the same chemical composition alternate, with the interfaces parallel to the X_1OX_2 plane. The effective constants of the porous composites are determined it two steps. First, the effective constants of the spherical-pore medium (Fig. 1; insets 1, 3) are found by averaging the elastic, $c_{ab}^{FC,E}$; piezoelectric, e_{ij}^{FC}; and dielectric, $\varepsilon_{kk}^{FC,\xi}$, constants of the FPC material or by averaging the elastic, c_{ab}^{p}, and dielectric, ε_{ff}^{p}, constants of the polymer using formulas for connectivity 0–3 (the self-consistency method for piezopassive [23] and piezoactive [24] media). The effective constants of the 2–2 matrix (Fig. 1, inset 1) are calculated from formulas for laminated media [17, 25]; and those for the 0–3 matrix with parallelepiped-like pores (Fig. 1, inset 4), using the matrix method [3, 26]. Then, the effective elastic moduli c_{ab}^{*E}, piezoelectric constants e_{ij}^{*} and d_{ij}^{*}, and dielectric permittivities $\varepsilon_{ff}^{*\sigma}$ of the modified fiber composite are found by averaging the associated constants for the filaments and the matrix using the self-consistency method [9] for connectivity 1–3. Subsequently, the effective constants thus obtained are employed to determine the reception parameters [from (1)] and other material properties.

From experimental data for $c_{ab}^{FC,E}$, e_{ij}^{FC}, $\varepsilon_{kk}^{FC,\xi}$, c_{ab}^{p}, and ε_{kk}^{p} [25, 27], and also from the results of computation [18], it follows that $(Q_{33}^{*})^2$ and $(Q_h^{*})^2$ peak at a filament concentration $m \ll 1$. Because of the smallness of m and the resulting high anisotropy, $|\zeta_e^{*}| \gg 1$ [4, 6, 18], expressions (1) can be recast as

$$\begin{aligned}(Q_h^{*})^2 &\approx \eta_{\text{elas},h}^{*}(e_{33}^{*}/c_{33}^{*E})^2/\varepsilon_{33}^{*\sigma}\\ &\approx \eta_{\text{elas},h}^{*}[me_{33}^{FC}/(mc_{33}^{FC,E}+c_{33}^{M})]^2/(m\varepsilon_{33}^{FC,\sigma}),\\ (Q_{33}^{*})^2 &\approx (\eta_{\text{elas},33}^{*}/\eta_{\text{elas},h}^{*})(Q_h^{*})^2,\end{aligned} \tag{2}$$

where

$$\begin{aligned}\eta_{\text{elas},h}^{*} &= \{[1+(\beta_{12}^{*})^{-1}-2(\beta_{13}^{*})^{-1}]/[1+(\beta_{12}^{*})^{-1}\\ &\quad -2(\beta_{13}^{*})^{-1}(\beta_{33}^{*}/\beta_{13}^{*})]\}^2,\\ \eta_{\text{elas},33}^{*} &= \{[1+(\beta_{12}^{*})^{-1}]/[1+(\beta_{12}^{*})^{-1}\\ &\quad -2(\beta_{13}^{*})^{-1}(\beta_{33}^{*}/\beta_{13}^{*})]\}^2,\end{aligned} \tag{3}$$

$\beta_{ab}^{*} = c_{11}^{*E}/c_{ab}^{*E}$, and c_{33}^{M} is the elastic modulus of the matrix.

From (2), we obtain that the reception parameters of the composite depend on the ratio

$$\gamma = me_{33}^{FC}/(mc_{33}^{FC,E}+c_{33}^{M}), \tag{4}$$

the permittivity $\varepsilon_{33}^{FC,\sigma}$ of the FPC material, and the differences between the elastic moduli of the FPC filaments, $c_{ab}^{FC,E}$, and the matrix, c_{qr}^{M}. Among perovskite-like FPCs [25, 27], PKR-7M (one of the so-called Rostov piezoceramics) has the highest ratio $e_{33}^{FC}/c_{33}^{FC,E}$ at room temperature; hence, high values of γ in (4) at $m \ll 1$. An elastomer–araldit laminated structure (Fig. 1, inset 2) features a large difference between c_{11}^{M}/c_{33}^{M} and c_{12}^{M}/c_{13}^{M} [18], so that $\eta_{\text{elas},h}^{*}$ and $\eta_{\text{elas},33}^{*}$ in (3) are appreciable. Below, we will consider how the shape of pores in the filaments and the matrix affects parameters (1) of the composite based on the PKR-7M ceramic. The electromechanical constants of the ceramics and polymers are listed in Table 1.

Table 2. Correlation between the elastic parameters of the laminated porous matrix, coefficient $\eta^*_{\text{elas, }h}$ in (3), and the peaks of the functions $(Q^*_h)^2_{1-0-3}(m, \nu, f)$ and $\max(Q^*_{33})^2_{1-0-3}(m, \nu, f)$ all calculated for the 1–0–3 "PKR-7M filaments–laminated porous elastomer matrix with spherical pores" composite

f	0.90		0.80			0.70		
ν	0.10	0.20	0.10	0.20	0.30	0.10	0.20	0.30
c^M_{11}/c^M_{12}	1.71	1.88	1.52	1.69	1.80	1.43	1.57	1.67
c^M_{11}/c^M_{13}	3.15	4.86	2.07	2.81	3.19	1.71	2.14	2.50
c^M_{11}/c^M_{33}	2.45	3.59	1.61	2.09	2.43	1.34	1.60	1.78
$\eta^*_{\text{elas, }h}$	0.552	0.657	0.339	0.465	0.528	0.250	0.331	0.401
$\max(Q^*_h)^2_{1-0-3}$, (PPa)$^{-1}$	4.94×10^5	1.11×10^6	1.67×10^5	4.00×10^5	6.10×10^5	7.97×10^4	1.89×10^5	3.10×10^5
$\max(Q^*_{33})^2_{1-0-3}$, (PPa)$^{-1}$	1.38×10^6	2.09×10^6	9.76×10^5	1.32×10^6	1.72×10^6	8.31×10^5	1.04×10^6	1.25×10^6

f	0.60				0.50				
ν	0.10	0.20	0.30	0.40	0.10	0.20	0.30	0.40	0.50
c^M_{11}/c^M_{12}	1.37	1.48	1.57	1.64	1.33	1.42	1.49	1.56	1.62
c^M_{11}/c^M_{13}	1.53	1.80	2.04	2.23	1.42	1.60	1.76	1.89	1.99
c^M_{11}/c^M_{33}	1.20	1.35	1.46	1.53	1.12	1.21	1.28	1.32	1.33
$\eta^*_{\text{elas, }h}$	0.175	0.254	0.291	0.335	0.137	0.181	0.225	0.242	0.268
$\max(Q^*_h)^2_{1-0-3}$, (PPa)$^{-1}$	4.47×10^4	1.01×10^5	1.64×10^5	2.29×10^5	2.77×10^4	5.73×10^4	9.10×10^4	1.25×10^5	1.56×10^5
$\max(Q^*_{33})^2_{1-0-3}$, (PPa)$^{-1}$	7.62×10^5	9.00×10^5	1.03×10^6	1.18×10^6	7.18×10^5	8.10×10^5	9.07×10^5	1.00×10^6	1.09×10^6

Note: The coefficient $\eta^*_{\text{elas, }h}$ in (3) was calculated for volume concentrations m_j, ν, and f that correspond to $\max(Q^*_h)^2_{1-0-3}$. The mean porosity of the matrix is specified in the interval of $0 < \nu f \le 0.25$.

The introduction of FPC filaments with spherical pores into the homogeneous polymer matrix (Fig. 1, inset 1) causes an increase in $e^{FC}_{33}/c^{FC,E}_{33}$. As a result, e^*_{33}/c^{*E}_{33} and the hydrostatic piezoelectric constant g^*_h of the 0–1–3 composite also grow. These circumstances, as well as the high anisotropy $|\zeta^*_e| \ge 10^2$ at $m \ll 1$, raise, according to formulas (2) and (3), $\max(Q^*_h)^2_{0-1-3}(m, t)$ and $\max(Q^*_{33})^2_{0-1-3}(m, t)$ (Figs. 2a, 2b) in comparison with $\max(Q^*_h)^2_{1-3}(m) = 6240$ (PPa)$^{-1}$ and $\max(Q^*_{33})^2_{1-3}(m) = 6.24\times10^5$ (PPa)$^{-1}$ for the 1–3 "PKR-7M–elastomer" composite. However, the improvement of parameters (1) of the 0–1–3 porous filaments embedded in the homogeneous matrix is less appreciable than in the case when the homogeneous filaments of the same composition are embedded in the heterogeneous (laminated [18] or porous) polymer matrix. This statement is confirmed by the calculation of $(Q^*_h)^2_{0-1-3}$ and $(Q^*_{33})^2_{0-1-3}$ for 1–0–3 composites with spherical pores distributed throughout the polymer matrix (Fig. 1, inset 3; Figs. 2c, 2d) or over similar layers (Table 2). In the latter case, the laminated porous elastomer matrix has the structure shown in inset 2 of Fig. 1 and the layers of concentration ν have pores with the volume concentration f (Fig. 1, inset 3).

The calculated values of $\max(Q^*_h)^2_{1-0-3} = (Q^*_h)^2_{1-0-3}(m_0, f)$ and $\max(Q^*_{33})^2_{1-0-3} = (Q^*_{33})^2_{1-0-3}(m_0, f)$ of the 1–0–3 "PKR-7M filaments–spherical-pore elastomer matrix" composites correlate, at f = const, with the ratios $c^M_{11}/c^M_{12} = c^M_{11}/c^M_{13} = c^M_{33}/c^M_{13}$ of the elastic

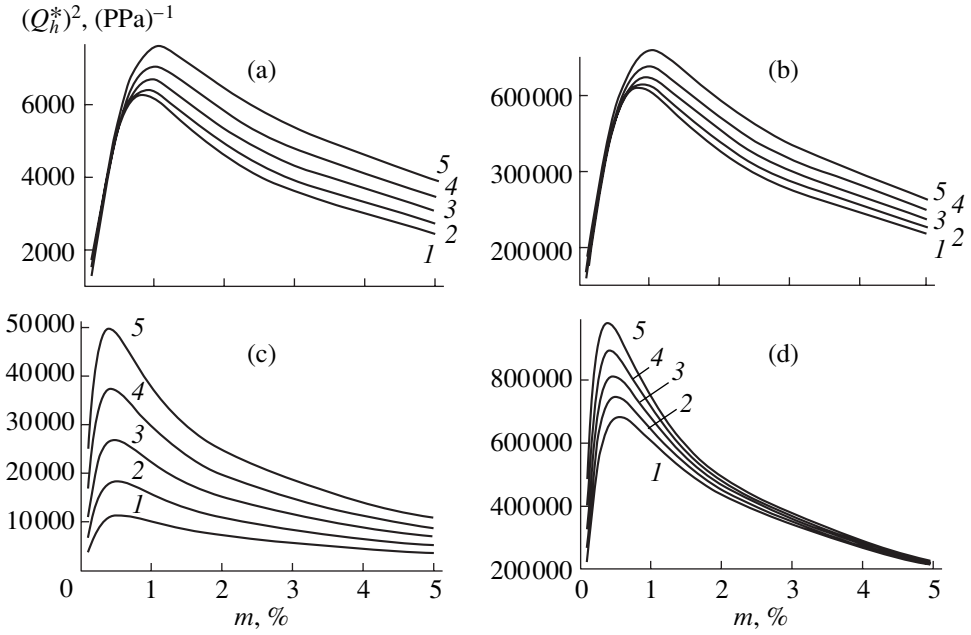


Fig. 2. Concentration dependences of the reception parameters squared (a) $(Q_h^*)^2_{0\text{–}1\text{–}3}(m, t)$, (b) $(Q_{33}^*)^2_{0\text{–}1\text{–}3}(m, t)$, (c) $(Q_h^*)^2_{0\text{–}1\text{–}3}(m, f)$, and (d) $(Q_{33}^*)^2_{0\text{–}1\text{–}3}(m, f)$ calculated for the (a, b) 0–1–3 "PKR-7M spherical-pore filaments–elastomer matrix" composite and (c, d) 1–0–3 "PKR-7M filaments–porous elastomer matrix" composite. (a, b) t = (*1*) 0.05, (*2*) 0.10, (*3*) 0.15, (*4*) 0.20, and (*5*) 0.25. (c, d) f = (*1*) 0.05, (*2*) 0.10, (*3*) 0.15, (*4*) 0.20, and (*5*) 0.25.

moduli of the isotropic porous matrix and with $\eta^*_{\text{elas},h}$ in (3). For example, as f increases from 0.05 to 0.25, c_{33}^M/c_{13}^M monotonically grows from 1.30 to 1.59 and $\eta^*_{\text{elas},h}(m_0, f)$, grows from 0.0721 to 0.109. The values of $\max(Q_h^*)^2_{1\text{–}0\text{–}3}$ and $\max(Q_{33}^*)^2_{1\text{–}0\text{–}3}$ also rise monotonically (Figs. 2c, 2d). One reason for the correlation between these concentration dependences is that $(Q_h^*)^2$ strongly depends on $\eta^*_{\text{elas},h}$ according to (2). In addition, as follows from (3), $\eta^*_{\text{elas},h}$ considerably depends on the ratio $\beta_{33}^*/\beta_{13}^*$, which is reduced to

$$\beta_{33}^*/\beta_{13}^* \approx c_{13}^*/c_{33}^M \qquad (5)$$

at $m_0 \approx 10^{-2}$ and f = const.

Similar considerations hold for the coefficient $\eta^*_{\text{elas},33}$ in (3), which governs $(Q_{33}^*)^2$ [see formulas (2)] and the values of $\max(Q_{33}^*)^2_{1\text{–}0\text{–}3}$ calculated.

The computed values of $\max(Q_h^*)^2_{1\text{–}0\text{–}3}$ and $\max(Q_{33}^*)^2_{1\text{–}0\text{–}3}$ in Table 2 are much larger than those in Figs. 2c and 2d, because the elastic modulus ratios c_{11}^M/c_{12}^M, c_{11}^M/c_{13}^M, and c_{11}^M/c_{33}^M differ in the laminated porous matrix. Comparing the data in Table 2 that were computed for various concentrations m_j, ν, and f corresponding to $\max(Q_h^*)^2_{1\text{–}0\text{–}3}$ and $\max(Q_{33}^*)^2_{1\text{–}0\text{–}3}$, we conclude that c_{11}^M/c_{33}^M has the most significant effect on $\eta^*_{\text{elas},h}$ in (3) and parameters (1), while the effect of c_{11}^M/c_{13}^M is somewhat weaker. This directly follows from relationship (5) and is obvious, for example, for $\max(Q_h^*)^2_{1\text{–}0\text{–}3} = (Q_h^*)^2_{1\text{–}0\text{–}3}(m_1;\ 0.70;\ 0.10)$ and $\max(Q_h^*)^2_{1\text{–}0\text{–}3} = (Q_h^*)^2_{1\text{–}0\text{–}3}(m_2;\ 0.50;\ 0.20)$ or for $\max(Q_h^*)^2_{1\text{–}0\text{–}3} = (Q_h^*)^2_{1\text{–}0\text{–}3}(m_3;\ 0.80;\ 0.10)$ and $\max(Q_h^*)^2_{1\text{–}0\text{–}3} = (Q_h^*)^2_{1\text{–}0\text{–}3}(m_4;\ 0.70;\ 0.20)$ (in every case $10^{-3} < m_j < 10^{-2}$). As for the differences between the values of $\max(Q_h^*)^2_{1\text{–}0\text{–}3}(m, \nu, f)$ that are obtained at nearly equal c_{11}^M/c_{33}^M and $\eta^*_{\text{elas},h}$ (cf., e.g., the estimates for f = 0.60, ν = 0.20 and f = 0.70, ν = 0.10 in Table 2), they can be explained by the effect of $\varepsilon_{33}^{*\sigma}(m, \nu, f)$ on $(Q_h^*)^2$ at $m \ll 1$ [see formulas (1) and (2)]. Note also that the values of $\max(Q_h^*)^2_{1\text{–}0\text{–}3}(m, \nu, f)$ and $\eta^*_{\text{elas},h}$ at $\nu \geq 0.70$ given in Table 2 are considerably larger than $\max(Q_h^*)^2_{1\text{–}2\text{–}2}(m, \nu)$ and $\eta^*_{\text{elas},h}(m, \nu)$ for the 1–2–2

Table 3. Correlation between the elastic parameters of the porous matrix, coefficient $\eta^*_{elas,\,h}$ in (3), and the peaks of the functions $(Q^*_h)^2_{1-0-3}(m, n, r)$ and $\max(Q^*_{33})^2_{1-0-3}(m, n, r)$ all calculated for the 1–0–3 PKR-7M filaments–elastomer matrix with plane-parallel pores composite

n	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
					$r = 0.99$					
c^M_{11}/c^M_{12}	2.16	2.19	2.20	2.21	2.21	2.21	2.22	2.22	2.22	2.22
c^M_{11}/c^M_{13}	33.8	66.5	99.1	132	164	197	230	262	295	327
c^M_{11}/c^M_{33}	27.4	53.2	78.6	103	128	151	175	197	219	241
$\eta^*_{elas,\,h}$	0.933	0.961	0.972	0.978	0.982	0.985	0.986	0.988	0.989	0.990
$\max(Q^*_h)^2_{1-0-3}$, 10^6 (PPa)$^{-1}$	3.96	5.05	5.57	5.90	6.15	6.35	6.52	6.66	6.80	6.92
$\max(Q^*_{33})^2_{1-0-3}$, 10^6 (PPa)$^{-1}$	4.31	5.28	5.74	6.04	6.27	6.45	6.61	6.75	6.88	7.00
					$r = 0.90$					
c^M_{11}/c^M_{12}	2.16	2.19	2.20	2.21	2.21	2.21	2.21	2.22	2.22	2.22
c^M_{11}/c^M_{13}	30.9	60.5	90.2	120	150	179	209	239	268	298
c^M_{11}/c^M_{33}	25.0	48.6	71.7	94.4	117	138	160	181	201	221
$\eta^*_{elas,\,h}$	0.928	0.958	0.970	0.977	0.981	0.983	0.985	0.987	0.988	0.989
$\max(Q^*_h)^2_{1-0-3}$, 10^6 (PPa)$^{-1}$	3.79	4.91	5.45	5.80	6.05	6.24	6.41	6.55	6.69	6.81
$\max(Q^*_{33})^2_{1-0-3}$, 10^6 (PPa)$^{-1}$	4.16	5.16	5.64	5.95	6.17	6.36	6.51	6.64	6.77	6.89

Note: The coefficient $\eta^*_{elas,\,h}$ in (3) was calculated for volume concentrations m_j, n, and r that correspond to $\max(Q^*_h)^2_{1-0-3}$. The mean porosity of the matrix is specified in the interval of $0 < nr < 0.10$.

"PKR-7M filaments–laminated elastomer–araldit matrix" composite [18].

The extremely high $(Q^*_{33})_{1-0-3}$ and $(Q^*_h)_{1-0-3}$ are observed in the case of plane-parallel air-filled pores in the homogeneous polymer matrix (Fig. 1, inset 4). When the concentration of these pores varies in the range $0.01 \le n \le 0.10$ and $0.90 \le r \le 0.99$, the values of $\max(Q^*_h)^2_{1-0-3}$ and $\max(Q^*_{33})^2_{1-0-3}$ change only slightly (Table 3): they remain roughly one order of magnitude higher than those computed for the laminated porous matrix with a mean porosity $\nu f \lesssim 0.10$ (Table 2). The disadvantages of the 1–0–3 composites with plane-parallel pores in the matrix are (1) small ($m_j < 0.5\%$) volume concentrations of the PKR-7M filaments at which $\max(Q^*_h)^2_{1-0-3}$ and $\max(Q^*_{33})^2_{1-0-3}$ are attained and (2) the possibility of electrical breakdown because of the electric field concentration at pore edges [28] normal to the OX_3 axis and the external field vector **E**. Known calculated data for $\max(Q^*_h)^2$ and $\max(Q^*_{33})^2$ indicate that the corresponding volume concentrations of the FPC filaments in 1–3 [9–13], 1–2–2 [18], and 1–0–3 [12, 18] composites are small ($m \approx 1\%$). This seems likely to retard the use of piezoelectric composites of this sort. In the 1–0–3 "PKR-7M filaments–elastomer matrix with plane-parallel pores" composite, $(Q^*_h)^2_{1-0-3}(m, n, r)$ and $(Q^*_{33})^2_{1-0-3}(m, n, r)$ vary (at $m > m_j$, $n \ll 1$, and $r \longrightarrow 1$) in such a way that parameters (1) remain roughly one order of magnitude higher than those for the 1–3 com-

posites [9–11, 13, 14] even at $m \approx 10\%$. A comparison with the calculated data for the 1–2–2 "PKR-7M–laminated elastomer–araldit matrix" [18] shows that $(Q_h^*)^2_{1-0-3} \approx 10 \max(Q_h^*)^2_{1-2-2}$ at $m = 2\%$ and $(Q_h^*)^2_{1-0-3} \approx 4 \max(Q_h^*)^2_{1-2-2}$ at $m = 5\%$.

The interrelation between c_{11}^M / c_{a3}^M ($a = 1, 3$), $\eta^*_{\text{elas}, h}$ in (3), and the parameters $(Q_{33}^*)^2$ and $(Q_h^*)^2$ in (1) and (2), which is observed in the composites with different pore configurations (Tables 2, 3), suggests that the difference between the elastic properties of the FPCs and the matrix are essential for the composites to have a high piezoelectric sensitivity. The presence of porous layers in the homogeneous matrix or the formation of plane-parallel pores in it (Fig. 1, insets 2, 4) causes a system of alternating compacted and rarefied nonpiezoelectric regions to occur along the OX_3 axis. Simultaneously, the elastic moduli c_{33}^M and c_{13}^M of the heterogeneous matrix noticeably decrease in comparison with c_{11}^p and c_{12}^p of the homogeneous polymer. As a result, the configuration of the mechanical stresses within the composite changes so that $\eta^*_{\text{elas}, h}$ and $\eta^*_{\text{elas}, 33}$ in (3) become larger than in the matrices shown in insets 2 and 3 in Fig. 1. From Tables 2 and 3, it follows that the highest values of $\eta^*_{\text{elas}, h} \approx 1$, $\max(Q_h^*)^2_{1-0-3}$, and $\max(Q_{33}^*)^2_{1-0-3}$ are observed in the porous matrix (Fig. 1, inset 4) with $c_{11}^M / c_{a3}^M \gg 10$ much higher than for the other matrices (Fig. 1, insets 2, 3).

The values of $\max(Q_h^*)^2_{1-0-3}$ (Table 3) for the concentration ranges $0.01 \le n \le 0.10$ and $0.90 \le r \le 0.99$ exceed $\max(Q_h^*)^2_{GT} = 2.200 \times 10^6$ (PPa)$^{-1}$, which has been obtained [12] for "PZT-5A–porous polyurethane matrix" composites at $\sigma_{13}^M > 0$. According to our estimates, for a similar PZT-5A-based composite whose matrix contains plane-parallel pores with $0.01 \le n \le 0.10$ and $r = 0.99$ (Fig. 1, inset 4), $\max(Q_h^*)^2_{1-0-3} = 3.26 \times 10^6$ (PPa)$^{-1}$ > $\max(Q_h^*)^2_{GT}$ at $n = 0.10$, $\max(Q_h^*)^2_{1-0-3} = 2.90 \times 10^6$ (PPa)$^{-1}$ > $\max(Q_h^*)^2_{GT}$ at $n = 0.05$, and $\max(Q_h^*)^2_{1-0-3} = 1.94 \times 10^6$ (PPa)$^{-1}$ < $\max(Q_h^*)^2_{GT}$ at $n = 0.01$. Thus, the assumption that $\max(Q_h^*)^2_{GT}$ is the upper theoretical limit of the hydrostatic sensitivity of the porous composite with $\sigma_{13}^M > 0$ needs correction. In our opinion, the authors of [12] have underestimated the effect of pores, which may markedly decrease c_{13}^M and c_{33}^M and increase $\eta^*_{\text{elas}, h}$ in (3) (see, e.g., Tables 2, 3). In addition, the ratio $\max(Q_h^*)^2_{1-0-3} / (d_h^{FC} g_h^{FC})$ estimated by us at 8.24×10^4 for $n = 0.10$ and $r = 0.99$ ("PKR-7M filaments–elastomer matrix with plane-parallel pores" composite) also exceeds the value $\max(Q_h^*)^2_{GT} / (d_h^{FC} g_h^{FC}) = 3.33 \times 10^4$ obtained in [12]. Finally, for the "PKR-7M filaments–polyurethane matrix with plane-parallel pores" composite, $\max(Q_h^*)^2_{1-0-3}$ at $n = 0.10$ and $r = 0.99$ is no higher than 7.20×10^6 (PPa)$^{-1}$; this slight rise (cf. data in Table 3) in the piezoelectric sensitivity is explained primarily by the fact the permittivity ε_{kk}^p of polyurethane is smaller than that of the elastomer (Table 1).

CONCLUSIONS

(1) Factors that govern the reception parameters Q_{33}^* and Q_h^* of "porous FPC filaments–polymer matrix, FPC filaments–porous polymer matrix, and FPC filaments–porous laminated polymer matrix" were studied. It was shown that $\max(Q_{33}^*)^2$ and $\max(Q_h^*)^2$ strongly depend on $e_{33}^{FC} / c_{33}^{FC, E}$, c_{11}^M / c_{33}^M, and c_{11}^M / c_{13}^M and to a lesser extent on $\varepsilon_{kk}^{FC, \sigma}$, ε_{kk}^p, and the difference between $c_{ab}^{FC, E}$ and c_{ab}^p.

(2) For the "PKR-7M filaments–elastomer matrix with plane-parallel pores" composite, the values of $\max(Q_{33}^*)^2_{1-0-3}$ and $\max(Q_h^*)^2_{1-0-3}$ calculated indicate that our model involving the more effective redistribution of internal mechanical stresses in the composite than that considered in [12] is valid. It was demonstrated that the coefficient $\eta^*_{\text{elas}, h}$ can be used as a quantitative measure of this redistribution. The correlation between c_{11}^M / c_{a3}^M, $\eta^*_{\text{elas}, h}$ and $\max(Q_{33}^*)^2_{1-0-3}$, $\max(Q_h^*)^2_{1-0-3}$ that was established for the different pore configurations and concentrations in the matrix will help to optimize the search for FPC and polymer materials that provide high piezoelectric sensitivity of related composites.

REFERENCES

1. R. E. Newnham, D. P. Skinner, and L. E. Cross, Mater. Res. Bull. **13**, 525 (1978).
2. R. E. Newnham, Mater. Res. Bull. **22** (5), 20 (1997).
3. F. Levassort, V. Yu. Topolov, and M. Lethiecq, J. Phys. D **33**, 2064 (2000).
4. H. Jensen, IEEE Trans. Ultrason., Ferroelectr. Freq. Control **38**, 591 (1991).
5. H. Taunaumang, I. L. Guy, and H. L. W. Chan, J. Appl. Phys. **76**, 484 (1994).
6. V. Yu. Topolov and A. V. Turik, J. Appl. Phys. **85**, 372 (1999).

7. M. J. Haun and R. E. Newnham, Ferroelectrics **68**, 123 (1986).
8. J. Mendiola and B. Jimenez, Ferroelectrics **53**, 159 (1984).
9. A. A. Grekov, S. O. Kramarov, and A. A. Kuprienko, Mekh. Kompoz. Mater., No. 1, 62 (1989).
10. G. Hayward, IEEE Trans. Ultrason., Ferroelectr. Freq. Control **43**, 98 (1996).
11. J. Bennett and G. Hayward, IEEE Trans. Ultrason., Ferroelectr. Freq. Control **44**, 565 (1997).
12. L. V. Gibiansky and S. Torquato, J. Mech. Phys. Solids **45**, 689 (1997).
13. L. Li and N. R. Sottos, J. Appl. Phys. **77**, 4595 (1995).
14. S. Schwarzer and A. Roosen, J. Eur. Ceram. Soc. **19**, 1007 (1999).
15. V. M. Petrov, Zh. Tekh. Fiz. **57**, 2273 (1987) [Sov. Phys. Tech. Phys. **32**, 1377 (1987)].
16. V. Yu. Topolov and A. V. Turik, Pis'ma Zh. Tekh. Fiz. **24** (11), 65 (1998) [Tech. Phys. Lett. **24**, 441 (1998)].
17. V. Yu. Topolov and A. V. Turik, J. Electroceram. **3**, 347 (1999).
18. V. Yu. Topolov and A. V. Turik, Pis'ma Zh. Tekh. Fiz. **27**, 81 (2001) [Tech. Phys. Lett. **27** (2001)].
19. W. Wersing, in *Proceedings of the 6th International Symposium on Applications of Ferroelectrics, ISAF'86, New York, 1986*, p. 212.
20. H. Banno, Ceram. Bull. **66**, 1332 (1987).
21. H. Banno, Jpn. J. Appl. Phys., Part 1 **32** (9B), 4214 (1993).
22. C.-W. Nan, J. Appl. Phys. **76**, 1155 (1994).
23. M. Dunn, J. Appl. Phys. **78**, 1533 (1995).
24. V. I. Aleshin, Candidate's Dissertation (Rostov. State Univ., Rostov-on-Don, 1990).
25. V. Yu. Topolov and A. V. Turik, J. Phys. D **33**, 725 (2000).
26. F. Levassort, M. Lethiecq, D. Certon, and F. Patat, IEEE Trans. Ultrason., Ferroelectr. Freq. Control **44**, 445 (1997).
27. *Landolt-Börnstein: Zahlenwerte und Funktionen aus Naturwissenschaften und Technik. Neue Serie* (Springer-Verlag, Berlin, 1984; 1990), Gr. III, Bd. 18; Bd. 28.
28. E. I. Bondarenko, V. Yu. Topolov, and A. V. Turik, Kristallografiya **37**, 1572 (1992) [Sov. Phys. Crystallogr. **37**, 852 (1992)].

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1101–1105. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 33–37.

GAS DISCHARGES, PLASMA

High-Pressure-Induced Baroelastic Effects and Martensitic Transformations in Two-Layer Micro- and Nanocomposites

G. A. Malygin

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

e-mail: malygin.ga@pop.ioffe.rssi.ru

Received December 28, 2000

Abstract—The theory of smeared diffusionless martensitic transitions is applied to analyzing martensitic transformation and relaxation of baroelastic stress in a thin shape-memory alloy layer included in a two-layer microcomposite. When omnidirectional pressure is applied to the composite, baroelastic stress arises in the alloy because of the different bulk compression moduli of the alloy and the substrate material. Baroelastic strain in the microcomposite undergoing martensitic transformation is found to acquire nonlinear and hysteretic properties, which can be used in pressure microtransducers and special-purpose miniature actuators. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In [1], martensitic transformation in a constrained shape-memory alloy (SMA) [2] is treated within the theory of smeared martensitic transitions. The constraint arises, for example, in a two-layer microcomposite that has the form of a narrow strip and consists of a thin (50–10^3 nm thick) SMA layer applied on a ≈100-μm-thick substrate that does not undergo structural transformation under the given conditions [3]. The difference in the thermal expansion coefficients of the composite components causes thermoelastic stress in a composite when its temperature varies. The theory of smeared martensitic transitions [4, 5] makes it possible to self-consistently calculate the stress relaxation through martensitic transformation taking place in the thin layer under temperature (and stress) variations and thus to determine the strain in the strip [1].

Recently, much attention has been given to thin-film two- and multilayer microcomposites, since they offer promise for microtransducers and microactuators in microelectromechanical systems (MEMS) [6]. SMAs are usually employed in these systems as active temperature [6–9], stress [5, 8], pressure [10, 11], and field- (magnetic and electric) sensitive elements [12, 13]. A variety of external and internal factors that can control SMA properties provides the basis for designing smart multifunctional microdevices with intriguing characteristics resulting from reversible martensitic transformations in their active elements.

It would be appropriate to calculate the characteristics of these devices using a theory of diffusionless martensitic transformations. The theory of smeared martensitic transitions [4, 5] seems to be the most suitable and efficient tool for the calculation. In this study, it will be used to investigate the deformation of the two-layer microcomposite under high pressure. Nonlinear characteristics exhibited by microcomposites with a thin SMA layer can be used to develop pressure microtransducers and special-purpose microactuators.

Section 1 of this paper contains the basic formulas of the theory of smeared martensitic transitions, which are necessary to calculate the martensitic relaxation of the baroelastic stress that arises in SMAs from the difference in the bulk compression moduli of the alloy and the substrate material. In Section 2, the results of calculation are outlined, and Section 3 is devoted to relaxation-induced size effects.

1. MATHEMATICAL BACKGROUND

We will consider a microcomposite in the form of a narrow strip of length l and width $w \ll l$ composed of a thin SMA layer of thickness h and a substrate of thickness $H \gg h$. The substrate does not undergo the transformations in a given temperature range. Because of the difference in the bulk compression moduli K_i (i = 1.2) of the composite components, placing the strip in a high-pressure vessel generates a baroelastic stress $\sigma_P = (1/3)Y_1\Delta\varepsilon$ in the thin SMA layer. Here, $\Delta\varepsilon = (K_1^{-1} - K_2^{-1})P$ is the difference in the volume compressive strains in the thin SMA layer and the substrate under pressure P; $Y_1 = E_1/(1 - \nu_1)$; E_1, K_1, and ν_1 are the Young modulus, the bulk compression modulus, and Poisson's ratio of the SMA layer; K_2 is the bulk compression modulus of the substrate. With the assumption that both materials are isotropic, the baroelastic stress is

given by

$$\sigma_P = \delta_K P, \quad \delta_K = \frac{1}{3}\left(\frac{K_2 - K_1}{K_2}\right)\frac{Y_1}{K_1}. \tag{1}$$

Similarly to the case of thermoelastic stress [12, 13], the strip with baroelastic stress bends to a radius of curvature of R and its edges deflect from the initial plane by the distance z:

$$R = \frac{H^2}{6h}\left(\frac{Y_2}{\sigma_P}\right), \quad z = \frac{l^2}{8R} = \frac{3h}{4}\left(\frac{l}{H}\right)^2\left(\frac{\sigma_p}{Y_2}\right), \tag{2}$$

where $Y_2 = E_2/(1 - \nu_2)$, E_2, and ν_2 are the respective parameters of the substrate.

As follows from (2), the deflection of the strip edges increases directly with baroelastic stress and, hence, with pressure P.

A rise in pressure (and stress) may initiate martensitic transformation and thereby the relaxation of the baroelastic stress in the thin alloy. The reduction of the stress, in turn, changes the amount of martensite in the alloy. To perform the self-consistent calculation of the pressure and temperature dependences of the baroelastic stress, as well as the magnitude of its relaxation, we will make use of the theory of smeared martensitic transitions [1, 4, 5].

According to the theory, for given temperature T, stress σ, and pressure P, the relative volume fraction of martensite φ_M in the material is defined as

$$\varphi_M(T, \sigma, P) = \left[1 + \exp\left(\frac{\Delta U}{kT}\right)\right]^{-1}. \tag{3a}$$

Here, $\Delta U = \omega \Delta u$; ω is the elementary volume of transformation, which depends on the bulk density of obstacles that hinder the motion of the phase boundaries;

$$\Delta u = q\frac{T - T_{c0}}{T_{c0}} - \xi(m\sigma \pm \tau_f) - \delta_0 P \tag{3b}$$

is the change in the internal energy of the SMA unit volume as the alloy passes from the austenitic to the martensitic state; q and T_{c0} are, respectively, the heat of transition and the critical (characteristic) temperature of transition at $\sigma = \tau_f = P = 0$; ξ is the spontaneous shear deformation of the lattice undergoing martensitic transformation; m is the crystallographic orientational factor of the martensite type most favorable for relaxation of the baroelastic stress; τ_f is the dry friction stress, which arises upon moving the phase boundaries and specifies the force hysteresis of the transformation; and δ_0 is the lattice dilatation upon transformation.

Expressions (3) suggest that the amount of martensite in the alloy is a function of both the absolute value and the sign of the energy Δu. Therefore, it also depends on the crystal temperature, as well as on the stress and pressure applied. Austenite (martensite) dominates in the crystal if $\Delta u > 0$ ($\Delta u < 0$). From (3), the pressure range (smearing) of the transformation $\Delta P_{M0} = |d\varphi/dP|^{-1}_{P = P_{c0}}$, its hysteresis ΔP_{f0}, and the critical pressure P_{c0} can be expressed at $\sigma = 0$ as

$$\Delta P_{M0} = \left(\frac{q}{\delta_0}\right)\frac{4}{B}, \quad \Delta P_{f0} = \pm\frac{\xi}{\delta_0}\tau_f, \quad P_{c0} = \left(\frac{q}{\delta_0}\right)\frac{T - T_{c0}}{T_{c0}}. \tag{4}$$

Here, $B = q\omega/kT_{c0}$ is the structure-sensitive parameter defining the smearing of the transition. The curves for the direct and reverse martensitic transformations proceeding in free conditions ($\sigma = 0$) under the action of the dimensionless pressure P/K_1 are shown in Fig. 1a. The curves were calculated according to formulas (3) with the dimensionless parameters $B = 50$, $T/T_{c0} = 1.15$, $c = (\delta_0/q)K_1 = 1.5$, $P_{c0}/K_1 = 0.1$, $\Delta P_{f0}/K_1 = \tau_f/c\tau_M = 0.013$, $\tau_f/\tau_M = 0.02$, and $\tau_M = q/\xi$.

2. MARTENSITIC RELAXATION OF BAROELASTIC STRESS

The reversible shear strains attendant to martensite formation, $\varepsilon_M = \varepsilon_m\varphi_M$, where $\varepsilon_m = m\xi$, cause the relaxation of the baroelastic stress in the thin SMA layer by $\sigma_M(T, \sigma, P) = Y_1\varepsilon_m\varepsilon_M(T, \sigma, P)$. Therefore, in view of (1), we can write for the current stress in the layer:

$$\sigma = \sigma_P(P) - \sigma_M(T, \sigma, P). \tag{5}$$

Since both parts of Eq. (5) contain the stress σ, a self-consistent solution should be found for every value of pressure P. It is convenient to reduce (5) to the equation for the part of the stress that has relaxed:

$$\sigma_M(P) = Y_1\varepsilon_m\left(1 + \exp\left(B\left(\frac{T - T_{c0}}{T_{c0}} - \frac{\delta_0}{q}P - \frac{\varepsilon_m}{q}\left(\delta_K P - \sigma_M(P) \pm \frac{\tau_f}{\tau_M}\right)\right)\right)\right)^{-1}. \tag{6a}$$

Introducing the dimensionless stress $S_M = |\sigma_M|/Y_1|\varepsilon_m|$, pressure $p = P/K_1$, and temperature $t = T/T_{c0}$, and putting $\delta_K = -|\delta_K|$ in (6a), one can rearrange Eq. (6a) to the form

$$S_M(p) = \left(1 + \exp\left(B\left(t - 1 - (c - a)p - b(S_M(p) - S_{M0}) \pm \frac{\tau_f}{\tau_M}\right)\right)\right)^{-1}, \tag{6b}$$

where

$$a = \frac{\varepsilon_m}{q}|\delta_K|K_1, \quad b = \frac{\varepsilon_m^2}{q}Y_1,$$
$$\frac{a}{b} = \frac{|\delta_K|}{\varepsilon_m}\left(\frac{K_1}{Y_1}\right), \quad S_{M0} = S_M(0). \tag{6c}$$

The graphical solution of Eq. (6b) is illustrated by Fig. 2, where straight line *1* shows the dependence $L(S_M)$ on the left-hand side of the equation and curves *2–4* depict the functions $R(S_M)$ on its right-hand side for different dimensionless pressures with $B = 50$, $c = 1.5$, $a = 0.7$, $b = 0.07$, and $\tau_f = 0$. The intersection of line *1* with curves *2–4* yields the amount of the martensitic relaxation of the baroelastic stress under a given pressure.

The solutions of Eq. (6b) for direct and reverse martensitic transformations are presented in Fig. 1b by curves *1* and *2*, respectively. Bearing in mind that $S_M(p) \equiv \varphi_M(p)$, one can treat the curves as the pressure dependences of the amount of martensite emerging in the constrained alloy during transformation. When Fig. 1b is compared with Fig. 1a, it is apparent that under constraint (1), the curves $\varphi_M(p)$ shift to higher pressures; (2) the martensitic transformation occupies a narrower temperature range, i.e. the smearing decreases; and (3) the hysteresis loop widens.

In fact, the differentiation of Eq. (6b) with respect to p yields the dimensionless values of smearing Δp_M, hysteresis $2\Delta p_f$, and critical (characteristic) pressure p_c of the transition in the following form:

$$\Delta p_M = \frac{4 - bB}{(c-a)B}, \quad \Delta p_f = \pm\frac{1}{c-a}\left(\frac{\tau_f}{\tau_M}\right),$$
$$p_c = \frac{t - 1 - 0.5b}{c-a}. \tag{7}$$

Putting $a = b = 0$ in (7), we arrive at the expressions for martensitic transformation under free conditions (4).

According to (5), the dimensionless total stress developing in the thin SMA layer at a given pressure is defined as

$$\frac{\sigma(p)}{Y_1\varepsilon_m} = S(p), \quad S(p) = \frac{a}{b}p - S_M(p). \tag{8}$$

Figure 3 illustrates the associated dependence for direct and reverse martensitic transformations (curves *1* and *2*, respectively). It is evident that microcomposites with the active SMA layer acquires nonlinear and hysteretic properties, which are absent in those without the SMA layer.

One more point to mention is that in the pressure range of martensitic transformation, the sensitivity of stress to pressure changes the sign. Differentiating

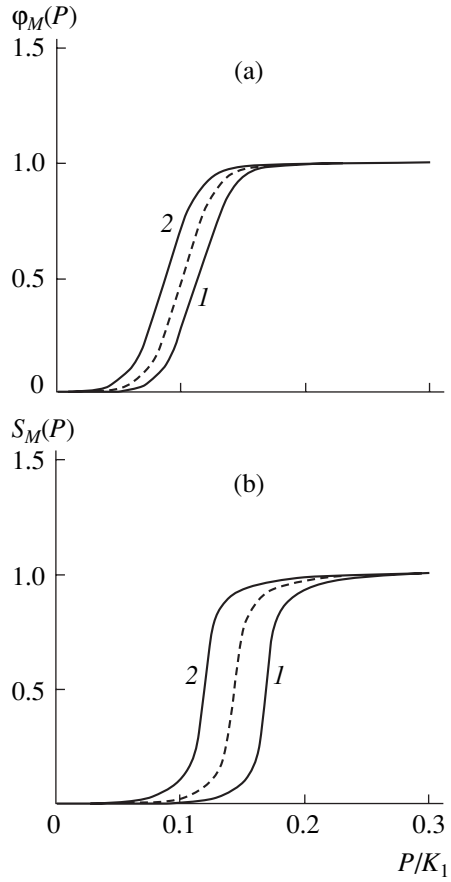


Fig. 1. Pressure dependences of the relative volume fraction of martensite at (*1*) direct and (*2*) reverse martensitic transitions under (a) free conditions and (b) under constraint. The dashed curves depict the transformation in the absence of hysteresis.

Eqs. (5) and (6a) with respect to P yields the sensitivity (at $P = P_c$)

$$\left.\frac{d\sigma}{dP}\right|_{P=P_c} = \frac{|\delta_K| - \delta_0\varepsilon_m(Y_1/4q)B}{1 - \varepsilon_m^2(Y_1/4q)B}. \tag{9}$$

In the absence of martensitic transformation ($\varepsilon_m = 0$), the sensitivity depends only on δ_K, given by (1), whereas during the martensitic relaxation of the baroelastic stress, it is governed by the combination of the parameters δ_0, ε_m, and B, which are related to the thermodynamics and the kinetics of the transition.

3. SIZE EFFECTS

In [3], the size effect observed in a thin NiTi alloy during the martensitic relaxation of thermoelastic stress was associated with the effect of the layer thickness h on the relaxation parameters. As the thickness decreased from 1 μm to 50 nm, the transformation became more and more smeared and the degree of stress relaxation progressively lowered. It was argued [1] that the size effects may arise because the elemen-

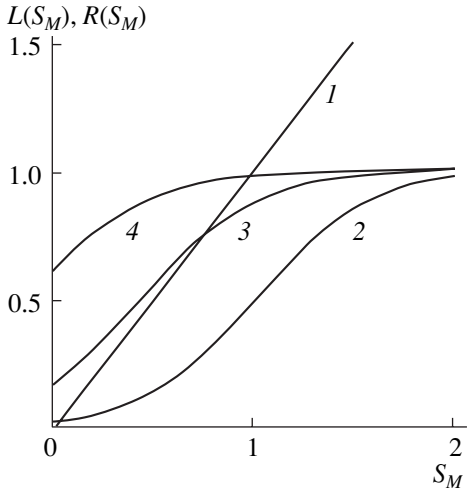


Fig. 2. Dependences of (*1*) left-hand and (*2–4*) right-hand sides of Eq. (6b) on the stress $S_M = \sigma_M/Y_1\varepsilon_m$ for the pressure P/K_1 = (*2*) 0.1, (*3*) 0.15, and (*4*) 0.2.

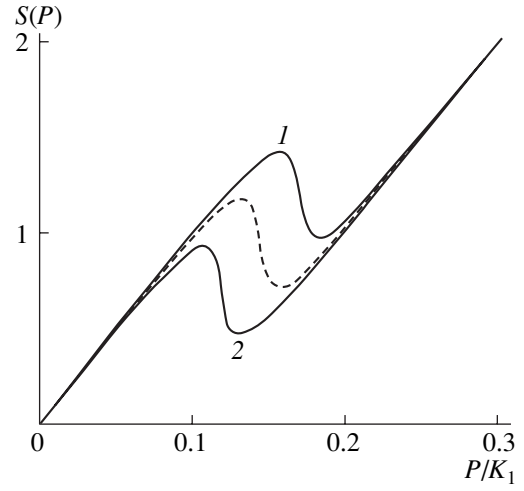


Fig. 3. Pressure dependences of the baroelastic stress for (*1*) direct and (*2*) reverse martensitic transformations. The dashed curve corresponds to the transformation in the absence of hysteresis.

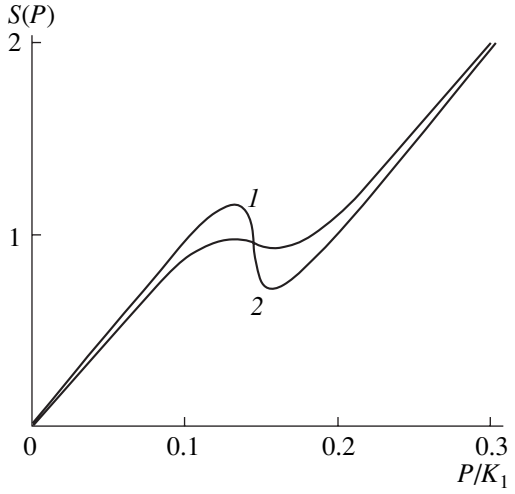


Fig. 4. Pressure dependences of the baroelastic stress for SMA layers with the thickness h/λ_m (*1*) $\gg$ 1 and (*2*) = 1.5.

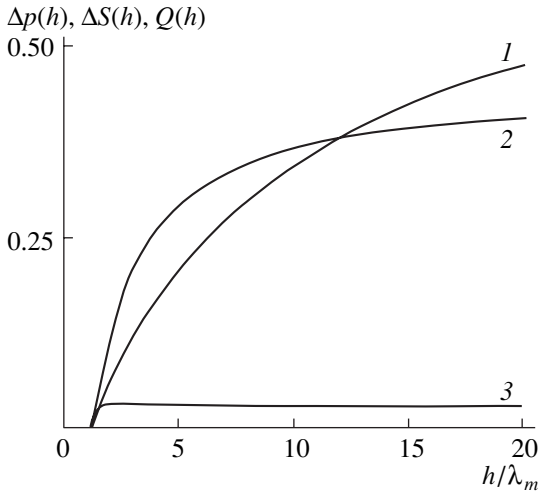


Fig. 5. Dependences of (*1*) the pressure range of stress relaxation, (*2*) the amount of stress relaxation, and (*3*) the coefficient Q on the layer thickness.

tary volume of transformation ω becomes dependent on the alloy thickness when the latter is comparable to the average distance λ_m between obstacles limiting the mobility of the phase boundaries. Then, for the parameter $B \sim \omega$, we have [1]

$$B(h) = B_m \frac{h/\lambda_m}{1 + h/\lambda_m}, \qquad (10)$$

where B_m is the value of B in a thick ($h \gg \lambda_m$) layer.

The pressure dependences of the baroelastic stress for thick ($B = B_m = 50$) and thin ($B = 30$, $h/\lambda_m = 1.5$) layers are presented in Fig. 4 (curves *1* and *2*, respectively). It is seen that in the thinner layer, the sensitivity of stress to pressure in the region of martensitic relaxation is smaller. Taking into account that $d\sigma/dP = (Y_1\varepsilon_m/K_1)(dS/dp)$, the sensitivity can be expressed as

$$\left.\frac{dS}{dp}\right|_{p = p_c} = Q(h) = \frac{4a - bcB(h)}{b[4 - bB(h)]}. \qquad (11)$$

The dependence of $Q(h)/Q_m$, where Q_m is the value of $Q(h)$ in the thicker layer, is shown in Fig. 5 by curve *3*.

As follows from (11), the sensitivity $Q < 0$ when $4 > bB > 4a/c$, goes to zero at $B = B_c = 4a/bc$, and may become positive in a very thin layer. In the opposite case of a relatively thick layer, Q also becomes positive when the conditions $bB > 4$ and $bB > 4a/c$ are met simultaneously. In this situation, in addition to the main hysteresis observed in the $S(p)$ curves (Figs. 3, 4), a

local loop should appear in the region of martensitic transformation, still further underlying the bistability of the curves.

Let us find the amount of martensitic relaxation of the baroelastic stress, i.e., the stress difference $\Delta\sigma = \sigma_1 - \sigma_2$, where $\sigma_1 = \sigma(P_1)$, $\sigma_2 = \sigma(P_2)$, and P_1 and P_2 correspond, respectively, to the maximum and minimum of the curve $\sigma(P)$. Differentiating Eq. (8) with respect to dimensionless pressure p in view of (6b) and equating the derivative of the total stress to zero, $dS/dp = 0$, we finally obtain the following system of two equations for dimensionless stresses S_1 and S_2 and pressures p_1 and p_2:

$$S_{1,2} = \left[\frac{a}{b}\left(t - 1 \pm \frac{\tau_f}{\tau_M}\right) - \frac{c}{1+z_{1,2}} - \frac{a}{bB}\ln\frac{z_1}{z_2}\right]\frac{1}{c-a}, \quad (12a)$$

$$p_{1,2} = \left[t - 1 \mp \frac{\tau_f}{\tau_M} - \frac{b}{1+z_{1,2}} - \frac{1}{B}\ln\frac{z_1}{z_2}\right]\frac{1}{c-a}, \quad (12b)$$

where

$$z_{1,2} = \frac{1}{2}\alpha - 1 \pm \left[\left(\frac{1}{2}\alpha - 1\right)^2 - 1\right]^{1/2}, \quad \alpha = \frac{bc}{a}B. \quad (12c)$$

According to (12), the amount of stress relaxation ΔS and the pressure range Δp where it occurs are

$$\Delta S = S_1 - S_2 = \frac{1}{\alpha}\left[(\alpha(\alpha-4))^{1/2} - \ln\frac{z_1}{z_2}\right]\frac{c}{c-a}, \quad (13a)$$

$$\Delta p = p_2 - p_1 = \frac{1}{\alpha}\left[\frac{c}{a}\ln\frac{z_1}{z_2} - (\alpha(\alpha-4))^{1/2}\right]\frac{b}{c-a}. \quad (13b)$$

Since $B = B(h)$, Eq. (12c) implies that $\alpha = \alpha(h)$; therefore, both the amount of relaxation and the associated pressure range depend on the SMA layer thickness. The functions ΔS and Δp [see (13)] are plotted in Fig. 5 (curves *1* and *2*, respectively).

In conclusion, we have demonstrated that, as in the case with thermoelastic stress [1], the theory of smeared martensitic transitions allows the calculation of the deformation characteristics of microdevices based on a microcomposite with a thin SMA layer subjected to high pressure. With the parameters of martensitic transformation shown in Figs. 3 and 4, the operating pressure range where the microcomposite can be used as a pressure microtransducer or microactuator is $(0.1–0.2)K_1 \approx 1–20$ GPa.

REFERENCES

1. G. A. Malygin, Fiz. Tverd. Tela (St. Petersburg) **43**, 822 (2001) [Phys. Solid State **43**, 854 (2001)].
2. K. Shimizu and K. Otsuka, *Shape Memory Effects in Alloys* (Nauka, Moscow, 1979) (translated from Japanese).
3. A. L. Roytburd, T. S. Kim, Q. Su, *et al.*, Acta Mater. **46**, 5095 (1998).
4. G. A. Malygin, Fiz. Tverd. Tela (St. Petersburg) **36**, 1489 (1994) [Phys. Solid State **36**, 815 (1994)].
5. G. A. Malygin, Zh. Tekh. Fiz. **66** (11), 112 (1996) [Tech. Phys. **41**, 1145 (1996)].
6. S. M. Spearing, Acta Mater. **48**, 179 (2000).
7. *Materials for Smart Systems II*, Ed. by E. P. George, R. Gotthardt, K. Otsuka, *et al.* (Materials Research Society, Pittsburg, 1997), Vol. 459.
8. J. E. Bidaux, W. J. Yu, R. Gotthardt, and J. A. Manson, J. Phys. IV **5** (C2), 453 (1995).
9. S. P. Belyaev, S. A. Egorov, V. A. Likhachev, and O. E. Ol'khovik, Zh. Tekh. Fiz. **66** (11), 36 (1996) [Tech. Phys. **41**, 1102 (1996)].
10. T. Kikashita, T. Saburi, K. Kindo, and S. Endo, Jpn. J. Appl. Phys. **36**, 7083 (1997).
11. R. D. James and K. F. Hane, Acta Mater. **48**, 197 (2000).
12. G. G. Stoney, Proc. R. Soc. London, Ser. A **82**, 172 (1909).
13. L. E. Andreeva, *Elastic Elements in Devices* (Mashinostroenie, Moscow, 1981).

Translated by A. Sidorova-Biryukova

Technical Physics, Vol. 46, No. 9, 2001, pp. 1106–1111. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 38–44.

GAS DISCHARGES, PLASMA

Interaction of a Long Pulsed High-Density Electron Beam with a Jet of Solid-Target Destruction Products

A. V. Arkhipov and G. G. Sominski

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

e-mail: arkhipov@twonet.stu.neva.ru, sominski@twonet.stu.neva.ru

Received December 6, 2000

Abstract—Results are presented from experimental studies of the destruction of a solid target by a high-density nonrelativistic electron beam at a deposited power density of 20 MW/cm^2 in millisecond pulses. Results of studying beam transportation under these conditions are presented as well. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

High-density electron beams are traditionally used for the transportation of energy to solid objects and for controlled energy release in the surface region. Besides classical applications of electron beams, such as electron-beam welding, cutting, and nonthermal hardening of metals and alloys, we can mention attempts using electron beams in nuclear fusion, the excitation of shock waves, the generation of intense X radiation, etc. Practical applications stimulated intensive studies in different ranges of the power density transported by the electron beam. Note, however, that the beams with extreme parameters have been studied most extensively. One of these cases is typical of technological applications: the relationship between the deposited energy density and the area of the irradiated region is such that, during the melting and evaporation of the material to be processed, the vapor density in front of the target is insufficiently high to significantly hamper electron transport toward the target surface [1–3]. On the other hand, in inertial fusion research [4, 5], the typical power densities are so high that the target material is completely and rapidly (in several nanoseconds or even faster) converted into an overdense, highly ionized plasma that absorbs the electron beam energy.

The processes occurring at intermediate beam power densities have been studied less intensively. The practical necessity of studying these processes arose in the early 1990s in connection with some new applications, such as the electronic ablation acceleration of macroparticles up to extremely high velocities [6, 7] (in particular, for the power supply and diagnostics in large tokamaks [8, 9]) and the modeling of the processes occurring on the first walls and divertors of such tokamaks during discharge disruptions [10, 11]. In these applications, the beam power density is typically on the order of several units to several tens of MW/cm^2 at a nonrelativistic electron energy in millisecond current pulses. Under these conditions, a layer consisting of vapor and other destruction products of the surface material is formed in front of the surface. This layer can substantially attenuate the electron beam. The vapor, to a greater or lesser extent, can be ionized thermally or via individual interactions of atoms with the beam electrons. Due to the long duration of irradiation pulses, the role of slow processes (such as heat conductivity, diffusion, and expansion of the vapor into a vacuum) increases substantially. Since the theoretical description of nonequilibrium systems including a solid body, an electron beam, and a multicomponent plasma is rather complicated, an experimental study of the interaction between a solid target and an electron beam in the given parameter range is of great importance.

EXPERIMENTAL TECHNIQUES

Experiments were carried out on the EPVP device of St. Petersburg State Technical University [12–15]. A schematic diagram of the experimental device is shown in Fig. 1. An electron beam with an electron energy up to 45 keV and a current up to 5 A was produced with a Pierce gun, transported through the drift channel in a pulsed magnetic field of a solenoid, and then deposited on a solid target in a mode of long (from tens to thousands of microseconds) single pulses.

One of the most serious engineering problems that we faced in our study was to ensure the electric strength of the high-voltage gap of the electron gun in the presence of vapor and plasma produced due to the intense destruction of the target. The solution to this problem was found empirically.

In the first version of the device, the distribution of the magnetic field $B(z)$ near the electron gun was controlled by an individual solenoid placed near the anode plane. The aim of the optimization was to match the electric and magnetic field lines in order to minimize electron losses at the entrance to the transport channel. However, it turned out that with this matching, the

1063-7842/01/4609-1106$21.00

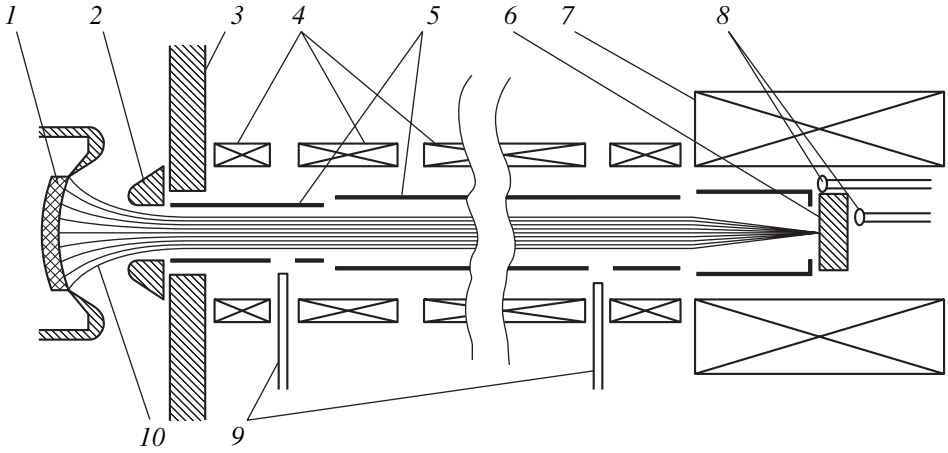


Fig. 1. Schematic diagram of the EPVP device: (*1*) cathode and (*2*) anode of the electron gun, (*3*) magnetic screen, (*4*) solenoid of the transport channel, (*5*) sections of the transport channel, (*6*) target, (*7*) solenoid of the target, (*8*) X-ray detectors, (*9*) optical detectors (fibers), and (*10*) electron beam.

problem of the electric strength of the high-voltage gun gap was aggravated. In contrast, if the magnetic field of the gun solenoid compensated for the field penetrating into the gun from the transport-channel solenoid, the electric breakdown occurred at higher values of the energy load on the target. This observation showed that the magnetically shielded gun was better protected from the penetration of the plasma from the target side: it was protected by the transverse (radial) components of **B** in the transient region, where the magnetic field lines arrived at the walls of the transport channel, thus guiding the expanding plasma to it. For this reason, when modifying the experimental device, we placed a 40-mm-thick ferromagnetic shield with a large-diameter aperture (20 mm) between the gun and the transport channel. This ensured a relatively slow increase in the magnetic field along the *z* axis and, consequently, a fairly long transient region. This innovation enhanced the electric strength of the system, although it posed some additional problems: an appreciable fraction of the beam could be lost at the entrance to the channel. We also protected the gun from the neutral component of the jet of the destruction products of the target. The target was placed a large distance away (about 1 m), which allowed us, with refractory targets, to substantially decrease the density of the vapor neutral component near the gun.

Due to the described modifications of the experimental device, the proper choice of the optimum operating modes, and the training of the components of the electronic-optical system that underwent the intense bombardment (see [14, 15] for details), we prevented breakdowns in the electron gun. The maximum duration of the beam-current pulse (~5 ms) was only limited by the parameters of the solenoid power supply system. The electrostatic focusing in the gun in combination with an additional compression by the magnetic field made it possible to compress the area of the electron beam by a factor of ~1500. Thus, at rather moderate levels of the beam current and accelerating voltage, we could obtain high values of the specific energy characteristics: the power density deposited on the target attained 20 MW/cm^2, and the total energy density per pulse attained tens of kJ/cm^2.

To obtain information on the parameters of the electron beam produced and its interaction with a medium formed in the device volume, we measured the electrode currents, signals from optical detectors positioned at different points of the drift channel, and signals from detectors of soft X radiation emitted from the vicinity of the target. The dynamics of target destruction was studied by measuring the shape and dimensions of craters (Fig. 2) produced by single current pulses of different duration τ_0. For this purpose, the irradiated region was displaced over the sample surface before every shot by varying the inclination of the axis of solenoid *7* (Fig. 1).

Most of the results presented below were obtained for graphite targets. This choice is motivated by the fact that carbon materials are commonly used in the production of electrodes undergoing high energetic loads and also by the convenience of operation with this material: graphite is nontoxic, does not produce irreversible poisoning of cathodes, has a low vapor pressure, and does not form a liquid phase.

EXPERIMENTAL RESULTS AND DISCUSSION

1. Dynamics of the Target Destruction in the Low-Current Mode

In the course of experiments, we measured the dependence of the mass *M* of the material lost by the target during one electron-irradiation pulse on the total energy transported by the electron beam:

$$w = U_0 I \tau_0, \quad (1)$$

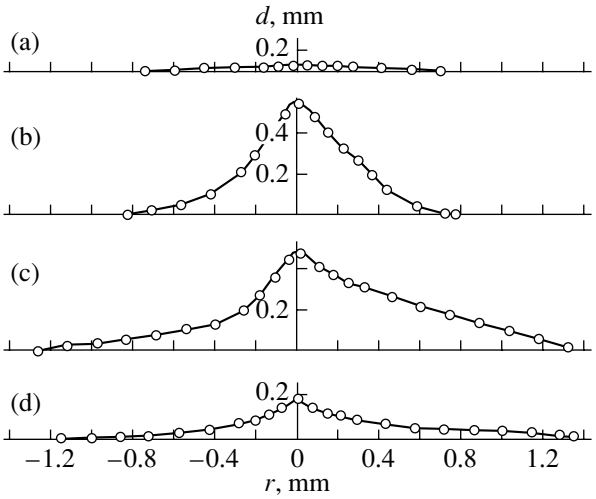


Fig. 2. Typical shapes of the craters produced by a pulsed electron beam in different sublimation regimes for U_0 = 36 kV and B = 0.8 T: (a) the irradiation dose is insufficient for the beginning of sublimation (I = 0.8 A, τ_0 = 0.24 ms, w = 6.9 J, and M = 31 μg); (b) the beam current is below the discharge threshold (I = 0.38 A, τ_0 = 1.6 ms, w = 21.6 J, and M = 323 μg); (c) the target is screened by the discharge (I = 0.96 A, τ_0 = 1.6 ms, w = 56.2 J, $\eta_i \approx 1.8$, and M = 908 μg); (d) the degree of target screening is substantial (I = 2.6 A, τ_0 = 1.45 ms, w = 135 J, $\eta_i \approx 10$, and M = 240 μg).

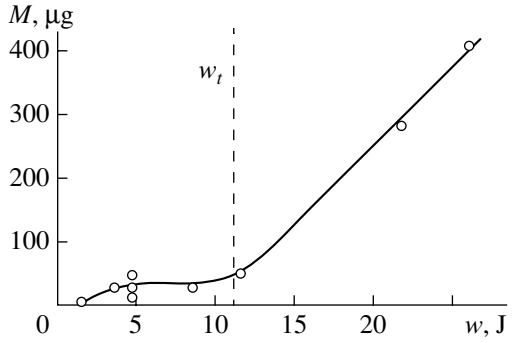


Fig. 3. Mass of material evaporated from the graphite target vs. the total energy of the pulsed electron beam in the absence of a discharge.

where U_0 is the voltage accelerating the electrons and I is the current of the beam fraction reaching the target. At given values of U_0 and I, the energy of the beam was varied by varying the pulse duration.

The destruction of the graphite target begins at very low values of the energy deposited on its surface (~10 J/cm^2). In a typical plot of the dependence $M(w)$ (Fig. 3), we can notice two almost linear segments separated by a characteristic bend at $w = w_t$. At $w < w_t$, the thickness of the evaporated layer of the target material d_p depends slightly on the irradiation dose and is equal to 5–30 μm for the entire surface irradiated by electrons (Fig. 2a). This value is comparable with the electron penetration depth into the target. After substituting the numerical values corresponding to the experiment, the empirical formulas [3] give ~5 and ~15 μm for the mean and maximum penetration depths, respectively. As is seen from the relief of the graphite surface, the material leaves the target in the form of grains (monocrystals). We believe that, at low deposited energy densities, the loss of the target material is caused by the cracking of the surface layer due to nonuniform thermal expansion [11]: only at $w > w_t$ does the material begin to evaporate, which is reflected as a bend in the $M(w)$ dependence. Intense sublimation occurs only after the surface layer of thickness d_t becomes heated to the critical temperature: for graphite, this temperature is equal to ~4000 K [16]; for metals, this is the temperature of the triple point. This occurs at

$$w = w_t \approx d_t S \rho c_p \Delta T, \tag{2}$$

where S is the irradiated area, ρ is the mass density of the target material, c_p is its specific heat, and ΔT is the difference between critical and room temperature.

Determining the value of the threshold dose ω_t from the plots, we can estimate d_t. The values obtained in this manner ($d_t \geq 100$ μm) substantially exceed the calculated electron penetration depth and depend on the deposited power density. They are close to the spatial scale length characteristic of thermal conductivity $(\sigma_t \tau)^{1/2}$, where σ_t is the thermal diffusivity and τ is the time needed to reach the dose w_t for given parameters of the electron beam.

The slope of the dependence $M(w)$ at $w > w_t$ corresponds to the specific energy spent on graphite destruction, ~40 kJ/g (or 5 eV/atom), which agrees well with the results of the thermodynamic analysis performed in [17, 18]. In particular, it was shown there that, soon after the beginning of evaporation, a time-independent distribution of all quantitative characteristics of the vapor–solid system (including the temperature) is established in the reference frame related to the phase boundary. This occurs because the propagation velocity of the heat conduction wave becomes equal to the velocity with which the phase boundary is displaced. As a result, the heat conduction ceases to affect the value of the specific energy spent on sublimation and the material (graphite) mass evaporated per unit time is related to the power P of the energy source (for an electron beam, this is $U_0 I$) by a simple equation:

$$dM/dt = P/(\Delta H_{\text{subl}} + c_p \Delta T) \tag{3}$$

or

$$M = (w - w_t)/(\Delta H_{\text{subl}} + c_p \Delta T), \tag{4}$$

where H_{subl} is the specific sublimation energy of graphite.

Since the term determining the energy spent on sublimation is dominant, it is expected that

$$dM/dw \approx 1/\Delta H_{\text{subl}}. \tag{5}$$

However, for graphite, the value of ΔH_{subl} is not uniquely determined [11]: it is equal to 59 kJ/g for atomic carbon sublimation, and it is only 30 kJ/g for sublimation in the form of three-atom aggregations. The value of dw/dM obtained in the experiment is nearly the average of these results, probably because of the complicated composition of the formed vapor.

The effect of surface-layer cracking, which causes the destruction of the target at small deposited energies, ceases to be important after sublimation begins. This can be explained, e.g., by the short sublimation time of the destroyed layer of thickness d_p (we assume that $d_p > d_t$)

$$\tau_p \approx d_p S \rho \Delta H_{\text{subl}}/P, \quad (6)$$

which is equal to several microseconds under our experimental conditions. The produced powder has no time to leave the region irradiated by electrons and, until it is completely evaporated, hampers the energy transfer to the subsequent target layers. For this reason, the specific destruction energy does not decrease.

Good agreement between the experimental value of dM/dw and the predictions of thermodynamic theory allows us to conclude that, in this case, we do not observe a substantial screening of the target by the products of its destruction, in spite of the density of the evaporated material in front of the target being very high. This density can be estimated as

$$n_0 = PN_A/\mu S v_j \Delta H_{\text{subl}}, \quad (7)$$

where N_A is Avogadro's number, μ is the material molar mass, and v_j is the vapor jet velocity.

According to [17, 18], the jet velocity depends only slightly on the parameters of the energy source and, over a wide range of parameters, is almost equal to the thermal velocity at the critical temperature. For graphite in the case of atomic sublimation, the mean value of the normal velocity component is $v_j \approx 3.5$ km/s. Near the target surface, at a distance z less than the size of the irradiated region r_0, the vapor density appears to be $n_0 \approx 5 \times 10^{19}$ cm^{-3} for the maximum deposited power density. Based on a simple model of the expansion of the products of target destruction (atomic sublimation, isotropic expansion within a solid angle of 2π, a high adhesion rate of vapor to the surfaces of surrounding electrodes, and the low degree of their ionization that eliminates the influence of the magnetic field), the decrease in the material density with distance from the target at $z > r_0$ can be assumed to be approximately quadratic. In this case, the total amount of the evaporated material in front of the target under steady-state conditions will be equivalent to a solid layer ~1 μm in thickness, which is significantly smaller than the electron penetration depth. Therefore, it is not surprising that the target screening was not observed in the experiment. Note that the violation of any one of the above conditions would lead to a stronger interaction of the beam with vapor and to a noticeable absorption of its energy, which was not observed.

2. Influence of the Discharge Phenomena on Beam Transportation and on the Dynamics of Target Destruction

Under our experimental conditions, the neutral component of the jet of the target destruction products is insufficiently dense for the parameters of the passing electron beam to change appreciably. However, efficient screening of the target can be detected due to the interaction of the beam with the ionized component of the target material vapor if an RF discharge develops in the system. The value of the threshold beam current I^* causing the RF discharge ignition depends on the electron energy, the magnitude and distribution of the magnetic field (which determines the irradiated area), and the target material. For example, for graphite at $eU_0 =$ 36 keV and $B = 0.8$ T (which corresponds to $r_0 \approx 1$ mm), the current I^* is equal to ~0.8 A, whereas at $B = 0.4$ T ($r_0 \sim 1.5$–2 mm), this current is equal to 1.5 A. Typical waveforms obtained for $I > I^*$ are shown in Fig. 4. During a rather long period of time ($\tau^* = 30$–1000 μs, depending on the experimental conditions), the signals remain constant. Then, the current is redistributed and a new quasi-steady state is established. This state is characterized by two effects: substantial noise components with frequencies from 1 MHz and higher appear in all the current signals, and a great number of slow particles with energies no higher than several electronvolts arrive at the electrodes. The current of positive ions onto the wall of the transport channel can be several times higher than the injected beam current. The slow plasma electrons escape from the device volume along the magnetic field lines onto the magnetic shield. During the transient process, the detectors of the integral optical radiation indicate the expansion of the high-density plasma along the transport channel at a

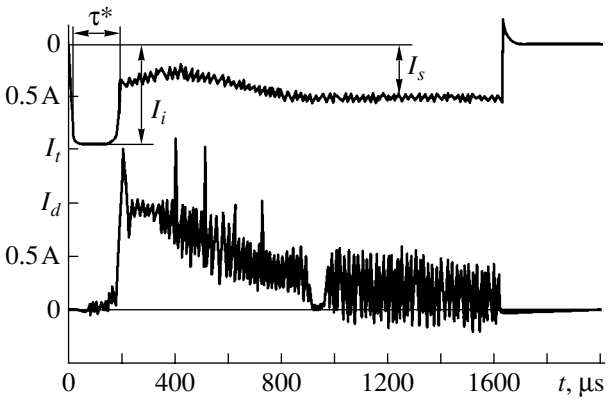


Fig. 4. Typical time dependences of the target current I_t and transport-channel current I_d in the presence of a discharge that screens the target.

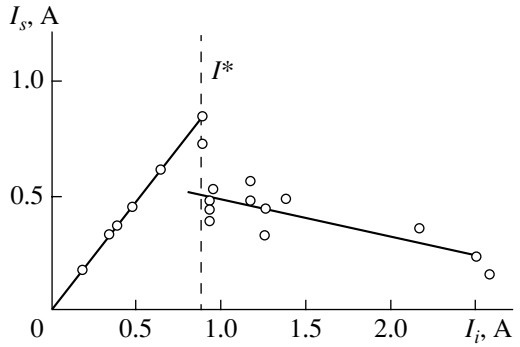


Fig. 5. Steady-state electron-beam current reaching the target, I_s, vs. the injected current I_i. In the absence of a discharge (at $I < I^*$), these currents coincide.

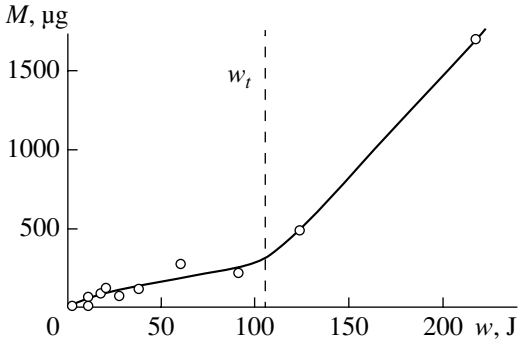


Fig. 6. Dependence $M(w)$ for a lower magnetic field (B = 0.4 T) and the beam current exceeding the threshold discharge current by 5–10%.

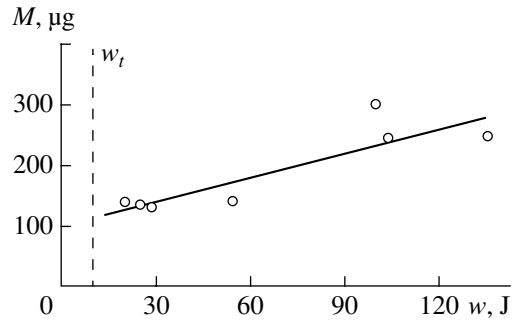


Fig. 7. Dependence $M(w)$ for B = 0.8 T and the beam currents substantially (by a factor of 1.8–2.6) exceeding the threshold discharge current.

velocity of ~10^6 cm/s. The time at which the plasma front arrives at the electron gun nearly coincides with the time at which the system comes to a quasi-steady state.

All these observations allow us to interpret the phenomenon observed as a beam–plasma discharge in the ionized plume of the target-destruction products. The fact that the time at which intense target sublimation begins, w_t/U_0I, does not coincide with the time of discharge ignition τ^* can be explained as follows. For the discharge to be ignited, it is necessary that, in front of the target, a sufficiently extended plasma be formed with the density corresponding to the efficient interaction with the electron beam (on the order of the beam electron density) and with a density gradient that is not too high. For such a plasma to be formed, the vapor should expand with the thermal velocity at least over a distance of several centimeters from the target.

The loss of the beam energy during the beam propagation through the discharge region is substantial and increases with increasing current. Figure 5 shows the steady-state value of the fraction of the beam current reaching the target, I_s, as a function of the electron gun current I_i, which varies only slightly during the pulse. When the injected current exceeds the threshold value I^*, the current that reaches the target under quasi-steady conditions decreases abruptly severalfold and continues to fall as I_i further increases. This means that the current loss in the developing discharge increases with I_i more rapidly than linearly. In the pulses characterized by a high current loss factor $\eta_I = I_i/I_s$, the decrease in the X-ray signal from the vicinity of the target, η_r, is even more pronounced ($\eta_r > \eta_I$), which can be explained by a decrease in the mean energy of electrons reaching the target.

Target screening by the discharge affects the dynamics of target disruption. Figure 6 shows the dependence $M(w)$ obtained at a reduced magnetic field B = 0.4 T and the beam current I_i exceeding the threshold discharge current I^* (which is equal to ~2 A in this case) by nearly 5–10%, which corresponds to the current loss factor η_I = 1.5–2. To determine the irradiation dose w, we used the value of the injected current I_i. The form of the dependence is similar to that shown in Fig. 3, but its slope at $w > w_t$ corresponds to the value $dw/dM \approx$ 80 kJ/g; i.e., the sublimation efficiency reduces by one-half. Figure 7 shows similar data obtained at B = 0.8 T for $1.5I^* < I < 2.6I^*$. As compared to the case illustrated in Fig. 3, the specific sublimation energy dw/dM at $w > w_t$ increases substantially (by a factor of 10–20).

Another specific feature of beam transportation in the presence of a discharge was the increase in both the crater size and the degree of its asymmetry (Figs. 2c, 2d); sometimes, the prints even had an annular shape [13]. These effects may be attributed to a lower efficiency of the magnetic confinement of the beam in the presence of the RF field of the discharge. In addition, the electrons at the periphery of the beam have a better chance of reaching the target and transfer a higher energy onto the target surface than the axial electrons, for which the trajectory of motion toward the target passes through the discharge region. We can assume that, using an electron beam with a larger cross section (for the same density and particle energy and an increased total current), we would observe even more efficient screening of the target by the discharge

because of a lower relative contribution from the peripheral regions.

3. Experiments with Metal Targets

In addition to the experiments with graphite described above, we attempted to carry out similar experiments with metal (copper) targets. As was expected, the methods used in the previous experiments for studying the formation of craters were difficult to employ because copper had a liquid phase. The craters formed were surrounded by a parapet whose volume could be several times greater than the volume of the material that left the target, which reduced the accuracy of measuring this volume. In addition, a substantial portion of the material left the target as drops of up to several tenths of a millimeter in size. Such drops were observed through the entire transport channel, which indicates a rather high velocity. Accordingly, the quantity of the material that left the target was several times greater than the estimates from formula (4).

The features of electron-beam transport in experiments with a metal target do not differ qualitatively from those observed for graphite. Here again, when the current exceeded the threshold level I^*, we observed the beam–plasma discharge absorbing a substantial fraction of the initial electron beam energy; however, other conditions being the same, the value of I^* for copper turned out to be lower.

CONCLUSIONS

In summary, one of the main results of this study is the production of an electron beam with high values of differential energy parameters (a power density up to 20 MW/cm^2 and a density of the energy transferred in a single millisecond pulse up to 100 kJ/cm^2), which were sustained under conditions when the beam-forming system was affected by the target-destruction products. The study of the physical processes occurring when such a dense, long, pulsed beam interacts with a target have allowed us to draw the following conclusions:

The steady state of this interaction is reached in tens or even hundreds of microseconds. Such high values of the delay times can be determined by thermal processes in the surface layers of the target or by the motion of particles evaporated from the surface over distances comparable to the characteristic scale length of the problem (e.g., the beam diameter or the length of the device).

The only effect hampering beam propagation toward the target under our experimental conditions is the beam–plasma discharge in the flow of the evaporated material. This effect results in a substantial (up to 90–95%) absorption of the energy of the electron beam.

For the beam current below the threshold discharge current, the beam propagates toward the target without appreciable losses even under conditions such that the evaporation velocity (i.e., the velocity at which the phase boundary of the solid and gas phases propagates into the target) attains 0.1–1 m/s and the vapor density in front of the target attains 3×10^{19} cm^{-3}.

In the interaction of an electron beam with a metal target, the dynamics of the melt filling the crater plays an important role.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 98-02-18323) and by the Ministry of Education of the Russian Federation (a grant from the 1997 Competition in the Field of Nuclear Engineering and Physics of Ionizing Radiation).

REFERENCES

1. V. V. Bashenko, *Electron-Beam Facilities* (Mashinostroenie, Leningrad, 1972).
2. S. Schiller, U. Heisig, and S. Panzer, *Elektronenstrahltechnologie* (Technik, Berlin, 1976; Énergiya, Moscow, 1980).
3. I. A. Abroyan, A. N. Andronov, and A. I. Titov, *Physical Principles of Electronic and Ionic Technology* (Vysshaya Shkola, Moscow, 1984).
4. *Generation and Focusing of High-Current Relativistic Electron Beams*, Ed. by L. I. Rudakov (Énergoatomizdat, Moscow, 1990).
5. V. P. Smirnov, Prib. Tekh. Éksp., No. 2, 7 (1977).
6. J. Linhart, in *The Physics of High Energy Density*, Ed. by P. Caldirola and H. Knoepfel (Academic, New York, 1971; Mir, Moscow, 1974).
7. F. V. Bunkin and A. M. Prokhorov, Usp. Fiz. Nauk **119**, 425 (1976) [Sov. Phys. Usp. **19**, 561 (1976)].
8. S. L. Milora, J. Vac. Sci. Technol. A **7**, 925 (1989).
9. B. V. Kuteev, Zh. Tekh. Fiz. **69** (9), 63 (1999) [Tech. Phys. **44**, 1058 (1999)].
10. B. Bazylev, L. Landman, and H. Wurz, in *Proceedings of the Conference "Physics and Technique of Plasma," Minsk, 1994*, p. 451.
11. V. Engelko, A. Andreev, O. Komarov, *et al.*, in *Proceedings of the 11th International Conference on High-Power Particle Beams, Prague, 1996*, Vol. II, p. 793.
12. A. V. Arkhipov and G. G. Sominski, Pis'ma Zh. Tekh. Fiz. **20** (11), 6 (1994) [Tech. Phys. Lett. **20**, 431 (1994)].
13. A. V. Arkhipov and G. G. Sominski, in *Proceedings of the 11th International Conference on High-Power Particle Beams, Prague, 1996*, Vol. 2, p. 789.
14. A. V. Arkhipov and G. G. Sominski, in *Proceedings of the 18th International Symposium on Discharges and Electrical Insulation in Vacuum, Eindhoven, 1998*, Vol. 2, p. 762.
15. A. V. Arkhipov and G. G. Sominski, IEEE Trans. Dielectr. Electr. Insul. **6**, 491 (1999).
16. *Handbook of Physical Quantities*, Ed. by I. S. Grigor'ev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991).
17. I. V. Nemchinov, Prikl. Mat. Mekh. **31** (2), 300 (1967).
18. Yu. V. Afanas'ev and O. N. Krokhin, in *The Physics of High Energy Density*, Ed. by P. Caldirola and H. Knoepfel (Academic, New York, 1971; Mir, Moscow, 1974).

Translated by N. Larionova

Technical Physics, Vol. 46, No. 9, 2001, pp. 1112–1115. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 45–48.

GAS DISCHARGES, PLASMA

The Effect of Ultraviolet Irradiation on the Properties of Zr/Si Contacts

I. S. Nuprienok and A. N. Shibko

Institute of Electronics, Belarussian Academy of Sciences, Minsk, 220841 Belarus

e-mail: inel@inel.bas-net.by

Received December 13, 2000

Abstract—The effect of ultraviolet (UV) irradiation on the properties of Zr/Si contacts subjected to heat treatment is investigated. It is established that, due to variations in the phase composition of the contact, its electrophysical parameters are modified. The application of the combined treatment allows one to form Zr/Si contacts with specified properties. © *2001 MAIK "Nauka/Interperiodica".*

Thin-film materials are widely used in semiconductor devices. In modern microelectronics, semiconductor elements based on refractory metals (including zirconium) are considered as promising high-temperature building blocks. They offer the electrophysical properties of the metals and the high melting points of their oxides and silicides. One of the methods modifying the electrophysical parameters of a metal–semiconductor system is heat treatment, which changes electron states at the interface. In polycrystalline films, heat treatment stimulates recrystallization processes, which bring the system to a thermodynamically more equilibrium state. Recently, however, nonthermal processes that occur in metal–semiconductor systems subjected to heat treatment and simultaneously irradiated by a photon beam with a certain energy have attracted much attention [1–3].

In this study, we investigated the phase transformations and modifications of the electrophysical parameters in the Zr/Si system subjected to thermal treatment and simultaneously exposed to UV radiation with λ = 0.312, 0.365, and 0.552 μm.

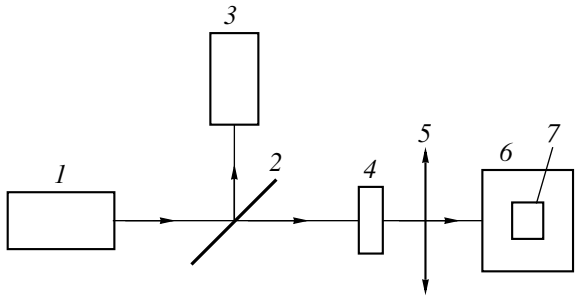
Zr films with a thickness of ≈80 nm were deposited onto Si(111) *n*-type substrates by electron-beam evaporation at a pressure of 3×10^{-5} Pa. During the process, the substrate temperature was kept at 373 K. Prior to deposition, the silicon wafer was chemically cleaned using the method described in [4]. The films grown had fine-grained polycrystalline structure with average grain sizes of 15–20 nm.

The samples were placed into a vacuum chamber where they were heat-treated at a pressure of 5×10^{-5} Pa at a temperature of 500°C for 1, 5, 10, 15, and 30 min. As a UV source, we used a DRSh-250 mercury–quartz lamp. The energy density of incident radiation was E = 0.01 J/cm^2. Its value was monitored by an IMO-2 power meter. The desired wavelengths were obtained with the help of special light filters. Our setup is shown in the figure. The annealing temperature was chosen from the results of tentative experiments. At annealing temperatures T > 500°C, the Zr/Si system was oxidized, whereas at T < 500°C, the phase modifications in it were virtually absent. Therefore, temperatures above 500°C are of interest in investigating physical processes occurring in the system under the joint action of the treatments.

The samples were examined by electron diffraction using a JEM-120 electron microscope and by electron spectroscopy for chemical analysis (ESCA). The Schottky barrier height was determined from the *I–V* curves by the method described in [5]. The Zr/Si contact area was equal to 0.7 mm^2.

To obtain deeper insight into the phase transformations in the Zr/Si system, we performed a similar treatment of thin zirconium films. For the purity of experiment, the Zr films were deposited on the as-cleaved surface of single-crystal NaCl in a single process cycle with the Zr/Si system. Next, they were removed from the NaCl crystal surface and placed on a molybdenum mesh.

It was found that the heat treatment (T = 500°C) of the Zr films causes the oxide phase to appear. It is worth



Setup for sample treatment: (*1*) DRSh-250 mercury–quartz lamp, (*2*) semitransparent mirror, (*3*) IMO-2 power meter, (*4*) light filter, (*5*) focusing lens, (*6*) vacuum chamber, and (*7*) annealing furnace.

1063-7842/01/4609-1112$21.00

noting that when the treatment time was 1–5 min, oxides with a small amount of oxygen, $ZrO_{0.35}$ or ZrO, were formed. However, when the annealing time was 10–30 min, the electron diffraction patterns contained the reflections of zirconia ZrO_2 alone.

Under the joint action of thermal annealing and UV irradiation in a vacuum, the phase composition of the films is modified. The zirconia phase appears even at annealing times of 5 and 1 min for the UV irradiation wavelengths $\lambda = 0.312$ and 0.552 μm, respectively. For $\lambda = 0.365$ μm, zirconia appears only after annealing for 30 min. Note that the kinetics of the phase transformations was virtually independent of the energy density of the incident radiation. The change in the phase composition of the Zr films under the treatments can be represented in the following way:

(1) heat treatment ($T = 500$°C) without UV irradiation:

$$Zr \xrightarrow{1\ \text{min}} ZrO_{0.35};$$

$$ZrO \xrightarrow{5\ \text{min}} ZrO_{0.35}; \quad ZrO_2 \xrightarrow{10\text{–}30\ \text{min}} ZrO_2;$$

(2) heat treatment with irradiation ($\lambda = 0.312$ μm):

$$Zr \xrightarrow{1\ \text{min}} ZrO_{0.35}; \quad ZrO \xrightarrow{5\text{–}30\ \text{min}} ZrO_2;$$

(3) heat treatment with irradiation ($\lambda = 0.365$ μm):

$$Zr \xrightarrow{1\ \text{min}} Zr; \quad ZrO_{0.35} \xrightarrow{5\ \text{min}} ZrO_{0.35};$$

$$ZrO \xrightarrow{10\text{–}15\ \text{min}} ZrO \xrightarrow{30\ \text{min}} ZrO_2;$$

(4) heat treatment with irradiation ($\lambda = 0.552$ μm):

$$Zr \xrightarrow{1\ \text{min}} ZrO; \quad ZrO_2 \xrightarrow{5\text{–}30\ \text{min}} ZrO_2.$$

The growth of the zirconium oxides is stimulated by the reaction of the film with oxygen that adsorbs during the deposition and diffuses from the environment. Since zirconium is a getter, the reaction between the film heated to 500°C and the oxygen proceeds with a high rate. From our results, it appears that during the combined treatment with $\lambda = 0.365$ μm, the chemical activity of oxygen decreases and the oxidation reaction slows down. This is explained by the fact that the zirconium film absorbs photons having an energy of 3.39 eV.

Our data also indicate that when the Zr/Si system is annealed without UV irradiation, the oxide phases with different contents of oxygen ($ZrO_{0.35}$, ZrO, and ZrO_2) and metal-enrinched silicides ($ZrSi_3$, Zr_3Si_2, ZrSi, and $ZrSi_2$) form on the surface. Their appearance depends on the annealing time (Table 1). As was noted, the oxides arise when zirconium combines with the oxygen that adsorbs on the film during deposition and diffuses from the environment. As is seen from Table 1, after annealing for 1 min, the electron diffraction patterns also contain reflections of the initial Zr. The formation of the silicides depends on the diffusion of silicon to the sample surface and its interaction with the Zr film. At an annealing time of 10 min, the silicides are enriched by the metal (Zr_5Si_3 and Zr_3Si_2). An increase in the annealing time to 15 min results in the formation of ZrSi; and to 30 min, of $ZrSi_2$.

Table 1. Modification of the phase composition in the Zr/Si system under thermal annealing at $T = 500$°C

d, Å	τ, min				
	1	5	10	15	30
3.61	–	–	–	–	$ZrSi_2$
3.28	–	–	–	–	$ZrSi_2$
3.21	–	–	Zr_5Si_3	–	–
2.92	–	ZrO_2	ZrO_2	ZrO_2	ZrO_2
2.81	$ZrO_{0.35}$	$ZrO_{0.35}$	–	–	–
2.79	Zr	–	–	–	
2.67	ZrO	ZrO	–	–	–
2.60	$ZrO_{0.35}$	$ZrO_{0.35}$	–	–	–
2.58	Zr	–	–	–	–
2.56	–	–	–	ZrSi	ZrSi
2.55	–	–	Zr_3Si_2	Zr_3Si_2	–
2.53	–	ZrO_2	ZrO_2	ZrO_2	ZrO_2
2.48	$ZrO_{0.35}$	$ZrO_{0.35}$	–	–	–
2.46	Zr	–	–	–	–
2.40	–	–	Zr_3Si_2	Zr_3Si_2	–
2.34	–	–	Zr_5Si_3	–	–
2.30	–	–	–	–	$ZrSi_2$
2.29	–	–	–	ZrSi	ZrSi
2.27	–	–	Zr_5Si_3	–	–
1.86	–	–	–	–	$ZrSi_2$
1.80	–	ZrO_2	ZrO_2	ZrO_2	ZrO_2
1.63	ZrO	ZrO	–	–	–
1.53	–	–	–	ZrSi	ZrSi
1.50	–	ZrO_2	ZrO_2	ZrO_2	ZrO_2
1.43	–	–	Zr_3Si_2	Zr_3Si_2	–
1.39	ZrO	ZrO	–	–	–

Note: *d* is the interplanar spacing; τ is the annealing time.

The combined vacuum treatment insignificantly modifies the phase composition of the surface layer (Table 2). However, note the appearance of the oxide Zr_3O at $\lambda = 0.365$ μm and $\tau = 5$ min. The higher oxide ZrO_2 appears only at a heat treatment time of 30 min. Also, at $\tau = 30$ min, the reflections of ZrSi appear but those of $ZrSi_2$ disappear. The combined treatment with $\lambda = 0.312$ and 0.552 μm (the photon energies are 3.96 and 2.25 eV, respectively) does not change the phase composition in the Zr/Si system as compared with thermal annealing alone (Table 2).

Table 2. Modification of the phase composition in the Zr/Si system under combined treatment at $T = 500°C$ and $E = 0.01$ J/cm^2

λ, μm	τ, min				
	1	5	10	15	30
Thermal annealing without irradiation	Zr, $ZrO_{0.35}$, ZrO	ZrO, $ZrO_{0.35}$, ZrO_2	ZrO_2, Zr_5Si_3, Zr_3Si_2	ZrO_2, ZrSi, Zr_3Si_2	ZrO_2, ZrSi, Zr_3Si_2
0.312	ZrO, $ZrO_{0.35}$	ZrO, ZrO_2	ZrO_2, ZrSi, Zr_3Si_2	ZrO_2, ZrSi	ZrO_2, $ZrSi_2$
0.365	ZrO, $ZrO_{0.35}$	Zr_3O, $ZrO_{0.35}$, ZrO	ZrO, Zr_5Si_3	ZrO, Zr_5Si_3, Zr_3Si_2	ZrO_2, ZrSi, Zr_3Si_2
0.552	ZrO, $ZrO_{0.35}$	ZrO_2, Zr_5Si_3	ZrO_2, Zr_3Si_2	ZrO_2, ZrSi, Zr_3Si_2	ZrO_2, $ZrSi_2$

Table 3. Modification of the phase composition in the Zr/Si system under combined treatment at $T = 500°C$ and $E = 0.05$ J/cm^2

λ, μm	τ, min				
	1	5	10	15	30
0.312	ZrO, $ZrO_{0.35}$	ZrO_2	ZrO_2, ZrSi, Zr_3Si_2	ZrO_2, $ZrSi_2$	ZrO_2, $ZrSi_2$
0.365	Zr, $ZrO_{0.35}$	ZrO, $ZrO_{0.35}$	ZrO, Zr_5Si_3	ZrO, Zr_3Si_2	ZrO_2, ZrSi
0.552	ZrO, ZrO_2	ZrO_2, Zr_3Si_2	ZrO_2, Zr_3Si_2, ZrSi	ZrO_2, ZrSi, Zr_3Si_2	ZrO_2, $ZrSi_2$

Table 4. Modification of the electrophysical parameters of the Zr/Si contact under combined treatment

τ, min	φ_1, eV	φ_2, eV	φ_3, eV	U_1, eV	U_2, eV	U_3, eV	n_1	n_2	n_3
As-prepared	0.57			6			1.16		
Thermal annealing									
1	0.55			6			1.18		
5	0.55			7			1.17		
10	0.57			8			1.14		
15	0.58			9			1.13		
30	0.59			11			1.10		
$E = 0.01$ J/cm^2									
1	0.56	0.55	0.56	6	6	8	1.17	1.18	1.18
5	0.56	0.55	0.57	7	7	8	1.15	1.17	1.16
10	0.58	0.56	0.58	10	7	9	1.12	1.16	1.13
15	0.59	0.57	0.59	11	8	10	1.10	1.14	1.12
30	0.60	0.57	0.60	12	8	11	1.09	1.14	1.10
$E = 0.05$ J/cm^2									
1	0.57	0.55	0.57	10	6	10	1.12	1.18	1.12
5	0.57	0.55	0.58	10	7	12	1.12	1.17	1.09
10	0.58	0.56	0.59	12	7	13	1.09	1.16	1.08
15	0.60	0.56	0.60	14	8	14	1.07	1.16	1.07
30	0.60	0.56	0.61	14	7	15	1.07	1.15	1.06

Note: τ is the treatment time; φ, Schottky barrier height; U, breakdown voltage; and n, ideality factor. Subscripts: 1, λ = 0.312; 2, 0.365; and 3, 0.552 μm.

By increasing the energy density of the incident radiation to 0.05 J/cm^2, the surface layer oxidizes more rapidly (Table 3). At λ = 0.312 and 0.552 μm, the higher zirconium oxide ZrO_2 forms when τ = 5 and 1 min, respectively. The UV irradiation of the Zr/Si system at λ = 0.365 μm (the photon energy $h\nu = 3.39$ eV) slows down the oxidation process and the formation of the silicides. As follows from Tables 2 and 3, the formation of zirconia takes place when the annealing time τ = 5 min or more (λ = 0.312 and 0.552 μm). Photons with $h\nu = 3.96$ and 2.25 eV affect the chemical activity of oxygen, increasing the oxidation rate. Because of the

gettering properties of zirconium, its deposition is accompanied by the dissolution of the residual gases including oxygen. The process of deposition goes in parallel with the ionization of the soluble elements by the metallic solvent. In the lattice of the metals, the cations of light elements are formed and their valence electrons pass into the collective state. The residual oxygen in the surface layer seems to be responsible for the lower oxides $ZrO_{0.35}$ and ZrO. The ordered solid solution of oxygen, Zr_3O, is associated with the redistribution of the oxygen cations over tetrapores.

During the irradiation of the heated Zr/Si system with a photon beam of a particular flow density, the radiation is absorbed by the oxygen present in the surface layer. This absorption is due to the transition $^1\Sigma_g^+ \longrightarrow {}^3\Sigma_g^-$, which is described by the diagram of energy levels for certain states of O_2 [6]. It should be noted that the flow density of photons with the given energy is insufficient to neutralize the chemical activity of the oxygen. This density stimulates the formation of the oxides in the surface layer of the Zr/Si system under combined treatment. To suppress the oxidation, a higher energy density of the incident radiation is required, because thermal annealing causes the oxygen to interact with the zirconium surface. Therefore, to neutralize the M–O bonds, a higher flow density of the photons is necessary. Under UV irradiation, the photochemical processes in the surface layer are associated with the electron transitions in the dissolved and adsorbed oxygen. The electron diffraction data correlate with those obtained by the ESCA method.

The change in the phase composition of the Zr/Si system results in the modification of the electrophysical parameters of the contact. The *I–V* characteristic of the initial contact is asymmetric, and the breakdown voltage is 6 V. This indicates that there is natural silicon dioxide SiO_2 between the Zr film deposited and the silicon, as demonstrated by the ESCA data. The electrophysical parameters after the thermal and combined treatments are listed in Table 4. The annealing decreases the concentration of different defects in the oxide film at the metal–semiconductor interface. The structure of the SiO_2 tends to a more equilibrium state, and the density of electron traps is reduced. However, oxygen diffusion makes the dielectric layer thicker. This, in turn, raises the breakdown voltage. As the annealing time increases, so does the Schottky barrier height. During the thermal annealing of the Zr/Si system, the density of surface states on the silicon at the metal–semiconductor contact changes, which affects the barrier height. The ideality factor *n* of the *I–V* characteristic is given in Table 4. The value of *n* was determined from the experimental *I–V* curves [5]. It is shown that, when the annealing time increases, the ideality factor decreases. This is associated with the change in the dielectric parameters of the oxide layer at the interface and with that in the concentration of the electron traps under the thermal treatment. Thus, as the annealing time increases, the *I–V* curves of the Zr/Si contact tend to their ideal shape. The modification of the electrophysical parameters can be related to the modification of the phase composition of the interface: from $Zr/SiO_2/Si$ for the as-prepared sample to Zr_xO_y/Zr_nSi_m or $ZrSi/SiO_2/Si$ for the sample treated.

Thus, the combined treatment of the Zr/Si system makes it possible to form rectifying contacts with the desired phase composition and, hence, with the desired electrophysical properties.

REFERENCES

1. D. T. Alimov, V. K. Tyugaĭ, P. K. Khabibulaev, *et al.*, Zh. Fiz. Khim. **61**, 3065 (1987).
2. A. M. Chaplanov and A. N. Shibko, Kvantovaya Élektron. (Moscow) **20**, 191 (1993).
3. A. M. Chaplanov and A. N. Shibko, Izv. Akad. Nauk SSSR, Ser. Fiz. **53**, 1111 (1989).
4. *Handbook of Thin-Film Technology*, Ed. by L. I. Maissel and R. Glang (McGraw-Hill, New York, 1970; Sov. Radio, Moscow, 1977), Vol. 1.
5. E. H. Rhoderick, *Metal–Semiconductor Contacts* (Claredon, Oxford, 1978; Radio i Svyaz', Moscow, 1982).
6. J. G. Calvert and J. N. Pitts, Jr., *Photochemistry* (Wiley, New York, 1966; Mir, Moscow, 1968).

Translated by Yu. Vishnyakov

Technical Physics, Vol. 46, No. 9, 2001, pp. 1116–1120. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 49–53.

GAS DISCHARGES, PLASMA

Suppressed Diffusion of Metal Atoms upon Excitation of Slightly Damped Plasmons

B. M. Gorelov, V. S. Mel'nikov, V. M. Ogenko, and G. M. Shalyapina

Institute of Surface Chemistry, National Academy of Sciences of Ukraine, Kiev, 252022 Ukraine

e-mail: user@surfchem.freenet.kiev.ua

Received December 5, 2000

Abstract—In the $YBa_2Cu_3O_{7-\delta}$ superconductor, the excitation of slightly damped plasmons upon the nonuniform heating of electron gas in a microwave electromagnetic field and upon perturbation of the charge density by the thermal desorption of oxygen atoms from the Cu1–O plane suppresses surface and volume diffusion of nickel and gold interstitials. The interstitial atoms aggregate in the surface layer of the crystallites. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In high-temperature superconductors (HTSCs), whose structures are characterized by a combination of wide and narrow bands [1, 2], low-frequency collective excitations, acoustic plasmons, may exist in the heavy-carrier (*h*) subsystem. Under normal conditions, due to light- (*l*) and *h*-carrier Landau damping, the number of plasmons is apparently small, their propagation in the crystal is hardly probable, and their experimental observation is difficult. However, slightly damped plasmons propagating in an HTSC with the velocities

$$\mathbf{v}_{Fh} \ll \mathbf{u} \ll \mathbf{v}_{Fl} \tag{1}$$

or

$$\mathbf{u} \gg \mathbf{v}_{Fh}, \mathbf{v}_{Fl} \tag{2}$$

may manifest themselves in the subthreshold formation of defects, the aggregation of a large number of defects in the surface crystallite layer, and suppressed diffusion of metal atoms. Here, $\mathbf{v}_{Fh}$ and $\mathbf{v}_{Fl}$ are the Fermi velocities of *h* and *l* carriers, $\mathbf{u} = \Omega_\mathbf{q}/\mathbf{q}$ are the excitation rates, $\Omega^2 = 4\pi e^2 n_h/m_h^*$, n_h and m^* are the concentration and the effective mass of *h* carriers, and $\mathbf{q}$ is the wave vector. Of special interest is the suppression of volume diffusion of nickel atoms [3]. If this is the case, the suppression of metal diffusion may indicate the presence of variously excited slightly damped plasmons.

In order to initiate slightly damped excitations, external fields can be used. These fields must not cause defect formation, as well as the transport and aggregation of atoms in the surface layer. Also, they must not suppress diffusion. Under normal conditions, $\mathbf{v}_{Fh} \simeq \mathbf{v}_{Fl}$. Therefore, a microwave electromagnetic field can be employed to stimulate plasmons with velocities satisfying (1). In an electric field $\mathbf{E} = \mathbf{E}_0 \sin\omega_0 t$ at $\omega_0 \ll \upsilon_{h,l}$ ($\upsilon_{h,l}$ are the collision frequencies of *h* and *l* carriers), the *l* and *h* carriers located in nonoverlapping quasi-two-dimensional bands acquire different velocities

$$\tilde{\mathbf{v}}_{h,l}(t) = -\frac{e\mathbf{E}_0}{m^*_{h,l}\omega_0\upsilon_{h,l}}\cos\omega_0 t,$$

because $m_h^* \gg m_l^*$ (m_l^* is the effective mass of the *l* carriers) [4]. The *l*-carrier Landau damping decreases with increasing field amplitude. At a certain $\mathbf{E}_0$, the collective excitation modes of the *h* carriers with phase velocities

$$\tilde{\mathbf{v}}_h(t) \ll \mathbf{u} \ll \tilde{\mathbf{v}}_l(t) \tag{3}$$

can propagate in the crystal nondissipatively.

In order to excite plasmons with velocities satisfying (2), oxygen atoms in $YBa_2Cu_3O_7$ are removed from the Cu1–O chains by heating. When the 01 atoms are thermally desorbed, the perturbations $\delta Z_i e$ of the electron concentration appear near the resulting vacancy in the interlayers (δZ_i is the effective perturbation charge near the *i*th 01 vacancy). The perturbations produce the shielding effect, which increases the spectral density of the plasmon states [5, 6]:

$$S(\omega, t) = -\frac{1}{\pi}\sum_{i,q} V_c(\mathbf{q}) Im\varepsilon^{-1}(\mathbf{q}, \omega).$$

Here, $V_c(\mathbf{q}) = 4\pi\delta Z_i e^2/\mathbf{q}^2$ is the permittivity, $N(\omega) = \iint S(\omega, T) f(\omega, T) d\omega d\mathbf{r}$ is the number of slightly damped plasmons, and $f(\omega, T)$ is the Bose–Einstein distribution function with $\omega \gg \mathbf{v}_{Fh}\mathbf{q}, \mathbf{V}_{F,l}\mathbf{q}$.

It is worth noting that the numbers of slightly damped plasmons $N(\omega)$ probably differ for the two ways of excitation. In the microwave field, the plasmons form only in the skin layer, where Landau damping is weak and the plasmons are excited by local field

effects. In the case of thermal desorption of 01 atoms, the plasmons are excited in the whole volume due to enhanced electron–plasmon interaction. However, irrespective of the way of excitation, the propagation of the slightly damped plasmons in an $YBa_2Cu_3O_7$ crystal may make itself evident in the formation of defects, the transport and aggregation of atoms in the surface layer, and suppressed diffusion. Note that both microwave irradiation and the oxygen thermal desorption do not directly generate defects (except for 01 vacancies), since the energy of a microwave field photon and the kinetic energy of the 01 atoms are small compared with the energy required for Y, Ba, Cu, or O(2, 3) atoms to leave the lattice sites (5.5–20 eV [7–9]) and with the same energy for the 01 and 04 atoms (4.5–10 eV [9]). The energy of the microwave photon is also smaller than their migration energy, which excludes the aggregation of the atoms near the surface and the suppression effect. However, in $YBa_2Cu_3O_{7-\delta}$, oxygen absorption and desorption may induce the diffusion of Ba and Cu atoms to and from the surface, respectively [10]. This may enhance or weaken the suppression effect induced by collective excitations. In addition, the desorption of 01 atoms destroys the band formed by the *pd* orbitals of the 01, 04, Cu1, and Cu2 atoms where the *h* carrier excitations may propagate. The desorption may even eliminate high-temperature superconductivity because of a decrease in the number of holes in the cuprate layers. Therefore, the excitation of slightly damped plasmons is possible only in the range of oxygen concentration $0 < \delta < 0.3$, where the effect of high-temperature superconductivity is likely to take place [11].

The aim of this work is to observe the effect of suppressed diffusion of metal atoms in $YBa_2Cu_3O_{7-\delta}$ when slightly damped acoustic plasmons are excited by microwave irradiation or the thermal desorption of 01 atoms. Ni and Au atoms were used as diffusants. The diffusion of these atoms has both surface and the volume components [12, 13]. This made it possible to observe the suppression of the volume diffusion and the arrest of surface diffusion.

EXPERIMENTAL

We studied polycrystalline single-phase (according to X-ray diffraction analysis) $YBa_2Cu_3O_7$ samples with a density of 5.5 g/cm^3 and a grain size of 5–15 μm. They were produced by solid-state synthesis from a mixture of Y_2O_3, Ba_2CO_3, and CuO powders. The lattice parameters were **a** = 3.821 Å, **b** = 3.889 Å, and **c** = 11.667 Å. The parameter $\eta = (C01 - C05)/(C01 + C05)$ was equal to 0.09 (C01 and C05 are the oxygen concentrations in the 01 and 05 states, respectively). This parameter characterizes the filling of these states with oxygen [14].

The samples were exposed to a pulsed microwave (9.4 GHz) field for 10 min at room temperature. The pulse amplitude, duration, and repetition rate were, respectively, 10^4 W, 2.5 μs, and 400 Hz. Then, the samples were placed in the antinode of the microwave electric field of a single-pass cavity.

Desorption of the oxygen atoms from $YBa_2Cu_3O_{7-\delta}$ was accomplished by heating the samples with $\delta = 0$ at a rate of 5°C/min. We removed the 01 atoms in amounts of $\approx 1.7 \times 10^{21}$ and 2.6×10^{21} cm^{-3}, which corresponds to $\delta \simeq 0.3$ and 0.45, respectively. The amount of the oxygen removed was determined by X-ray diffraction analysis and from data on oxygen absorption due to heating with a Q-1500 derivatograph.

Water molecules were adsorbed on $YBa_2Cu_3O_{6.55}$ samples (preheated in a vacuum at ≈140°C for 2.5 h) from saturated steam at room temperature for 360 min. The annealing of water molecules was performed by heating the samples to 400°C.

The diffusion coefficient of nickel and gold atoms was measured by the radioactive tracer (^{63}Ni and ^{195}Au) method with successive layer removal. The diffusion to a depth of 150–250 μm proceeded in air at temperatures of 200–500°C for 5–45 h. The concentrations were determined with a step of 3–5 μm.

RESULTS AND DISCUSSION

The temperature dependences of the diffusion coefficient for the gold atoms are shown in Fig. 1. In the as-prepared samples, the slow and rapid components of surface diffusion between 200 and 410°C are

$$D_s^s = 2.8 \times 10^{-11} \exp(-0.07/kT)$$

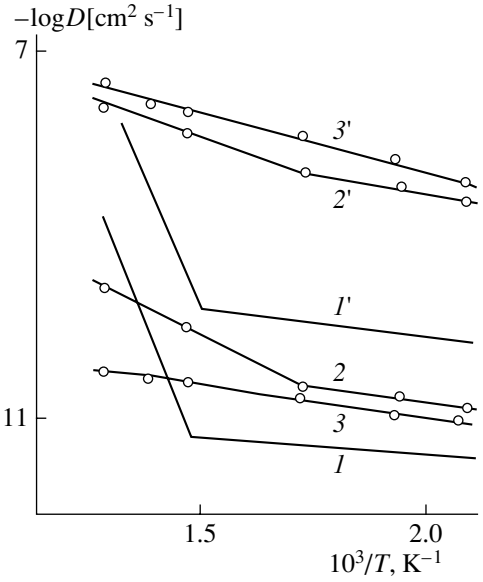


Fig. 1. Temperature dependences of the diffusion coefficient for Au atoms (*1*, *1*') before and (*2*, *2*') after 10-min microwave irradiation and (*3*, *3*') of thermal desorption of oxygen atoms.

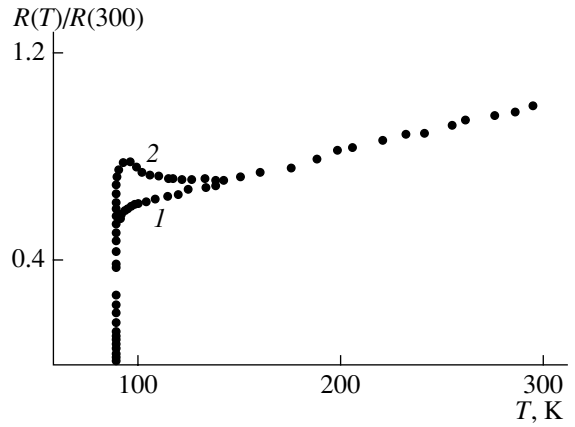


Fig. 2. Temperature dependences of the $YBa_2Cu_3O_7$-resistance (*1*) before and (*2*) after microwave irradiation for 10 min.

and

$$D_s^r = 1.9 \times 10^{-9} \exp(-0.13/kT). \quad (4)$$

For volume diffusion [13], which is observed at temperatures of the thermal desorption of the 01 atoms, $T \geq 410$°C,

$$D_v^s = 6.6 \exp(-1.24/kT)$$

and

$$D_v^r = 1.9 \times 10^{-2} \exp(-1.08/kT). \quad (5)$$

The microwave irradiation arrests the volume diffusion of Au and suppresses both $D_v^{s,r}$ components. In addition, the preexponentials and the energy of activation of the surface diffusion coefficient grow. The components of the surface diffusion coefficient are given by

$$D_s^{s1} = 8.0 \times 10^{-10} \exp(-0.17/kT)$$

and

$$D_s^{r1} = 1.8 \times 10^{-7} \exp(-0.18/kT) \quad (6)$$

in the temperature range 200–305°C, and by

$$D_s^{s2} = 2.3 \times 10^{-7} \exp(-0.45/kT)$$

and

$$D_s^{r2} = 2.0 \times 10^{-6} \exp(-0.30/kT) \quad (7)$$

at $T \geq 305$°C (Fig. 1, curves *2*, *2*'). Note that the irradiation affects neither the temperature nor the rate of thermal desorption of the 01 atoms from the bulk of $YBa_2Cu_3O_7$. This follows from the fact that the curves describing weight losses upon heating the irradiated and the as-prepared samples are identical, which is indicative of the constant oxygen concentration in the samples. The arrest of volume diffusion upon the microwave irradiation implies the formation of volume defects and the aggregation of a large number of atoms in the surface layer of the grains.

Note that after microwave irradiation, the dc resistance R near the normal-superconducting transition increases, whereas the transition temperature T_c does not vary (Fig. 2). The irradiation does not influence the lattice parameters **a** and **b**, whereas the parameter **c** either remains constant (after 1 min of exposure) or decreases by 0.022 Å (after 5 and 30 min of exposure). The parameter η varies from 0.05 to 0.36. The increase in R means that the microwave irradiation generates extra defects, since [6]

$$R(T) = \frac{4\pi}{\Omega_{pl}^2}[\tau_d^{-1} + \tau_{ph}^{-1}(T) + \gamma_q(T)]. \quad (8)$$

Here, Ω_{pl} is the plasma frequency of the l carriers, τ_d and τ_{ph} are the times of scattering by defects and phonons, and γ_q is the coefficient of damping of the l carriers by acoustic plasmons. Note also that $\tau_{ph}^{-1} \sim T$, $\gamma_q \sim T$ for nondegenerate and $\gamma_q \sim T^2$ for generate h carriers. The increase in R at $T_c < T < 150$ K is caused by the rise in both τ_d^{-1} and the number of defects. On the other hand, the migration of oxygen between the 01 and 05 states with an energy of activation of 2.03 eV [15] indicates that microwave irradiation triggers the low-energy mechanism of defect migration. The formation and redistribution of the defects may have no effect on T_c, which depends on the density p of holes [16]:

$$T_c(p) = T_{cm}[1 - 82.6(p - 0.16)^2], \quad (9)$$

where T_{cm} is the maximal value of T_c at a constant number of holes in the cuprate layers, which are responsible for high-temperature superconductivity.

The effect of vacancies of the 01 atoms on the diffusion of Au atoms is similar to that of microwave irradiation. The desorption of the 01 atoms in an amount of 1.7×10^{21} cm^{-3} suppresses the volume diffusion of the Au atoms and increases both the preexponential and the energy of activation of the surface diffusion:

$$D_s^{s3} = 4.0 \times 10^{-10} \exp(-0.16/kT)$$

and

$$D_s^{r3} = 1.45 \times 10^{-6} \exp(-0.24/kT) \quad (10)$$

(Fig. 1, curves *3*, *3*'). The introduction of the vacancies in an amount of 2.6×10^{21} cm^{-3} into the Cu1–O chains has a similar effect on the diffusion of the Ni atoms (Fig. 3). In the as-prepared $YBa_2Cu_3O_7$ samples, the diffusion has a surface component

$$D_s = 3.16 \times 10^{-10} \exp(-0.17/kT)$$

in the temperature range 200–410°C and a volume component

$$D_v = 1.0 \times 10^{-2} \exp(-0.13/kT)$$

at $T > 410$°C [12]. Once the oxygen has been desorbed, the volume diffusion of Ni is suppressed and the surface diffusion is characterized by the slow and rapid components,

$$D_s^{r4} = 2.0 \times 10^{-6} \exp(-0.21/kT)$$

and

$$D_s^{s4} = 1.4 \times 10^{-9} \exp(-0.27/kT) \quad (11)$$

in the whole temperature range (Fig. 3, curves *2*, *2*'). These surface components are apparently due to diffusion in various crystallographic directions [17]. Note that the effect of the slightly damped plasmons excited by the perturbations δZe on the grain surface on Ni diffusion is similar [3].

The arresting effect shows up after the desorption of 01 atoms in amounts of more than 0.9×10^{21} cm^{-3}. This may be achieved by heating to $T > 500$°C. The desorption of a lesser number of 01 atoms or the generation of a lesser number of perturbations $\sum_i \delta Z_i e$ upon heating to $T \simeq 500$°C does not affect the diffusion of the Ni and Au atoms. In addition, in the $YBa_2Cu_3O_{7-\delta}$ samples with $\delta = 0.3$ and 0.45, the diffusion is accompanied by oxygen absorption in the range $200 < T \le 350$°C and oxygen desorption at $T > 400$°C (Fig. 3, insert). However, the formation of oxygen vacancies in the planes Cu1–O in an amount of $\approx 2.6 \times 10^{21}$ cm^{-3} and their filling are not reflected in the temperature dependences of $D_s^{r,s}$ and the volume diffusion D_v does not recover in this case. At the same time, atom diffusion due to oxygen desorption and adsorption during the formation of the diffusion profile [10] seems to insignificantly affect the number of atoms in the surface layer. This indicates that the arresting effect after both microwave irradiation and oxygen desorption is caused by atoms that are transferred to the surface with a rate far exceeding the diffusion rate at the temperatures of diffusion profile formation. This follows from the fact that this effect takes place after the 10-min irradiation or heating for no more than 50 min at a rate of 5°C/min. The transfer may be associated with the migration of atoms in the electric field of the propagating weakly damped plasmons [3].

Thus, both the microwave irradiation and the introduction of oxygen vacancies into the Cu1–O chains affect the diffusion of Au and Ni atoms in a similar way. Namely, they arrest volume diffusion, suppress the components D_v and $D_v^{s,r}$, and increase the preexponential and activation energy of surface diffusion. The D_s^s and D_s^r components of the surface diffusion have close activation energies that are substantially smaller than

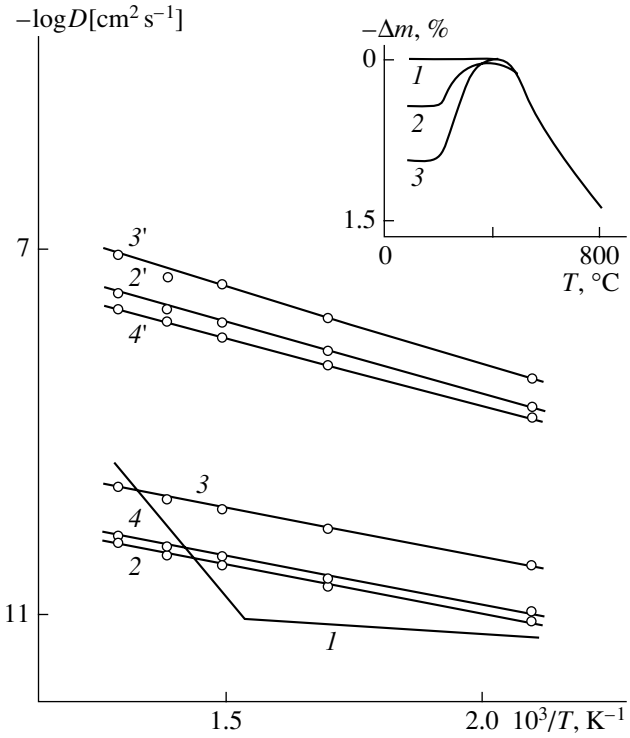


Fig. 3. Temperature dependences of the diffusion coefficient for Ni atoms (*1*) before and (*2*, *2*') after thermal desorption of oxygen atoms, (*3*, *3*') after adsorption of water molecules for 360 min, and (*4*, *4*') after subsequent annealing of water at 400°C. The insert shows the temperature dependences of the $YBa_2Cu_3O_{7-\delta}$ weight change. δ = (*1*) 0, (*2*) 0.3, and (*3*) 0.45.

the activation energy of D_v and are independent of the concentration of the oxygen vacancies in the Cu1–O layers inside the grains.

The suppression of the volume diffusion, which can be described by the expression [18] $D_v = D_0(1 - q)\exp(-E/kT)$ (here, $q = m/M$, m is the number of filled interstices, and M is the total number of interstices), may be explained by an increase in the factor q when the interstices are filled by defects in the surface layers of the grains, as well as by the fact that $D_v \longrightarrow 0$ for $q \longrightarrow 1$. The rise in the preexponentials in the expressions for $D_s^{r,s}$ can also be explained by filling the interstices with the defects. For instance, the introduction of water molecules in an amount of $\approx 1.4 \times 10^{20}$ cm^{-3}, which occupy the interstices of the $YBa_2Cu_3O_{6.55}$ lattice [19], results in an increase in the preexponentials for D_s^s and D_s^r and weakly increases the activation energy:

$$D_s^{r5} = 6.3 \times 10^{-6} \exp(-0.30/kT)$$

and

$$D_s^{s5} = 7.6 \times 10^{-9} \exp(-0.22/kT) \quad (12)$$

(Fig. 2, curves *2*, *2*'). Surface diffusion is nearly completely recovered annealing of the water molecules:

$$D_s^{r6} = 1.2 \times 10^{-6} \exp(-0.26/kT)$$

and

$$D_s^{s6} = 1.6 \times 10^{-9} \exp(-0.20/kT); \quad (13)$$

however, the volume component D_v does not recover (Fig. 2, curves *3*, *3*'). Consequently, the localization of defects in the interstices increases both the preexponential and the activation energy of $D_s^{s, r}$.

In the case of diffusion by two types of interstices [18],

$$D_s^{s, r} = \alpha l^2 \omega \left(\frac{\lambda + q\mu - K}{q\mu^2} \right) \exp(-E/kT), \quad (14)$$

where α is the geometric factor, l is the transfer length, ω is the oscillation frequency of the atom in an interstice, $\mu = 1 - \varepsilon$, $K = \sqrt{(\lambda + 3q\mu)^2 - 12q\mu}$, $\lambda = 1 + 2\varepsilon$, $\varepsilon = \exp[(u_0 - u_T)/kT]$, and u_0 and u_T are the potential energies of the atom in different interstices. The increase in the preexponential factor for $q \longrightarrow 1$ results from the changes in the potential barriers u_0 and u_T.

If the atoms diffuse from a site to an interstice [20],

$$D_s^{s, r} = \frac{1}{6} l^2 (g\tau_g\tau_z)^{-1/2} \exp(-E/kT), \quad (15)$$

where g is the number of vacant sites and τ_g and τ_z are potential energy and interstice, respectively. The behavior of D_s^s and D_s^r is defined by a decrease in τ_g and τ_z when the defects fill the interstices, as well as by an increase in g due to the formation of the defects upon irradiation and thermal desorption.

Thus, nonuniform heating of the *l* and *h* carriers in the microwave field and the perturbations of the charge density in the $YBa_2Cu_3O_7$ interlayers suppress the volume and surface diffusion of metal atoms. The suppression is apparently caused by the aggregation of a large number of interstitials in the surface layers of the grains and may indicate that both effects excite slightly damped collective perturbations of *h* carriers.

REFERENCES

1. G. P. Shveĭkin, V. A. Gubanov, A. A. Fotiev, *et al.*, *Electron Structure and Physicochemical Properties of High-Temperature Superconductors* (Nauka, Moscow, 1990).
2. *High-Temperature Superconductivity. Fundamental Research and Applications*, Ed. by A. A. Kiselev (Mashinostroenie, Leningrad, 1990).
3. B. M. Gorelov, Zh. Éksp. Teor. Fiz. **116**, 586 (1999) [JETP **89**, 311 (1999)].
4. A. F. Alexandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics* (Vysshaya Shkola, Moscow, 1978; Springer-Verlag, Berlin, 1984).
5. É. A. Pashitskiĭ, Fiz. Nizk. Temp. **21**, 995 (1995) [Low Temp. Phys. **21**, 763 (1995)].
6. É. A. Pashitskiĭ, Fiz. Nizk. Temp. **21**, 1092 (1995) [Low Temp. Phys. **21**, 837 (1995)].
7. R. C. Baetzold, Phys. Rev. B **38**, 11304 (1988).
8. N. N. Degtyarenko, V. F. Elesin, and V. L. Mel'nikov, Sverkhprovodimost: Fiz., Khim., Tekh. **3**, 2516 (1990).
9. V. V. Kirsanov, N. N. Musin, and E. I. Shamarina, Sverkhprovodimost: Fiz., Khim., Tekh. **7**, 427 (1994).
10. S. I. Sidorenko, K. I. Barabash, S. M. Volos'ko, and B. V. Egorov, Metallofizika **14** (10), 81 (1992).
11. É. A. Pashitskiĭ, Fiz. Nizk. Temp. **21**, 405 (1995) [Low Temp. Phys. **21**, 315 (1995)].
12. P. P. Gorbik, V. V. Dyakin, F. A. Zaitov, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **3**, 1654 (1990).
13. V. N. Alfeev, P. P. Gorbik, V. V. Dyakin, *et al.*, Dokl. Akad. Nauk Ukr. SSR, No. 1, 41 (1991).
14. V. N. Lysenko, V. T. Adonkin, V. V. Dyakin, *et al.*, Supercond.: Phys., Chem., Technol. **5**, 341 (1992).
15. R. C. Baetzold, Phys. Rev. B **42**, 56 (1990).
16. G. V. M. Williams, J. L. Tallon, R. Michalak, and R. Dupree, Phys. Rev. B **54**, R6909 (1996).
17. V. T. Adonkin, A. P. Galushka, P. P. Gorbik, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **5**, 1901 (1992).
18. A. A. Smirnov, *Theory of Diffusion in Interstitial Alloys* (Naukova Dumka, Kiev, 1982).
19. B. M. Gorelov, D. V. Morozovskaya, V. M. Pashkov, and V. A. Sidorchuk, Zh. Tekh. Fiz. **70** (9), 50 (2000) [Tech. Phys. **45**, 1147 (2000)].
20. W. Seith, *Diffusion in Metallen: Platzwechselreaktionen* (Springer-Verlag, Berlin, 1939; Inostrannaya Literatura, Moscow, 1958).

Translated by M. Fofanov

Technical Physics, Vol. 46, No. 9, 2001, pp. 1121–1124. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 54–57.

SOLID-STATE ELECTRONICS

Semiconductor Compound Thin Films Obtained with Capillary Evaporators

B. N. Gritsyuk, A. A. Lyakhov, S. V. Mel'nichuk, and V. N. Strebezhev

Chernovtsy State University, Chernovtsy, 58012 Ukraine

e-mail: melnych@cv.ukrtel.net

Received September 22, 2000

Abstract—The thickness of decomposable semiconductor thin films that are obtained with capillary evaporators is measured. In the viscous flow approximation, an expression that relates the film thickness, the distance between the capillary and the deposition area, and the angle between the capillary axis and the flow direction is derived. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Obtaining semiconductor compound thin films by thermal evaporation is a challenge, since the components vaporize with various rates that depend on their vapor pressures and composition in the melt. The resulting composition and the properties of the condensed material differ from those of the initial components.

The use of specially designed evaporators that produce a steady vapor flow condensing to stoichiometric films is a promising approach in this field.

MATERIAL EVAPORATION USING CAPILLARY EVAPORATORS

Evaporators for obtaining films of multicomponent alloys and compounds have been described in [1–3]. An original design [4] involves a crucible in the form of a set of capillaries closed at one end, with their diameter such that it prevents convective mixing in the melt. Each of the capillaries is loaded by the initial multicomponent semiconductor charge and the crucible is heated by coaxial cylindrical furnaces. Since melt mixing in the fine capillaries is prevented, the melt becomes depleted in the high-volatility component over time and the steady state that provides its congruent evaporation is established.

However, the evaporator suggested in [4] is not free of disadvantages. To load the crucible, the initial single crystal should be finely ground. As a result, the total contaminated and oxidized surface area of the charge greatly increases; accordingly, this adversely affects the purity and the properties of the films. In addition, the melting time varies from capillary to capillary because of different charge weights and the nonuniform distribution of the thermal field over the capillaries. Because of this, the steady state in the capillaries is established at different time instants, which makes obtaining near-stoichiometric films difficult. The need load each of the capillaries complicates the preparatory stage. The charge weight is small; therefore, it rapidly evaporates, so that the service time of the evaporator is very short.

Capillary evaporators that do not suffer from these disadvantages and allow the growth of high-grade stoichiometric films have been reported in [5, 6]. The capillaries act as a single evaporating system, since the melt is delivered to the capillaries from the noncapillary container common to all of them. The container, much like communicating vessels, is made by connecting the bottoms of the capillaries and has a large volume. This makes it possible to extend the service time of the system and to reduce the number of the capillaries, since they do not contain a charge. The evaporator design can be both compound- and process-specific.

We experimented with capillary evaporators with crucibles made of quartz glass (Fig. 1a) and graphite (Fig. 1b). Capillaries *1* (inner diameter 0.8–1.0 mm) communicated with the charge container *2* (diameter 10–12 mm). The quartz container is sealed on the top and has one capillary. In the graphite container, which is plugged by spacer *3* and threaded stopper *4* from the bottom, the number of capillaries varies depending on the vapor flow rate and trajectory. The crucibles were heated by cylindrical tungsten heaters *5* surrounded by tantalum foil thermal shields *6*.

Once melt *7* has entered the capillary, the evaporating layer becomes depleted by the high-volatility component after a time and the steady state is established; that is, the difference in the vapor pressures of the components is balanced by their amounts in the melt. The outgoing vapor has an appropriate ratio of the high- and low-volatility components and condenses to the stoichiometric films. The evaporator with the quartz crucible is conveniently used to watch the melt level for selecting evaporation conditions and also when the contact of the melt with materials other than quartz is undesirable. However, the quartz crucible can be used only once and is difficult to make if the number of cap-

1063-7842/01/4609-1121$21.00

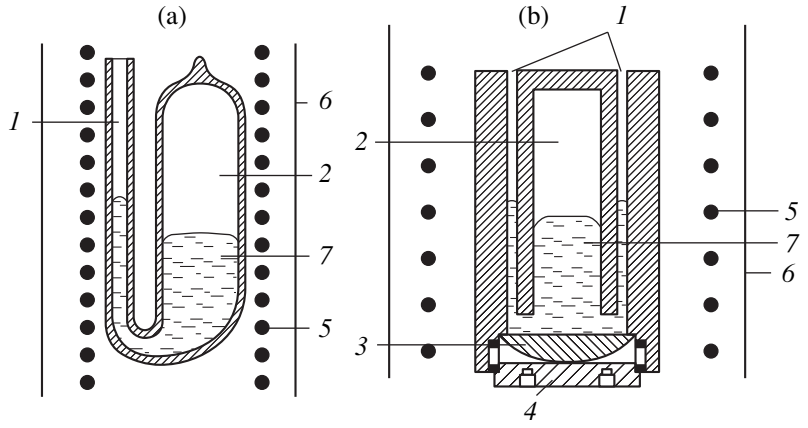


Fig. 1. (a) Single-capillary quartz and (b) four-capillary graphite evaporators.

illaries is more than one. Hence, the deposition rate of the films remains low.

The graphite crucible can be loaded many times. It is significant that the capillaries are filled by the melt simultaneously and evaporation begins from the steady state. The design of the evaporator is known to markedly influence the condensate thickness over the substrate. By the vapor direction and the vapor flow rate, the evaporators are classified into surface-type evaporators, open-type crucible evaporators, closed-type pulsed evaporators, and quasi-closed-type crucible evaporators [7]. Capillary evaporators fall into the last category. The distribution of the condensate thickness over the substrate for the evaporation from a point source, a small-area evaporator, ring- and disk-shaped evaporators, as well as from actual evaporators (effusion cells and cone crucibles), has been reported in [8]. At the same time, data on the condensate thickness for capillary evaporation are lacking in the literature. In this article, we try to bridge this gap, since thickness uniformity is very important for optical and other applications.

CALCULATION OF THICKNESS DISTRIBUTION

The film thickness as a function of the source type and the position of the source relative to the deposition area has been considered in detail in [8] for point, one-dimensional (wire), and planar sources. For the planar source, it is assumed that evaporation obeys Knudsen's cosine law. The mass dM deposited per unit time on the surface area dS whose position is defined by the angle Θ is given by

$$dM = \frac{m\cos^2\Theta}{\pi r^2}dS, \qquad (1)$$

where m is the mass of the material evaporated per unit time and r is the distance from the source to the deposition site.

In our case, Knudsen's law fails, since the material evaporates from the capillary interior rather than from the open surface. Before entering the vacuum chamber, the vapor should pass through the capillary.

Consider the capillary flow of the vapor of a binary compound (e.g., CdSb). Let the capillary radius and length be a and L, respectively. For a vapor pressure of about 1 torr, the ratio of the free path of the molecule to the capillary diameter is 0.017 in our case; therefore, we must consider viscous flow [10]. In addition, the flow velocity is much less than the sound velocity, so that the Mach number, the ratio of the flow and sound velocities, is very small, indicating that the flow is incompressible. At such pressures, the Reynolds number is also very small (Re < 10); hence, the flow is laminar. The distance L_e over which the flow becomes completely equilibrium is given by $L_e = 0.227a\mathrm{Re}$ [9]. In our case, it is much shorter than the capillary part along which the flow moves. However, even for such low pressures, the flow velocity near the walls is other than zero [10] and comprises about 6% of the velocity calculated. We define the flux Q as

$$Q = PSv, \qquad (2)$$

where v is the flow velocity, S is the cross section of the capillary, and P is the mean pressure that specifies this velocity.

Under the above conditions, Poiseuille's law is valid:

$$Q = \pi a^4 P(p_1 - p_2)/(8\eta L), \qquad (3)$$

where p_1 and p_2 are the vapor pressures near the evaporating area and at the end of the capillary, respectively, and η is the dynamic viscosity of the vapor.

The pressure p_2 at the exit from the capillary is far smaller than the pressure p_1 near the evaporating area,

so that we can put $P \approx p_1$. Another expression for the flux is

$$Q = kT\frac{N_A}{\mu}\frac{dm}{dt}, \quad (4)$$

where k is the Boltzmann constant, N_A is the Avogadro number, μ is the atomic mass of the evaporating material, and dm/dt is the evaporation rate. The values of P and v are found from expressions (2)–(4).

Let our coordinate system move with the flow velocity v. Separate an elementary flow layer bounded by two concentric spheres with radii r and $r + dh$, where dh is a small increment. The number of atoms in this layer is

$$dN = 4\pi r^2 n dh, \quad (5)$$

where n is their surface concentration.

In the layer, the number of atoms remains constant (dN = const); therefore,

$$n = \frac{dN}{4\pi dh}\frac{1}{r^2}.$$

The atoms are deposited on the substrate; consequently, for the thickness of the deposit, we have

$$H = \frac{C}{r^2}, \quad r = Vt. \quad (6)$$

Here,

$$C = \frac{dm}{dt}\frac{1}{\pi\rho}$$

is an evaporation-rate-dependent constant, t is the time it takes for an atom to reach the substrate, V is the thermal velocity of the atoms, and ρ is their density inside the capillary. From Fig. 2,

$$vt = L - \sqrt{(Vt)^2 - x^2}. \quad (7)$$

Solving (7) for t, we find

$$t = -L^2 v^2 + \sqrt{L^2 v^2 - (V^2 - v^2)(x^2 + L^2)}. \quad (8)$$

Here, we leave only the positive root. As V, we can take the mean velocity of atoms; that is, $V = (3RT/\mu)^{1/2}$, where R is the gas constant. The thermal velocities of cadmium and arsenic atoms differ only by 4.37%. Therefore, in calculations, they were set equal to each other: $V = (V_{Sb} + V_{Cd})/2$. Taking into consideration that $x^2 + L^2 = L^2/\cos^2\Theta$, we obtain from (6) and (8)

$$H(\Theta, L) = \frac{C(V^2 - v^2)^2\cos^4\Theta}{V^2L^2(\cos 2\Theta v^2 - 2\cos\Theta v\sqrt{V^2 - v^2\sin^2\Theta + V^2})}, \quad (9)$$

where L is the capillary end–substrate distance and Θ is the angle between the capillary axis and the radius vector from the center of the evaporating area (from the center of the capillary end) to a point on the deposition area where the film thickness $H(\Theta, L)$ is measured.

If the flow velocity v inside the capillary is zero, expression (9) passes to Knudsen's law for the evaporating area:

$$H = \frac{C\cos^4\Theta}{L^2} = \frac{C\cos^2\Theta}{r^2}.$$

EXPERIMENTAL DATA

Using the original capillary evaporators, we deposited II–V semiconductor (CdSb and ZnSb) films. When these compounds evaporate from evaporators of usual design (boats, helical and strip evaporators, and crucibles), their dissociation takes place [7]. As a result, early in the process, the vapor is enriched by the high-volatility component (Zn or Cd). Then, multiphase mixtures of metastable cadmium and arsenic compounds form. Near-stoichiometric CdSb films have been obtained by pulsed evaporation [11], double-source evaporation [12], and laser-induced evaporation [13]. These techniques are, however, very difficult to apply in practice. CdSb and ZnSb films with properties close to those of the starting single crystals have been obtained in our previous works by using quartz [5, 14] and graphite [6] capillary evaporators. The properties and the composition of the films successively applied in a single process cycle were shown to vary from metal to semiconducting during the transient in the capillary evaporator. Once the steady-state evaporation conditions have been established in the evaporator, the properties of the films become identical to those of CdSb and ZnSb single crystals. This effect has been most vividly demonstrated with temperature dependences of the conductivity that were taken from the films in the order of their application [14].

In this work, we studied the thickness uniformity of CdSb films deposited on pyroceramic substrates from a four-capillary graphite evaporator. The process was carried out in a VUP-5 installation at a pressure of 1×10^{-6} torr, and the substrates were 12 cm away from the

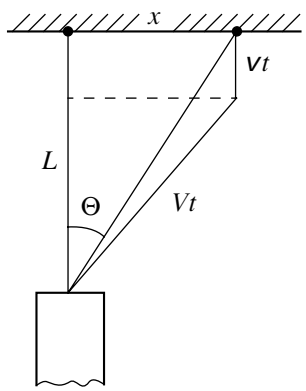


Fig. 2. Particle deposition for capillary evaporation.

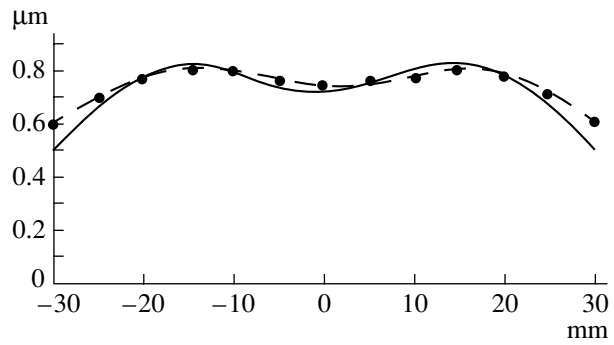


Fig. 3. Film thickness distribution over the substrate surface.

open end of the capillary. The film thickness distribution was determined by an MII-11multiple-beam interferometer and also by examining the substrate–film interface (on the cross section of the structure) with an RÉM-100U scanning electron microscope.

Figure 3 shows the film thickness distribution over the substrate (dashed line). The continuous line represents the calculated distribution of the condensate thickness for the case when four capillaries are arranged at the vertices of a square. It is seen that four-capillary evaporation improves the film thickness uniformity over a substantially large area. The quality of the films obtained, as well as agreement between the experimental and calculated thickness distribution curves, suggests that the use of capillary evaporators is a promising approach and that our technique for determining film thickness is appropriate.

REFERENCES

1. Yu. M. Bukker, A. S. Valeev, A. A. Vanin, *et al.*, USSR Invetror's Certificate No. 659642, Byull. Izobret., No. 16 (1979).
2. V. M. Kuznetsov, *Vacuum Technology* (VINITI, Moscow, 1977), No. 8.
3. B. A. Chenov, S. P. Oles'kiv, and B. S. Oles'kiv, Fiz. Élektron. (Lvov) **38**, 66 (1989).
4. G. F. Dolidze, D. V. Yakashvili, V. N. Vigdorovich, and G. A. Ukhlinov, USSR Invetror's Certificate No. 544711, Byull. Izobret., No. 4 (1977).
5. B. N. Gritsyuk and V. N. Strebezhev, Prib. Tekh. Éksp., No. 5, 157 (1997).
6. B. N. Gritsyuk and V. N. Strebezhev, *Problems of Atomic Science and Technology* (VANT, Kharkov, 1998).
7. *Handbook of Thin-Film Technology*, Ed. by L. I. Maissel and R. Glang (McGraw-Hill, New York, 1970; Sov. Radio, Moscow, 1977), Vol. 1.
8. L. Holland and W. Steckelmacher, Vacuum **11** (4), 346 (1952).
9. H. L. Langhaar, J. Appl. Mech. **9**, A-55 (1942).
10. S. Dushman, *Scientific Foundations of Vacuum Technigue*, Ed. by J. M. Lafferty (Mir, Moscow, 1960; Wiley, New York, 1962).
11. S. I. Baranov, I. V. Glavatskiĭ, K. R. Zbigli, and A. G. Cheban, *Thermoelectric Devices and Films: Proceedings of the All-Union Conference, Leningrad, 1976*.
12. M. Komatsu, N. Matsuda, Y. Kashiwaba, and H. Saito, Mater. Bes. Bull. **13**, 835 (1978).
13. B. N. Gritsyuk and S. V. Nicheĭ, Prib. Tekh. Éksp., No. 2, 114 (1997).
14. B. N. Gritsyuk, V. V. Zolotukhina, I. M. Rarenko, and V. N. Strebezhev, Fiz. Élektron. (Lvov) **39**, 19 (1989).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1125–1127. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 58–60.

InAs/InAsSbP Light-Emitting Structures Grown by Gas-Phase Epitaxy

E. A. Grebenshchikova, N. V. Zotova, S. S. Kizhaev[1], S. S. Molchanov, and Yu. P. Yakovlev

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

[1]e-mail: serguie@mail.ru

Received October 16, 2000

Abstract—Using the metal-organic chemical vapor decomposition technique, light-emitting diodes based on InAs/InAsSbP double heterostructures emitting in a wavelength range around 3.3 μm have been fabricated. The external quantum yield of the diodes is 0.7%. In laser diodes, stimulated emission at a wavelength of 3.04 μm has been obtained at $T = 77$ K. *© 2001 MAIK "Nauka/Interperiodica".*

At present, a demand exists for devices that monitor the maximum admissible concentrations of a number of hydrocarbons (methane, propane, ethylene, and others). Molecules of methane, as well as other hydrocarbons, can absorb infrared radiation. Methane has strong absorption bands in a wavelength range around 3.3 μm [1]; therefore, semiconductor light-emitting diodes (LEDs) emitting in this spectral range can be used as radiation sources for portable gas analyzers. Compounds based on lead salts (IV–VI) [2], as well as narrow band HgCdTe (II–VI) semiconductors [3], have low thermal conductivity and considerable metallurgical instability, which makes them less suitable for the fabrication of infrared emitters than III–V solid solutions. Light-emitting structures for the 3.3 μm spectral range based on III–V semiconductors are usually grown by the liquid-phase epitaxy (LPE) method [4–6] or sometimes by the method of metal-organic chemical vapor deposition (MOCVD) [7] and molecular beam epitaxy (MBE) [8].

The main drawback of the LEDs for the 3–5 μm spectral range is that their power output is insufficient for practical applications, being ≤0.1 mW in the continuous mode and ~ 1 mW in the pulsed mode. This work is a continuation of our research on LEDs for the 3–5 μm spectral range grown by MOCVD [9], with the aim of increasing the LED efficiency by about an order of magnitude compared with LEDs grown by LPE. Our expectations are based on the possibility of more efficient use of the potential of the MOCVD technique in growing InAsSb/InAsSbP heterostructures, first of all by growing InAsSbP layers having a wider bandgap (for example, in the immiscibility region); also, heterostructures with better electron and optical confinement and more perfect morphology can, hopefully, be grown. This work deals with the use of MOCVD for the fabrication of LEDs emitting at 3.3 μm with high emitted optical power.

Consider a symmetric double heterostructure grown for use in fabricating LEDs and lasers (Fig. 1). The structure consisted of an InAs(111)B substrate ($n \sim 3 \times 10^{17}$ cm^{-3}) on which were grown a 1.2 μm thick unintentionally doped InAsSbP layer ($n \sim 10^{17}$ cm^{-3}), an unintentionally doped active layer of *n*-InAs (1 μm thick), and a *p*-InAsSbP layer doped with zinc to $p \sim 8 \times 10^{17}$ cm^{-3} of a thickness 1.2 μm. Subsequent measurements have shown that zinc diffusion from the last barrier layer of InAsSbP took place, making the conductivity of the InAs active layer the *p*-type. The *p–n* junction was formed in the first barrier layer of InAsSbP 0.5 μm from the *n*-InAs substrate. The phosphorus content in the barrier layers was 25%. The bandgap (E_g) of the InAsSbP solid solution calculated with the use of data in [10] was found to be 580 meV.

InAsSbP/InAs/InAsSbP structures were grown by MOCVD in a standard horizontal-type reactor under atmospheric pressure. The reactor was similar in design to a system considered earlier [11]. The rate of hydrogen flow through the reactor was 18 l/min. The sources of indium, arsenic, antimony, and phosphorus were,

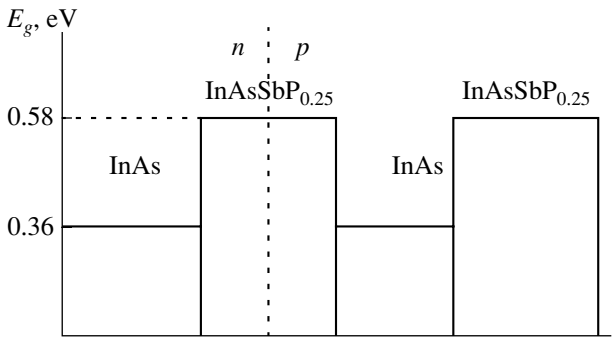


Fig. 1. Energy diagram of the grown symmetrical InAsSbP/InAs/InAsSbP double heterostructure at room temperature.

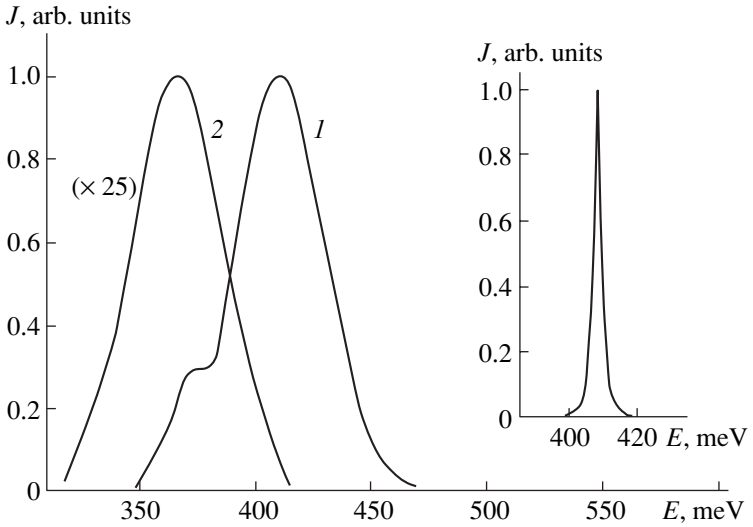


Fig. 2. Electroluminescence spectra of the diodes at different temperatures. T = (*1*) 77, (*2*) 300 K. Inset: stimulated emission spectrum of the laser diode at T = 77 K.

respectively, trimethylindium (TMIn), arsine (AsH_3) diluted to 20% with hydrogen, trimethylstibin (TMSb), and phosphine (PH_3) diluted to 20% with hydrogen.

The InAsSbP barrier layers were grown at a substrate temperature of 580°C. The flow rates of hydrogen through bubblers with TMIn and TMSb were 435 and 50 cm^3/min, respectively. TMIn and TMSb were maintained at temperatures of 27 and 6°C, respectively. The flow rates of AsH_3 (20%) and PH_3 (20%) were 6 and 50 cm^3/min, respectively. InAsSbP was made *p*-type by doping with zinc. The zinc source was diethylzinc (DeZn). The bubbler with DeZn was held at a temperature of 4.7°C. The hydrogen flow through the DeZn vaporizer was 20 cm^3/min.

The InAs active layer was grown at a substrate temperature of 620°C and the V/III ratio in the gas phase equal to 40.

The LEDs were fabricated in the form of mesas using standard photolithography. The mesa diameter was 300 μm. A solid ohmic contact was applied on the side of *p*-InAsSbP. The diameter of the point contact to the substrate was 100 μm. Ohmic contacts were prepared by evaporaton of gold with tellurium (onto the layer of *n*-type conductivity) or gold with zinc (onto the layer of *p*-type conductivity).

The laser diodes had a mesa-stripe width of 30 μm. The contact stripe was made on the *p*-InAsSbP layer. The width of the contact stripe was 5 μm. Resonators 300 μm in length were made by cleaving. Properties of the structures grown were studied using electroluminescence (EL). EL was registered by a cooled InSb photodiode in the phase-lock detection mode.

First, consider the characteristics of the LEDs. Figure 2 shows EL spectra at 77 and 300 K. The diodes were powered with a pulsed current of 1 A (t = 5 ms, f = 500 Hz). In the spectra taken at 77 K, two peaks are clearly distinguished. The short-wavelength peak corresponds to the *n*-InAs substrate ($h\nu_{max}$ = 410 meV, $Dh\nu_{1/2}$ = 50 meV) with an electron concentration of ~3×10^{17} cm^{-3} [12], and the other peak is due to the *p*-InAs active layer ($h\nu_{max}$ = 380 meV, $Dh\nu_{1/2}$ = 26 meV) with $p \sim 8 \times 10^{17}$ cm^{-3} [13]. These results suggest that zinc diffusion took place from the top barrier layer into the semiconductor bulk and the contributions into radiative recombination come from the *n*-InAs substrate and the InAs active layer having *p*-type conductivity due to zinc diffusion. To confirm these observations, the location of the *p*–*n*-junction was determined using scanning electron microscopy. The *p*–*n*-junction was found to be located in the first wide-bandgap InAsSbP layer at a distance of 0.5 μm from the substrate (Fig. 1). At T = 295 K, the peak in the EL spectrum corresponds to $h\nu_{max}$ = 364 meV and $Dh\nu_{1/2}$ = 56 meV. At room temperature, InAs has Eg = 360 meV [10]; the shift of the peak to higher energies confirms that the heavily doped *n*-InAs substrate contributes to the radiative recombination.

W–*I* characteristics of the diodes were measured in pulsed and continuous modes. Shown in Fig. 3 is a *W*–*I* characteristic of the diode in pulsed mode (t = 5 ms, f = 500 Hz). Nonlinearity of the characteristic is not caused by heating of the diode, as evidenced by the coincidence of the characteristics measured in continuous and pulsed current modes. The external quantum yield of the diodes was ~0.7%. Fast saturation of the power versus the pump current curve is evidently due to the small thickness of the active region (~1 μm), causing a rapid rise in the charge carrier concentration and, as a consequence, enhancement of the Auger recombination.

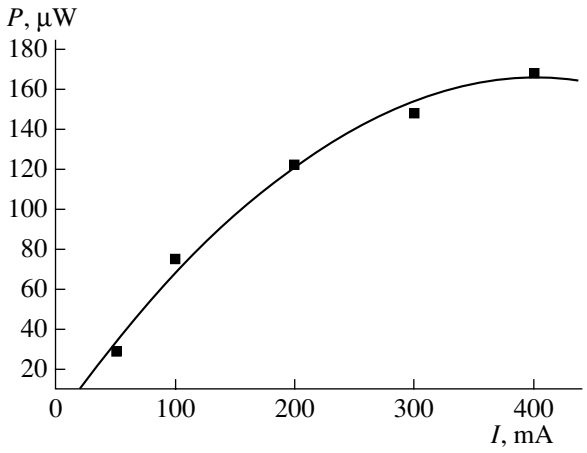


Fig. 3. *W–I* characteristic of the laser diode measured in pulsed regime (τ = 5 μs, f = 500 Hz). P is the optical power.

The efficiency of the diodes can be improved by optimizing the DHS parameters. So, in [14], where the recombination mechanism of nonequilibrium carriers in $InAs/InAs_{0.16}Sb_{0.84}$ structures (with the active region of *p*-InAs) was considered, it was found that at small injection currents when the radiative recombination efficiency is independent of current, the efficiency in *p*-InAs reached 24% at 300 K and the external quantum yield was ~9%. These values were measured for an optimum hole concentration of $p \sim 3 \times 10^{17}$ cm^{-3} and an active region thickness of ~3 μm. With increasing active region thickness, reabsorption of the emission increased (absorption length in *p*-InAs at the emission spectrum maximum was ~4.8 μm) and the efficiency dropped. At less than optimum thicknesses, the recombination at the InAs/AlAsSb interface was more pronounced.

The inset in Fig. 2 shows the stimulated emission spectrum at T = 77 K of a laser diode fabricated using the structure grown. Its peak position at 3.04 μm corresponds to an *Eg* value in InAs of 408 meV for this temperature. The observation of stimulated emission provides evidence of a perfected InAs/InAsSbP heterointerface. The high value of the threshold current J_{th} = 330 mA is explained by insufficient thickness of the InAsSbP barrier layers, nonoptimal for the laser structure size of the active region and the position of the *p*–*n*-junction.

In the future, in order to increase the efficiency of the light-emitting diodes based on the InAs/InAsSbP DHS, the active layer thickness must be increased to the optimal value and the hole concentration reduced. Also, because of the higher probability of radiative recombination in *n*-InAs than in *p*-InAs [13], the active region should be made *n*-type. In addition, with the MOCVD technique, InAsSbP solid solutions in the immiscibility region can be grown [11, 15, 16]; these can be used to enhance the optical and electronic confinement in InAs/InAsSbP DHSs. Although the diode structure investigated is far from optimal, the emission power of the first LEDs grown by MOCVD is found to be comparable to LEDs prepared by other methods [4–8].

ACKNOWLEDGMENTS

The authors are grateful to M. A. Remenny for his interest in this work and measurements of the emission power of the LEDs; T. B. Popov for determining the chemical composition of InAsSbP; and V. A. Soloviev for determining of the position of the *p–n*-junction. S. S. Kizhaev thanks the Robert Haveman Foundation for financial support during this study.

REFERENCES

1. L. S. Rothman, R. R. Gamache, R. H. Tipping, *et al.*, J. Quant. Spectrosc. Radiat. Transf. **48**, 469 (1992).
2. Z. Feit, D. Kostyk, R. J. Woods, and P. Mak, Appl. Phys. Lett. **58**, 343 (1991).
3. E. Hadji, J. Bleuse, N. Magnea, and J. L. Pautrat, Appl. Phys. Lett. **67**, 2591 (1995).
4. M. Aĭdaraliev, N. V. Zotova, S. A. Karandashev, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 102 (2000) [Semiconductors **34**, 104 (2000)].
5. A. A. Popov, M. V. Stepanov, V. V. Sherstnev, and Yu. P. Yakovlev, Pis'ma Zh. Tekh. Fiz. **23** (21), 24 (1997) [Tech. Phys. Lett. **23**, 828 (1997)].
6. M. K. Parry and A. Krier, Electron. Lett. **30**, 1968 (1994).
7. A. Stein, D. Puttjer, A. Behres, and K. Heime, IEE Proc.: Optoelectron. **145** (5), 257 (1998).
8. B. Grieteus, S. Nemeth, and G. Borghs, in *Proceedings of the International Conference on Midinfrared Optoelectronics, Materials and Devices, Lancaster, 1996.*
9. N. V. Zotova, S. S. Kizhaev, S. S. Molchanov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 1462 (2000) [Semiconductors **34**, 1402 (2000)].
10. S. Adachi, J. Appl. Phys. **61**, 4869 (1987).
11. W. J. Duncan, A. S. M. Ali, E. M. Marsh, and P. C. Spurdens, J. Cryst. Growth **143**, 155 (1994).
12. A. A. Allaberenov, N. V. Zotova, D. N. Nasledov, and L. D. Neuĭmina, Fiz. Tekh. Poluprovodn. (Leningrad) **4**, 1939 (1970) [Sov. Phys. Semicond. **4**, 1662 (1970)].
13. N. P. Esina and N. V. Zotova, Fiz. Tekh. Poluprovodn. (Leningrad) **14**, 316 (1980) [Sov. Phys. Semicond. **14**, 185 (1980)].
14. M. J. Kane, G. Braithwaite, M. T. Emeny, *et al.*, Appl. Phys. Lett. **76**, 943 (2000).
15. M. J. Jou, Y. T. Cherng, H. R. Jen, and G. B. Stringfellow, J. Cryst. Growth **93**, 62 (1988).
16. A. Behres, D. Puttjer, and K. Heime, J. Cryst. Growth **195**, 373 (1998).

Translated by B. Kalinin

Technical Physics, Vol. 46, No. 9, 2001, pp. 1128–1132. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 61–65.

SOLID-STATE ELECTRONICS

Transition Processes Occurring under Continuous and Stepwise Heating of GaAs Surface-Barrier Structures

Yu. A. Gol'dberg and E. A. Posse

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

Received October 31, 2000.

Abstract—Evolution of the capacitance–voltage (*C–U*) and current–voltage (I_f–*U* and I_r–*U*) characteristics of solid metal–semiconductor structures (Ni/GaAs) in the process of their continuous and stepwise heating are studied. Properties of the initial structures obey the theory of thermionic emission. It has been shown that as a result of continuous heating, the rectifying structures become ohmic at a temperature of T_{Ohm} = 720 K, which is substantially lower than the melting points of the metal or the metal–semiconductor eutectic. For comparison, properties of the structures annealed at different temperatures T_{ann} are measured after cooling to room temperature (stepwise heating). In this case, *I–U* characteristics are closer to the initial ones for annealing temperatures $T_{ann} < T_0$ = 553 K; for $T_{ann} > T_0$, the characteristics display excess currents; and, finally, for T_{ann} exceeding T_0 by 200–300 K, the characteristics become purely ohmic. It is suggested that these effects are due to a chemical interaction between Ni and GaAs, which changes the properties of the semiconductor surface. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known that a metal–semiconductor contact is either rectifying if there is a potential barrier between the metal and the semiconductor impervious to tunneling, or ohmic if there is no potential barrier or the barrier is transparent for tunneling (see, for example, [1–3]). In our studies [4, 5] of liquid metal–semiconductor structures (Ga/GaP, In/GaP, Ga/GaAs), it was found that as a result of heating the structure, the rectifying contact becomes ohmic because the semiconductor surface layer by the metal dissolves. In [6], we showed that such a transition also occurs in the case of a semiconductor–solid metal contact.

In the present work, we studied the evolution of the capacitance–voltage (*C–U*) and current–voltage (I_f–*U* and I_r–*U*) characteristics of solid metal–semiconductor structures (Ni/GaAs) caused by continuous and stepwise heating.

EXPERIMENTAL TECHNIQUE

For preparation of the structures, epitaxial GaAs layers ($n = 10^{15}$ cm^{-3}) grown on GaAs substrates ($n = 10^{18}$ cm^{-3}) were used. Orientation of all crystals was in the (100) plane. First, an ohmic contact was applied by alloying In into a GaAs substrate. The epitaxial GaAs layer was processed in the way usually used in preparing rectifying surface-barrier structures: mechanical lapping, chemical lapping, and washing. Then a layer of Ni was deposited onto the chemically treated surface [7]. The obtained structures were rectifying.

To study the evolution of properties of the Shottky diodes, some specimens were heated continuously from room temperature to 870 K at a slow rate (5 deg/min) in a neutral atmosphere (helium); current–voltage (*I– U*) and capacitance–voltage (*C–U*) characteristics were measured in the course of heating (i.e., without cooling to room temperature).

Other specimens were subjected to a stepwise heating; i.e., they were annealed at different temperatures with cooling down to room temperature after annealing; then *I–U* and *C–U* characteristics were measured. The parameter characterizing the asymmetry of the *I–U* characteristic was the rectification coefficient $K_r = I_f/I_r$ measured at $U = \pm 0.5$ eV.

RESULTS AND DISCUSSION

1. Let us consider the evolution of the differential capacitance–voltage and current–voltage characteristics in the case of continuous heating of the structures.

1.1. *C–U* dependences in C^{-2}–*U* coordinates were linear over a temperature range of 290–470 K (Fig. 1a), in agreement with Shottky's theory

$$C^{-2} = \frac{2(U_d - U - kT/q)}{\varepsilon_s \varepsilon_0 S^2 q (N_d - N_a)}, \quad (1)$$

where ε_s is the static dielectric constant of the semiconductor; ε_0 is the dielectric constant of a vacuum; q is the electron charge; $N_d - N_a$ is the concentration of uncompensated donors in the semiconductor; U_d is the diffusion potential; k is the Boltzmann constant; and T is the temperature (K).

1063-7842/01/4609-1128$21.00

From these characteristics, the concentration of uncompensated donors, $N_d - N_a = 1–2 \times 10^{15}$ cm^{-3}, and values of the diffusion potential difference at various temperatures were determined. The potential barrier height not corrected for its lowering by image forces (φ_{BO}) was determined from the formula $\varphi_{BO} = U_d + \mu/q$, where μ is the Fermi level energy in the semiconductor. From the plots of $\varphi_{BO} = f(T)$ (Fig. 1b), the temperature coefficient $\alpha = 2.4 \times 10^{-4}$ V/deg, and φ_{BO} extrapolated to 0 K, $\varphi_{BO}(0\ K) = 0.96$ V, were determined.

The C^{-2}–U dependence measured at temperatures $T \geq 470$ K deviated from a straight line (Fig. 1a) and the measured capacitance was significantly lower. In this case the capacitance measured by the bridge method differed substantially from its true value because of the considerable drop in the differential resistance of the structure, giving evidence that the barrier contact became ohmic.

1.2. Measured curves of the forward current I_f vs. voltage (Fig. 2a) had an exponential portion in the temperature range T = 290–430 K,

$$I_f = I_0 \exp(qU/\beta kT), \tag{2}$$

where β = 1.02–1.03.

Analysis of these characteristics has shown that the dependence of I_f on U obeys the theory of thermionic emission with correction for the effect of image forces on the potential barrier height $\varphi_B(U)$:

$$I = I_0[\exp(qU/kT) - 1], \tag{3}$$

$$I_0 = A^*ST^2 \exp(-q\varphi_B/kT), \tag{4}$$

$$\varphi_B = \varphi_B(0K) - \alpha T, \tag{5}$$

where the Richardson constant is $A^* = 120 m_e/m_0$, m_e being the effective mass of majority charge carriers and m_0 the free electron mass in vacuum; and α is the barrier height temperature coefficient.

If it is taken into account that the true barrier height is

$$\varphi_B(U) = \varphi_{BO} - \Delta\varphi_B(U), \tag{6}$$

then the preexponential factor I_0 can be written as

$$I_0 = I_s \exp(q\Delta\varphi_B(U)/kT), \tag{7}$$

where the saturation current is

$$I_s = A^*ST^2 \exp(-q\varphi_{BO}/kT), \tag{8}$$

and the barrier lowering due to image forces is

$$\Delta\varphi_B = \left[\frac{q^3(N_d - N_a)}{8\pi^2\varepsilon_s\varepsilon_0^2}(U_d - U - kT/q)\right]^{1/4}. \tag{9}$$

To determine I_s at different temperatures, I_f–U characteristics were plotted in coordinates $I_f = f[U + \Delta\varphi_B(U)]$. Values of $\Delta\varphi_B(U)$ were determined for each value of U by Eq. 9. Values of I_s were determined by extrapolating to $[U + \Delta\varphi_B(U)] = 0$ the exponential portions of the dependences plotted in the form $I_f = f[U + \Delta\varphi_B(U)]$. The dependence of I_s/ST^2 on $1/T$ (Richardson's plot in Fig. 3) turned out to be linear, in accordance with the theory, with the Richardson constant equal to $A^* = (8.2 \pm 1.0)$ A/cm^2 deg^2, as predicted by the theory (the electron effective mass for GaAs is $m_e^* = 0.068 m_0$). The potential barrier height not corrected for the image forces determined from Richardson's plot,

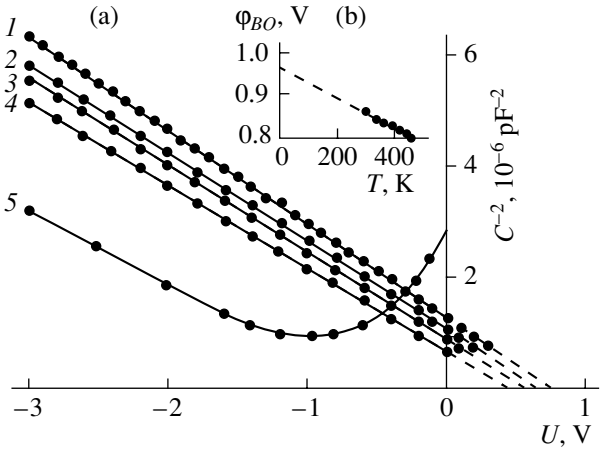


Fig. 1. (a) Dependence of the capacitance C on voltage U for structure N1 at temperatures T, K: *1*—295; *2*—368, *3*—413; *4*—465; *5*—493. (b) The temperature dependence of the potential barrier height φ_{BO} calculated from C^{-2}–U characteristics by the cut-off voltage on the ordinate.

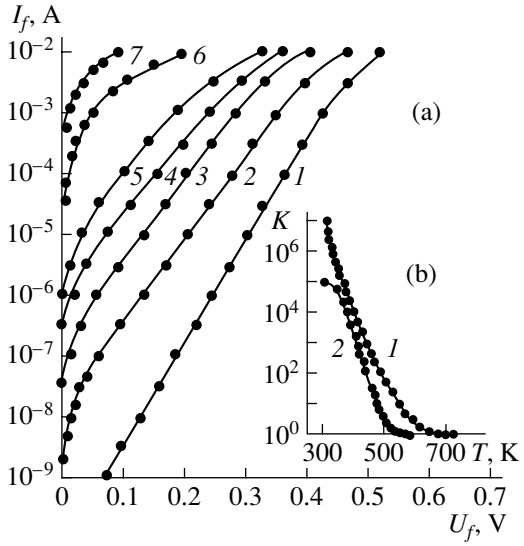


Fig. 2. (a) Dependences of the forward current I_f on voltage U for structure N1 at temperatures T, K: *1*—296; *2*—343, *3*—373; *4*—393; *5*—428; *6*—473; *7*—563. (b) The temperature dependence of the rectification coefficient K_r for (*1*) a Ni/GaAs structure and (*2*) a Ga/GaAs liquid metal-semiconductor structure [4].

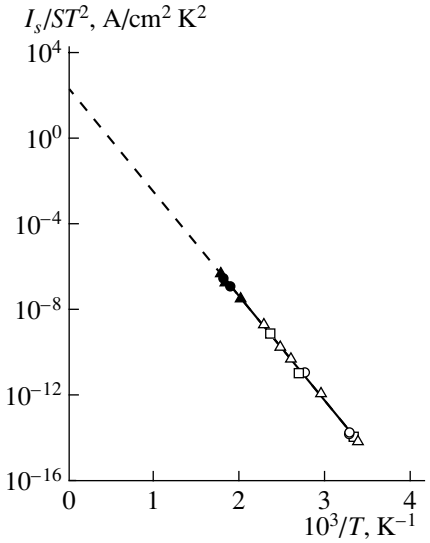


Fig. 3. A Richardson plot drawn using measured I_f–U data (△, ○, and □) and I_r–U data (▲, ●) for the Ni–GaAs structures.

$\varphi_{BO}(0\ K)$ = 0.96 V, coincided with the value obtained from the *C*-*V* characteristics.

At $T > 430$ K, the exponential portion could not be observed because the voltage dropped almost entirely across the residual resistance (resistance of the semiconductor bulk and the ohmic contact) as a result of a substantial reduction in the differential resistance of the structure.

1.3. The dependence of the reverse current (I_r) on voltage (Fig. 4) was measured in the temperature interval 290–580 K. At $T < 350$ K, the reverse current was very low. At $350 \le T \le 450$ K, the reverse current was strongly dependent on voltage, probably because of leakage currents. At $490 \le T \le 550$ K, the I_r–U characteristics had a portion close to saturation (Fig. 4).

In order to determine the reverse current mechanism corresponding to these portions of the characteristic, a Richardson plot was built for the reverse current. I_s values that would have been in effect in the absence of the image forces were determined by dividing the reverse current values at different voltages by $\exp(q\Delta\varphi_B(U)/kT)$ (solid lines in Fig. 4). The plot of I_s/ST^2 as a function of $1/T$ for the reverse current turned out to be a linear continuation of the Richardson plot for the forward current (Fig. 3). This means that the reverse current in the structures studied follows the thermionic emission theory at temperatures up to T = 550 K. At $T > 550$ K, an excess bulk current appeared, caused by the irreversible transition of the barrier contact into an ohmic one.

Thus, in the Ni/GaAs surface-barrier structures studied, the forward current at $T \le 430$ K and the reverse current $T \le 550$ K are due to thermal electron emission. At higher temperatures, excess currents arise that are caused by transition of the contact from barrier to ohmic.

1.4. Let us consider next the behavior of the ratio of the forward and reverse currents (rectification ratio, $K = I_f/I_r$) in the process of continuous heating of the structures. The measurement results for $K = f(T)$ at $U = \pm 0.5$ V (Fig. 2b) can be summarized as follows.

(a) Similar to semiconductor–liquid metal contacts [4], when a certain temperature is reached in the process of continuous heating, the Ni/GaAs rectifying contacts become ohmic (Fig. 2b). This transition takes place before the possible formation of a heavily doped or graded-gap recrystallized semiconductor typical of traditional ohmic contacts.

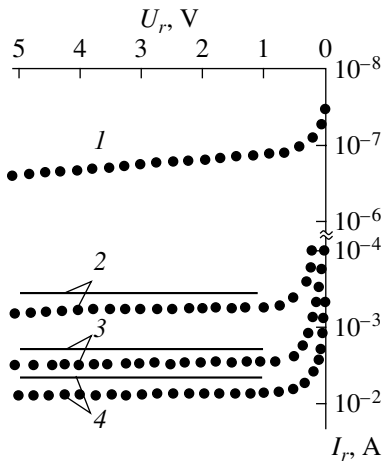


Fig. 4. Dependence of the reverse current I_r on voltage U for structure N1 at temperatures T, K: *1*—363; *2*—493, *3*—533; *4*—553. Points, experiment; curves, calculation by the thermionic emission theory.

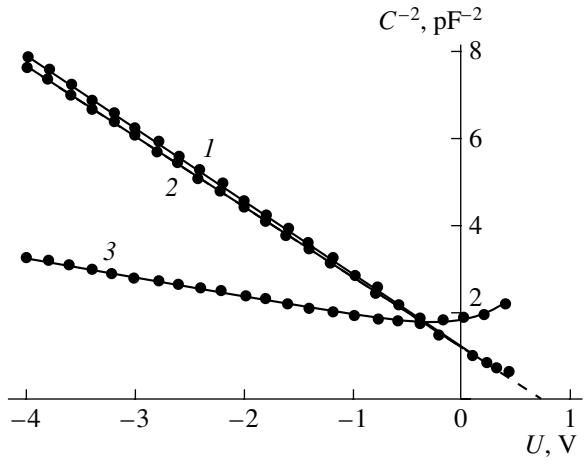


Fig. 5. Dependence of the capacitance C on voltage U for structure N2 measured immediately after fabrication (*1*) and after annealing. The annealing temperatures T_{ann}, K: *2*—573, *3*—582. Curves for annealing at temperatures 533, 543, 553, and 563 K coincided with curve *1*. Measured at 295 K.

(b) The temperature at which the rectifying Ni/GaAs contact became ohmic (T_{Ohm}) is 720 K, which is about 100 K higher than the temperature of such a transition for contacts between liquid metal and GaAs. At this temperature the contact metal does not melt (the melting temperature of Ni being T_m = 1453°C), so the dissolution of the semiconductor in the liquid metal is not a possibility. We suggest that at these temperatures a chemical interaction takes place between the semiconductor and the metal, resulting in the disappearance of the thin near-surface layer of the semiconductor, where the density of surface states is high and which is responsible for the rectifying properties of the contact.

Note that the interaction between Ni and GaAs, as determined from measurements of the backscatter of α particles [8], starts at ~470 K. As a result of this interaction, a new phase is formed, which appears at ~570 K at the structure surface.

In [9], it was found that this new phase is a metastable compound Ni_2GaAs; its formation raises the barrier height to 0.84 eV. At temperatures of 620–820 K, this phase decomposes into two compounds, NiGa and NiAs, and the barrier height drops significantly. At a temperature of 870 K, this process leads to the formation of the ohmic contact. According to our data, formation of the ohmic contact occurred at 720 K. However, this result was obtained not for annealing, as in [9], but for continuous heating.

2. Consider the evolution of the capacitance–voltage and current–voltage characteristics of the Ni/GaAs structures in the case of stepwise heating. The structures were heated to different temperatures and cooled to room temperature; measurements were then carried out.

2.1. The results on *C–U* characteristics can be summarized as follows.

(a) After annealing at temperatures of $T_{ann} \le 563$ K, the C^{-2}–*U* plots coincided with the starting characteristics (Fig. 5).

(b) After annealing at temperatures of $T_{ann} \ge 580$ K, the C^{-2}–*U* plots differed markedly from the starting characteristics (Fig. 5), which is evidence of the irreversible nature of the processes in the barrier contact.

2.2. The results on the evolution of the I_f–*U* characteristics are as follows.

(a) After annealing at temperatures of T_{ann} below a certain temperature T_0 = 553 K, the I_f–*U* characteristics coincided with the starting characteristics. (Fig. 6). Note that in the structures which had an ideality factor β for the starting I_f–*U* characteristic somewhat different from the theoretical value (for curve *1* β = 1.05), after annealing β was close to the theoretical value (curve *2*, β = 1.01–1.02) due to the disappearance of the intermediate dielectric layer.

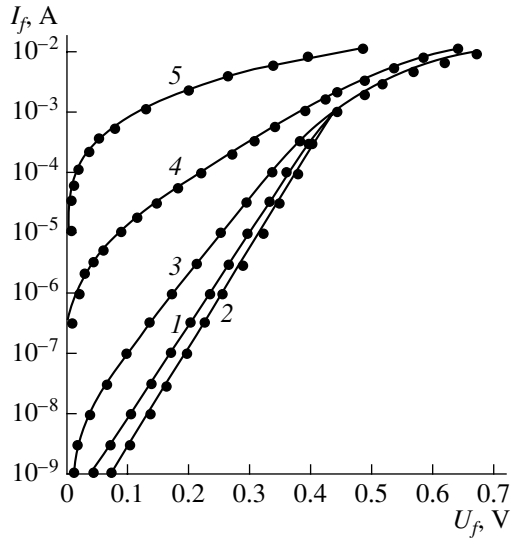


Fig. 6. Dependence of the forward current I_f on voltage *U* for structure N2 measured immediately after fabrication (*1*) and after annealing. T_{ann}, K: *2*—533, 543, and 553; *3*—563, *4*—573; *5*—582. Measured at 295 K.

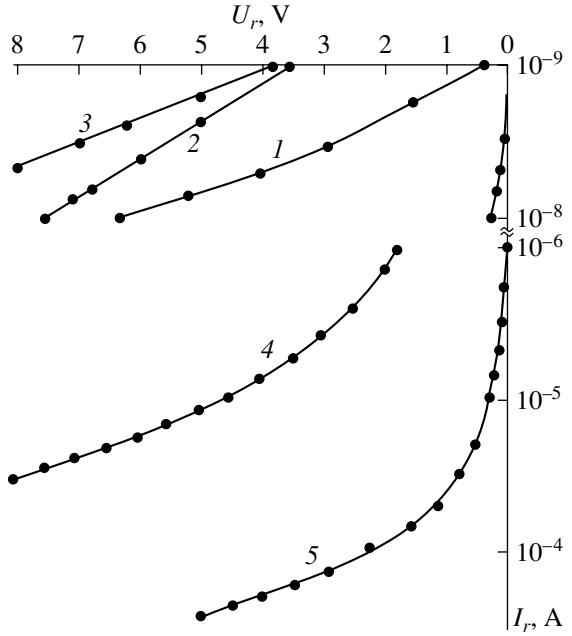


Fig. 7. Dependence of the reverse current I_r on voltage *U* for structure N2 measured immediately after fabrication (*1*) and after annealing. T_{ann}, K: *2*—543; *3*—553, *4*—563; *5*—573. Curve for annealing at the temperature 364 K coincided with curve *1*. Measured at 295 K.

(b) After annealing at temperatures of $T_{ann} \ge T_0$, the I_f–*U* characteristics showed excess currents although the structures remained rectifying (Fig. 6).

2.3. The results on the evolution of the I_r–*U* characteristics are as follows.

(a) After annealing at temperatures of T_{ann} below a certain temperature T_0 = 553 K, the I_r–*U* characteristics

were similar to the starting characteristics. Note that the structures in which an intermediate layer remained between the metal and the semiconductor during fabrication had lower reverse currents after annealing due to the disappearance of this layer during annealing (Fig. 7, curves *2* and *3*).

(b) After annealing at temperatures of $T_{ann} > T_0$, the reverse current increased by more than two orders of magnitude due to the beginning of an irreversible transition of the rectifying contact into an ohmic contact (Fig. 7). At temperatures of 200–300 K higher than T_0, the contact remained ohmic after cooling as well.

CONCLUSION

The metal–semiconductor contact, initially rectifying, becomes ohmic in the process of heating even before the recrystallized layer is formed. A chemical interaction takes place between the metal and the near-surface region of the semiconductor (in the case of Ni/GaAs structures) or dissolution of this near-surface region in liquid metal (in the case of Ga/GaAs structures [4]). From this fact, it follows that the newly formed surface of the semiconductor acquires properties varying from those of the starting surface. It can be assumed that on this surface, in contact with the metal, states differing from the initial ones arise that pin the surface Fermi level either in the conduction band or close to its bottom so that a potential barrier for the electrons does not form and the contact becomes ohmic.

REFERENCES

1. E. H. Rhoderick, *Metal–Semiconductor Contacts* (Claredon, Oxford, 1978; Radio i Svyaz', Moscow, 1982).
2. P. K. Kupta and W. A. Anderson, J. Appl. Phys. **69**, 3623 (1991).
3. N. Mochida, T. Honda, T. Shirasawa, *et al.*, J. Cryst. Growth **129/130**, 716 (1997).
4. Yu. A. Gol'dberg, E. A. Posse, and B. V. Tsarenkov, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 1510 (1986) [Sov. Phys. Semicond. **20**, 947 (1986)].
5. Yu. A. Gol'dberg, M. V. Il'ina, E. A. Posse, and B. V. Tsarenkov, Fiz. Tekh. Poluprovodn. (Leningrad) **22**, 555 (1988) [Sov. Phys. Semicond. **22**, 342 (1988)].
6. Yu. A. Gol'dberg and E. A. Posse, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 200 (1998) [Semiconductors **32**, 181 (1998)].
7. Yu. A. Gol'dberg, E. A. Posse, B. V. Tsarenkov, and M. I. Shul'ga, Fiz. Tekh. Poluprovodn. (Leningrad) **25**, 439 (1991) [Sov. Phys. Semicond. **25**, 266 (1991)].
8. V. G. Bozhkov, V. M. Zavodchikov, K. V. Soldatenko, *et al.*, Élektron. Tekh., Ser. 2: Poluprovodn. Prib., No. 7, 41 (1978).
9. F. Lahav, M. Eizenberg, and Y. Komem, J. Appl. Phys. **60**, 991 (1986).

Translated by B. Kalinin

Technical Physics, Vol. 46, No. 9, 2001, pp. 1133–1137. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 66–72.

OPTICS, QUANTUM ELECTRONICS

Optical Nonlinearity in Thin Films Exposed to Low-Intensity Light

A. V. Khomchenko and E. V. Glazunov

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

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

Received August 2, 2000; in final form, February 19, 2001

Abstract—Optical nonlinearity in semiconducting and insulating thin-film structures is studied by waveguide methods under self-action conditions at a light intensity less than 0.1 W/cm^2 and a wavelength of 630 nm. The optical properties vs. light intensity are similar for semiconducting and insulating films, multilayer thin-film structures, and semiconductor-doped glass films. It is demonstrated that the optical nonlinearity and the nonlinear optical constants depend on the interface condition. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Optical devices for data processing employ materials with the refractive index and the absorption coefficient dependent on the light intensity. The highest values of the nonlinear optical constants have been found in semiconductor materials. Therefore, at present, emphasis is on developing new semiconductor materials and structures with still higher optical nonlinearity [1]. At a relatively low intensity of incident light, one can easily achieve the desired level of the controllable power in optical waveguides owing to their small transverse dimensions [2]. That is why optical nonlinearity in waveguide semiconductor structures is of great interest.

In semiconductors, optical nonlinearities due to various mechanisms are observed at a light intensity of no less than ~10 W/cm^2. In this article, we observed the nonlinear variation in the refractive index and the absorption coefficient of thin films at an incident light intensity lower than 0.1 W/cm^2. The nonlinear refractive index and the absorption coefficient were measured to be ~10^{-3} cm^2/W.

EXPERIMENTAL

Amorphous and polycrystalline thin-film structures were obtained by vacuum deposition. The arsenic sulfide films were deposited by thermal evaporation. The zinc selenide, tin dioxide, and quartz glass films were produced by rf sputtering of ceramic targets. The GaSe films were prepared by electron-beam evaporation. The low-dimensional thin-film structures were obtained by sputtering semiconductor-doped glasses that were subsequently coated with lithium niobate, tin dioxide, and quartz glass. In the multilayer structures, the thicknesses of the quartz glass, lithium niobate, and tin dioxide layers were 70, 50, and 10–80 nm, respectively. The thickness of each of the layers was controlled by the evaporation time. The refractive indices of the quartz glass, lithium niobate, and tin dioxide films were 1.476, 2.160, and approximately 2.0, respectively. The tin dioxide and lithium niobate layers of the multilayer structure were obtained by sputtering the ceramic and single-crystal (*z* cut) targets, respectively. The glass layers were produced by sputtering OS-12 and KV glasses. KV and K8 optical glasses were used as substrates. In the course of deposition, the substrate temperature did not exceed 250°C. All films were applied in argon–oxygen (4 : 1) atmosphere except for the ZnSe films and the semiconductor-doped glass films, which were obtained in argon atmosphere.

The refractive indices and the absorption coefficients of the films were determined by the waveguide method at the wavelength of an He–Ne laser. The waveguide mode in the structure was excited by a coupling prism. We measured the spatial distribution of the intensity in the cross section of the beam reflected from the base of the coupling prism. A photodetector array measured the spatial distribution of the intensity of the reflected beam in the focal plane of the objective. From the distribution obtained, we calculated the real h' and the imaginary h'' parts of the propagation constant h of the waveguide mode ($h = h' + ih''$, where $h' = \mathrm{Re}h$ and $h'' = \mathrm{Im}h$) [3]. Using the values of the complex h for any two modes, one can determine the refractive index, the absorption coefficient, and the thickness of the film [4].

We determined the nonlinearity parameter of the waveguide structure by recording the shape of the signal when the intensity of the incident light I was gradually increased under self-action conditions. Based on the refractive index and the thickness of the film, and knowing the field of the waveguide mode, one can find the nonlinear refractive index n_2 and absorption coefficient k_2 [5]. The incident power was varied from 0.5 to 500 μW. The radius of the beam at the base of the coupling prism was no more than 200 μm.

1063-7842/01/4609-1133$21.00

RESULTS AND DISCUSSION

We measured the nonlinear optical constants of thin vitreous arsenic sulfide films at incident light intensities ranging from 10 to 100 W/cm^2. Experiments were carried out under self-action conditions at wavelength of 0.63 μm [6]. The value of the nonlinear constant $n_2 = 1.5 \times 10^{-5}$ cm^2/W agrees with data in [7]. However, the measurements at intensities lower than 0.1 W/cm^2 revealed a strong nonlinear dependence of the optical parameters of the thin films on the incident intensity. Figure 1 plots $\Delta h'(I) = h'(I) - h'(I_0)$ versus the incident light intensity for the TE polarization waveguide mode of zero order excited in the As_2S_3 film (I_0 is the highest intensity at which the nonlinear effects are not yet observed). Hereafter, we are dealing with the values of h' that are related to the measured resonance angle φ of waveguide mode excitation by the expression $h' = k_0 n_p \cos\varphi$, where n_p is the refractive index of the coupling prism and k_0 is the wave number. In the linear case, h' is the real part of the propagation constant of the waveguide mode. The nonlinear constant was found to be $n_2 = 2.65 \times 10^{-3}$ cm^2/W. Similar curves were obtained upon exciting waveguide modes of higher orders and other polarizations. The changes in the optical parameters of the film due to its heating by light absorption were five orders of magnitude smaller than their values [6]. This fact points to the nonthermal character of the nonlinearity observed.

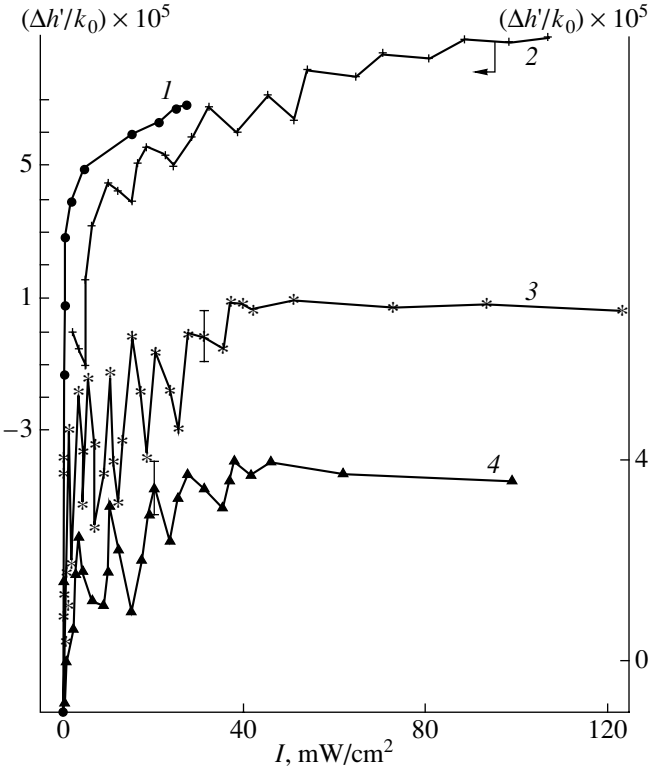


Fig. 1. Propagation constant versus intensity of incident light for (*1*) As_2S_3 film and ZnSe films deposited at the substrate temperature (*2*) 180, (*3*) 240, and (*4*) 250°C.

Thin polycrystalline zinc selenide films also exhibit nonlinear changes in their optical properties as the optical power was varied [8]. Curve *2* in Fig. 1 shows the propagation constant of the mode vs. incident light intensity. The significant spread of the h' values far exceeding the experimental error ($\delta h' = 5 \times 10^{-6}$) necessitated a more thorough study of the nonlinear properties of the zinc selenide films. The h' versus I curves were nonmonotonic, and their run was very complex. It was found that the nonmonotonicity of the curves and the value of the constant n_2 correlate with the crystal perfection of the film. All films were polycrystalline with their grains having cubic structure and preferred (022) orientation parallel to the substrate. Diffraction patterns did not show any other structures [8]. Figure 1 (curves *2–4*) plots $\Delta h'(I)$ versus I for the ZnSe films at substrate temperaturess of 140, 180, and 250°C, respectively. The mean grain sizes in the films varied between 19.7 and 12 nm.

The thin-film structures obtained by sputtering the semiconductor-doped OS-12 glasses and by sputtering the ceramic targets containing SiO_2 and CdSe also exhibit the nonmonotonic dependence of the waveguide properties on the incident intensity (Fig. 2). In these structures, h' also tends to decrease and the absorption tends to rise. The peaks in the curves again exceed the experimental error. The optical nonlinearity here depends on the size of the semiconductor grains in the glass matrix [9]. One can change the size of these grains by thermal annealing or by varying the conditions for film deposition. The curves $\Delta h'(I)$ in Fig. 2 were obtained for the OS-12 glass films prior to and after 6 h of thermal annealing at 400°C. The increase in the substrate temperature also affects the nonmonotonic dependence of the optical properties on the light intensity (curves *1–3* in Fig. 2).

Similar studies were carried out for glassy arsenic sulfide films; polycrystalline zinc selenide, zinc oxide, and gallium selenide films; and the doped glass films. The radiation wavelength 0.63 μm does not necessarily coincide with the absorption edge of these materials, yet all the structures exhibit nonlinear optical properties. Recently, Gaponenko [10] proposed that the reason for the optical nonlinearity is the effect of the surface and the interfaces and made an attempt to simulate such a nonlinear medium with a multilayered structure.

The structure was produced by sequentially depositing lithium niobate and fused quartz. Figure 3 plots h' versus the incident radiation intensity for the multilayer structures. Note the nonmonotonic run of these curves and also the fact that the number of peaks (five or six) coincide with the number of layers in the structure produced by sputtering the lithium niobate target. For the intensity ranges *I* and *II* indicated in Fig. 3, the nonlinear refractive index n_2 and the nonlinear absorption

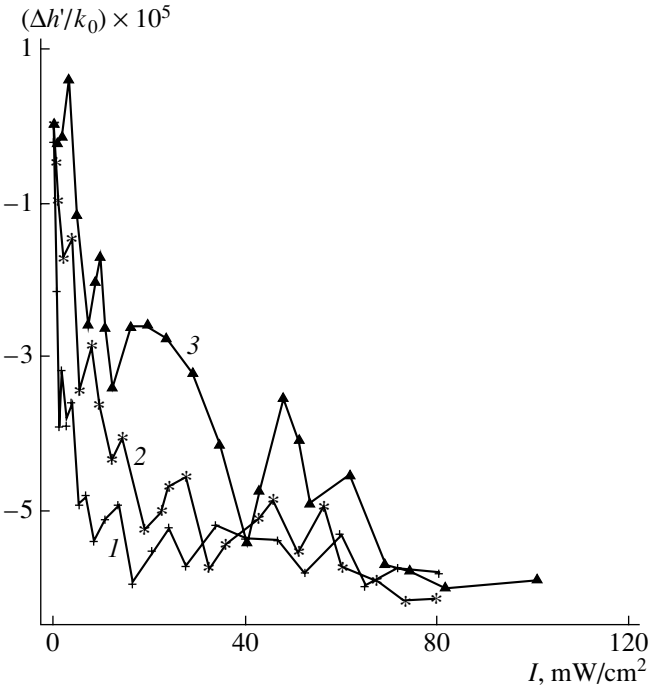


Fig. 2. $\Delta h'(I)$ versus intensity of incident light for the color glass films deposited at the substrate temperature (*1*) 140 and (*3*) 190°C. (*1*) Prior to and (*2*) after thermal annealing.

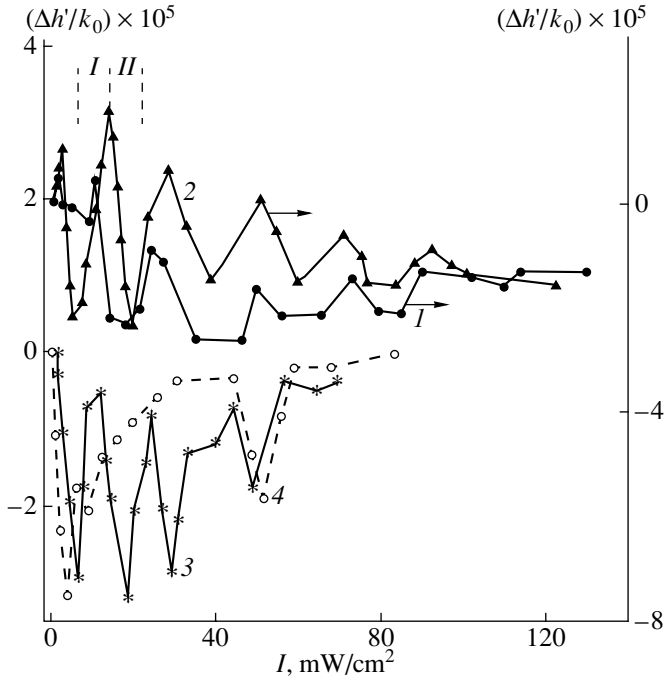


Fig. 3. $\Delta h'(I)$ versus intensity of incident light for the zero-order mode in multilayer structures containing (*1*) five and (*2*) six layers of lithium niobate, and (*3*) three and (*4*) one conducting SnO_2 layers.

coefficient k_2 were $n_2^{(I)} = -2.1 \times 10^{-3}$ cm^2/W, $k_2^{(I)} = -6.2 \times 10^{-3}$ cm^2/W, $n_2^{(II)} = 3.1 \times 10^{-3}$ cm^2/W, and $k_2^{(II)} = -5.1 \times 10^{-3}$ cm^2/W (curve *1*). The large values of n_2 and k_2 at low intensities make it possible to use these structures as nonlinear optical media.

Another type of multilayer structure was produced by the sequential deposition of the linear (in the usual sense) materials: conducting tin dioxide and insulating SiO_2. Such a structure with a thickness of the layers of about 10 nm simulates a low-dimensional nonlinear medium. In Fig. 3, $\Delta h'$ is plotted against the light intensity for a structure containing three tin dioxide layers separated by silicon dioxide layers. It is seen that the behavior of the curve correlates with the number of layers in the structure. The thicknesses of the conducting (tin dioxide) layers were 12, 24, and 36 nm. Curve *3* in Fig. 3 has three peaks with different widths. It can be concluded that the thicker the layer, the wider the corresponding peak. The third peak in curve *3* (d_{layer} = 36 nm) is similar in shape to that in curve *4* for the thicker film (d_{layer} = 120 nm).

The decrease in h' with increasing intensity corresponds to the decrease in the refractive index of the film. In this case, the absorption coefficient grows. The increase in the absorption coefficient and the decrease in the refractive index are usually related to the rise in the charge carrier concentration in the semiconductor material [11]. Based on our experimental data, we assume that the optical nonlinearity in the structures considered is associated with electron processes at the semiconductor–insulator interfaces.

If the condition $|k_2^{(I)}| > |k_2^{(II)}|$ is met, the processes taking place at the interfaces increase in the total absorption of the multilayer structure as the incident light intensity grows. The additive nature of this effect explains the tendency for the absorption to grow. In the opposite case ($|k_2^{(I)}| < |k_2^{(II)}|$), the film will become more and more transparent with an increase in the intensity. We suggest that these processes are responsible for the nonmonotonic behavior of the optical properties of the structures at low incident intensities.

For the multilayer structures, the nonlinearity of the optical properties substantially depends on the optical quality of the insulating layer. Figure 4 shows the $\Delta h'(I)$ curves for SnO_2 films deposited onto three different substrates in a single process cycle. The substrates were quartz glass and quartz glass covered by an SiO_x film. The film was applied by sputtering fused quartz under various conditions. The difference in the composition of the SiO_x films accounts for the difference in their absorption coefficients: 1.5×10^{-5} (curve *2*) and 5×10^{-6} (curve *3*). The thickness of these films was 1 μm. Even an inert amorphous substrate is known to affect the properties of thin-film waveguides [12]. In our case, the nonlinearity of the optical properties is also more pronounced in the waveguide structures where the sub-

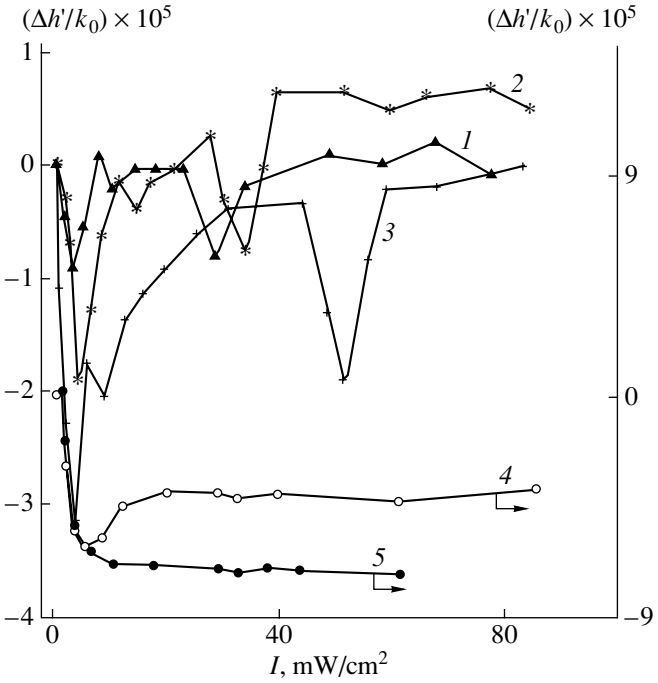


Fig. 4. $\Delta h'(I)$ versus intensity of incident light for multilayered structures (*1*) SnO_2–substrate, (*2*) SnO_2–SiO_x with the absorption coefficient of the SiO_x film 1.5×10^{-5}, and (*3*) SnO_2–SiO_x with the absorption coefficient of the SiO_x film 5×10^{-6}. Curves *4* and *5* refer to prism–buffer layer–SnO_2 structures in (*4*) air and (*5*) water vapor, respectively.

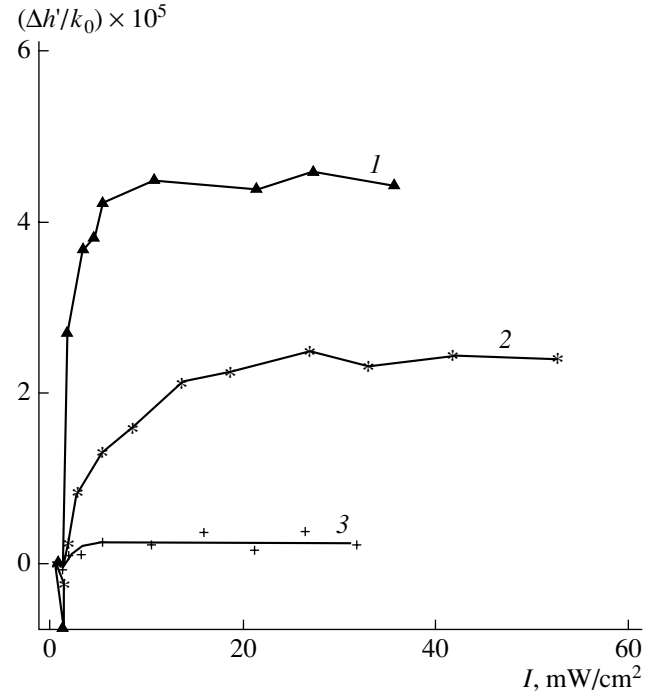


Fig. 5. $\Delta h'(I)$ versus intensity of incident light for the quartz films with the absorption coefficients (*1*) 2×10^{-5}, (*2*) 9×10^{-6}, and (*3*) 3×10^{-6}.

strate and the waveguide are separated by the imperfect buffer SiO_x films with the higher absorption coefficients.

Note that the insulating films also exhibit optical nonlinearity at low incident intensities. Nonlinear optical properties were observed in the films obtained by sputtering fused quartz. The value of the nonlinear constant n_2 (about 10^{-7} cm^2/W) was significantly smaller than those for the semiconductor films or for the semiconductor-doped glass films. However, the nonlinear dependence of the optical parameters on the light intensity was reliably detected only in the nonstoichiometric films. We did not observe this effect when the optical losses in the waveguide film were about 2 dB/cm (Fig. 5).

The range I of photoinduced absorption (Fig. 3) is most likely to be related to the capture of light-induced charge carriers by localized states in the imperfect SiO_2 film. The nature of the localized states is usually specified by defects in the film. Therefore, it is expedient to study the spatial distribution of defects near the interfaces in thin-film structures.

If the nonmonotonicity of the curves is related to the surface-induced modification of the electron states, the run of the curves can be varied by introducing gas admixtures into the medium surrounding the film. Gas molecules adsorbed on the surface of a thin-film structure can both induce and remove the surface states [13]. To study these phenomena, we made a waveguide structure consisting of silicon dioxide and tin dioxide films sequentially deposited on the base of a glass prism. In this case, the surface of the semiconductor film can be exposed to the gas admixtures. Water vapor is known to intensely passivate the surface states [14]. Therefore, it is logical to assume that the presence of water vapor would suppress the decrease in the absorption (and the increase in h'). Curves *4* and *5* in Fig. 4 show the variation in the parameters of the waveguide film with incident intensity in air and in the water vapor atmosphere, respectively. The low-intensity portion is where the refractive index drops and the absorption increases is retained. However, when exposed to water vapor, the film did not become transparent. Thus, this can be considered as proof that our assumption of the effect of the interfaces on the nonlinear behavior of the optical properties of the multilayer structures is valid.

In our opinion, the above considerations are also applicable to nonlinear structures based on doped glasses (Fig. 2). In this case, by a semiconductor–insulator interface, we mean the surface of the semiconductor grains embedded in the glass matrix. Immediately after deposition, the films have statistically equal grain sizes (curve *1*). Thermal annealing changes the grain size distribution (curve *2*). Similar charges are observed when the substrate temperature rises (curve *3*). By analogy with the multilayer structures (curve *3* in Fig. 3), we can assume that the wider the peaks, the larger the grain size.

Our results may indicate the similarity of the processes that take place in these thin-film structures and are responsible for the nonmonotonic dependence of the optical properties on the radiation intensity.

Irradiation of semiconductors by light generates electron–hole pairs, which diffuse into the semiconductor. As the thickness of the semiconductor film is much smaller than that of the screening layer, we are dealing with the surface generation of charge carriers or, more precisely, with the generation of carriers in the near-surface layer of the semiconductor. Then, the equation for the charge carrier balance at the surface is given by

$$g_s = \frac{1}{e} j_s(O) + S\delta n_s,$$

where g_s is the rate of carrier surface generation, $S\delta n_s$ is the rate of carrier surface recombination, and $j_s(O)$ is the current of the carriers near the surface [15].

This equation implies that g_s is independent of the absorption coefficient of the film. Since the wavelength of the incident light does not fall into the fundamental absorption band and the concentration of surface states can reach the concentration of atoms at the surface of the film, $j_s(O) < 0$. In this case, the charge carriers move toward the surface of the semiconductor. Filling the free energy levels of the surface states decreases the absorption in the film. Note also that the surface states do not all take part in recombination. Some of them are involved in band–energy level transitions, so that the energy level is a trap. Then, photoinduced (enhanced) absorption in the thin-film structures can be treated within the two-level model of optically recharging deep impurity levels in the band gap.

The competition of these processes at the interfaces gives rise to the nonmonotonic dependence of the optical properties on the intensity of the probing beam.

CONCLUSION

With the waveguide methods, we studied optical nonlinearity in semiconductor and dielectric thin-film structures at radiation intensities below 0.1 W/cm^2 and at a wavelength of 630 nm. The general tendencies in the dependences of the optical properties on the radiation intensity for multilayer structures, polycrystalline films, and semiconductor-doped glass films were considered. In low-dimensional semiconductor structures, optical nonlinearity depends on the quality of the semiconductor–dielectric interface; and in polycrystalline films, on the degree of crystallinity and the grain size. Our approach suggests that the nonlinear variation of the optical properties in the thin-film structures is due to the generation, capture, and recombination of charge carriers at the interfaces of the semiconductor film.

ACKNOWLEDGMENTS

We are grateful to A. I. Voĭtenkov for fruitful discussions. This work was supported by the Belarussian Foundation for Basic Research.

REFERENCES

1. H. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic, New York, 1985; Mir, Moscow, 1988).
2. H. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, Englewood Cliffs, 1984; Mir, Moscow, 1988).
3. V. P. Red'ko, A. A. Romanenko, A. B. Sotskiĭ, and A. V. Khomchenko, RF Patent No. 2022247, Byull. Izobret., No. 20 (1994).
4. A. B. Sotskiĭ, A. A. Romanenko, A. V. Khomchenko, and I. U. Primak, Radiotekh. Élektron. (Moscow) **44** (5), 1 (1999).
5. A. B. Sotskiĭ, A. V. Khomchenko, and L. I. Sotskaya, Pis'ma Zh. Tekh. Fiz. **20** (16), 49 (1994) [Tech. Phys. Lett. **20**, 667 (1994)].
6. A. B. Sotskiĭ, A. V. Khomchenko, and L. I. Sotskaya, Opt. Spektrosk. **78**, 502 (1995) [Opt. Spectrosc. **78**, 453 (1995)].
7. A. Yu. Vinogradov, É. A. Smorgonskaya, and E. I. Shifrin, Pis'ma Zh. Tekh. Fiz. **14**, 642 (1988) [Sov. Tech. Phys. Lett. **14**, 287 (1988)].
8. A. V. Khomchenko, Zh. Tekh. Fiz. **67** (9), 60 (1997) [Tech. Phys. **42**, 1038 (1997)].
9. S. Sh. Gevorkyan and N. V. Nikonorov, Pis'ma Zh. Tekh. Fiz. **16** (13), 32 (1990) [Sov. Tech. Phys. Lett. **16**, 494 (1990)].
10. S. V. Gaponenko, *Optical Properties of Semiconductor Nanocrystals* (Cambridge Univ. Press, Cambridge, 1998).
11. A. V. Khomchenko and V. P. Red'ko, Proc. SPIE **1932**, 14 (1993).
12. R. G. Hunsperger, *Integrated Optics: Theory and Technology* (Springer-Verlag, Berlin, 1984; Mir, Moscow, 1985).
13. A. Zangwill, *Physics at Surfaces* (Cambridge Univ. Press, Cambridge, 1988; Mir, Moscow, 1990).
14. *The Physics of Hydrogenated Amorphous Silicon*, Vol. 2: *Electronic and Vibrational Properties*, Ed. by J. D. Joannopoulos and G. Lucovsky, with contributions by D. E. Carlson *et al.* (Springer-Verlag, New York, 1984; Mir, Moscow, 1988).
15. V. L. Bonch-Bruevich and S. G. Kalashnikov, *Physics of Semiconductors* (Nauka, Moscow, 1977).

Translated by A. Chikishev

Technical Physics, Vol. 46, No. 9, 2001, pp. 1138–1142. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 73–78.

Optimization of the Transfer Function of an Acoustooptic Cell with an Apodized Piezoelectric Transducer

V. N. Parygin, A. V. Vershubskiĭ, and E. Yu. Filatova

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

Received October 27, 2000

Abstract—It is shown that an apodized piezoelectric transducer can significantly reduce the side lobe level of the acoustooptic cell transfer function. Series and symmetric connections of the transducer sections and measures to suppress the effect of spurious elements arising in the electric circuit are proposed. In particular, the effect of spurious capacitances and inductances on the frequency response of the transducer is studied. It is shown that they violate the optimal condition for the suppression of the transfer function side lobes, especially at high frequencies. It is shown by calculations that the effect of spurious elements can be eliminated by the insertion of additional capacitors at 80 and 150 MHz. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The field of knowledge of acoustooptics lies at the border between physics and technology. It studies the interaction between electromagnetic and acoustic waves and offers a variety of applications based on this phenomenon. The interaction between light and acoustic waves is used for controlling coherent light in modern optics, optoelectronics, and laser technology. Acoustooptic devices are capable of controlling the amplitude, frequency, polarization, spectrum, and direction of a light beam. An important field of application of acoustooptic effects is data processing systems, where microwave signals are processed in real time [1].

Acoustooptic effects are employed in active optical devices that can entirely control a light beam and process information carried by optical and acoustic waves. The core of these devices is an acoustooptic cell that consists of a working medium where light–acoustic wave interaction occurs and of an acoustic radiator (usually, a piezoelectric transducer). There are a number of acoustooptic devices designed for various purposes: deflectors, modulators, filters, processors, etc. [2–4].

The acoustooptic interaction phenomenon relies on the photoelasticity effect, i.e., on the property of a medium to change its refractive index under the action of elastic stress. Due to this effect, an acoustic wave propagating in an optically transparent medium is, in essence, a phase grating that moves with the speed of sound v. When passing through the acoustic field, the light is diffracted by refractive index nonuniformities.

The acoustooptic cell is characterized by its transfer function, i.e., by the dependence of the intensity of diffracted light on its wavelength at a given sound frequency f. Along with the main lobe, the transfer function usually exhibits significant side lobes. The side lobes substantially narrow the dynamic range of acoustooptic devices. Recently, a number of works [5–10] aimed at suppressing the side lobes have appeared.

Acoustooptic filters based on acoustooptic cells can be collinear and noncollinear. In the former, the transfer function is controlled by applying appropriate acoustic pulses instead of a continuous signal to the cell. The width of the control pulse specifies the filter passband, while its waveform defines the shape of the transfer function. This has been demonstrated experimentally and theoretically in [8, 10].

Unfortunately, these methods for optimizing the transfer function cannot be applied to the usual orthogonal geometry of the acoustooptic interaction. Therefore, in this paper, we consider the possibility of reducing the side lobes by apodizing the transducer with regard for the arising parasitic elements.

BASIC EQUATIONS FOR AN APODIZED TRANSDUCER

Consider the typical orthogonal geometry of the acoustooptic interaction, where the sound beam is generated by a rectangular piezoelectric transducer and the light wave is considered to be plane. Then, the system of equations that describes the acoustooptic interaction by relating the amplitudes of the transmitted, E_t, and diffracted, E_d, light waves is written as

$$\begin{aligned} \frac{dE_t}{dx} &= -\frac{q}{2}E_d(x)\exp(-j\eta x), \\ \frac{dE_d}{dx} &= \frac{q}{2}E_t(x)\exp(j\eta x). \end{aligned} \tag{1}$$

Here, q is a coefficient proportional to the acoustic wave intensity; $\eta = (\mathbf{k}_t + \mathbf{K} - \mathbf{k}_d) \cdot \mathbf{e}_x$ is the detuning parameter; $\mathbf{k}_t$, $\mathbf{k}_d$, and $\mathbf{K}$ are the wave vectors of the incident and diffracted light waves and of the acoustic

1063-7842/01/4609-1138$21.00

wave, respectively; $\mathbf{e}_x$ is the unit vector along the x axis, and $(\mathbf{k}_t + \mathbf{K} - \mathbf{k}_d) \cdot \mathbf{e}_x$ is the scalar product. A solution to system (1) under the boundary conditions

$$E_d(0) = 0, \quad E_t(0) = E_i \tag{2}$$

is given by

$$E_d(l) = \exp(j\eta l/2)\frac{qE_i}{\sqrt{q^2+\eta^2}}\sin\sqrt{q^2+\eta^2}\frac{l}{2},$$

$$E_t(l) = E_i\exp(-j\eta l/2)\left(\cos\sqrt{q^2+\eta^2}\frac{l}{2}\right. \tag{3}$$

$$\left. + \frac{j\eta}{\sqrt{q^2+\eta^2}}\sin\sqrt{q^2+\eta^2}\frac{l}{2}\right),$$

where E_i is the amplitude of incident light. Formulas (3) define the amplitude of the light waves at the exit from the cell (at $x = l$). The diffracted light intensity at the exit from the cell, $I_d = E_d(l)E_d^*(l)$, depends on the Raman–Nath parameter ql and the detuning parameter ηl. The latter depends on the sound frequency $f = Kv/2\pi$ and the wavelength of incident light $\lambda = 2\pi n_t/k_t$. Here, v is the acoustic velocity and n_t is the refractive index of the medium for incident light. Thus, $P(\eta l) = I_d(\eta l)/I_d(0)$ is the transfer function of the acoustooptic cell.

Relationship (1) shows that $E_d(\eta)$ is the Fourier transform of the product qE_t, i.e., of the product of the transverse amplitude distribution of the acoustic field by the amplitude of incident light. To change the transfer function, it is therefore necessary to vary the distribution of the acoustic field amplitude. In the case of orthogonal geometry, this problem is very difficult. However, we can divide the transducer into several sections and apply different voltages to them. With such an approach, the acoustic beam will consist of several layers, each having its own sound intensity.

Mathematically, the solution to the problem can be described, as before, by Eqs. (1) if q is assumed to be constant within each section. At the exit from the first section, the solution is thus given by formulas (3) by q and l replaced with q_1 and l_1, respectively. Then, Eqs. (1) can be solved for the case $q = q_2$ under the initial conditions $E_t = E_t(l_1)$ and $E_d = E_d(l_1)$. As a result, we obtain the amplitudes of the incident and diffracted light at the exit from the second section: $E_t = E_t(l_2)$ and $E_d = E_d(l_2)$, where $L_2 = l_1 + l_2$. These amplitudes are the initial conditions as applied to Eqs. (1) for the third section, and so on. The amplitudes at the exit from the $(k+1)$th section can thus be expressed in terms of the amplitudes $E_t(L_k)$ and $E_d(L_k)$ for the previous section by the recurrent formulas

$$E_t(L_{k+1}) = \exp(-j\eta l_{k+1}/2)\left\{E_t(L_k)\cos\xi_k\right.$$
$$\left. + \frac{j\eta E_t(L_k) - q_{k+1}E_d(L_k)e^{-j\eta L_k}}{\sqrt{q_{k+1}^2+\eta^2}}\sin\xi_k\right\},$$
$$E_d(L_{k+1}) = \exp(j\eta l_{k+1}/2)\left\{E_d(L_k)\cos\xi_k\right. \tag{4}$$
$$\left. - \frac{j\eta E_d(L_k) - q_{k+1}E_t(L_k)e^{j\eta L_k}}{\sqrt{q_{k+1}^2+\eta^2}}\sin\xi_k\right\}.$$

Here, q_k is a parameter proportional to the amplitude of the acoustic wave in the kth section, $\xi_k = \sqrt{q_{k+1}^2+\eta^2}(l_{k+1}/2)$, and $L_k = \sum_{i=1}^{k} l_i$ is the total length of k sections of the piezoelectric transducer.

Relationships (4) are the solution of system (1) at $q = q_{k+1}$ for arbitrary amplitudes of the incident and diffracted light at the entrance to the $(k+1)$th section of the piezoelectric transducer. For example, at $k = 0$ and $E_d(0) = 0$, formulas (4) yield solution (3). Relationships (4) are valid at any q_k. However, for weak acoustooptic interaction, characterized by a small Raman–Nath parameter ($q_k l_k \ll 1$), the terms containing $(q_k l_k)^2$ can be neglected and relationships (4) can be written in a simpler form

$$E_t(L_{k+1}) = E_t(L_k) - 0.5q_{k+1}l_{k+1}e^{-j\zeta}E_d(L_k)\mathrm{sinc}(\zeta/\pi),$$
$$E_d(L_{k+1}) = E_d(L_k) + 0.5q_{k+1}l_{k+1}e^{j\zeta}E_t(L_k)\mathrm{sinc}(\zeta/\pi), \tag{5}$$

where $\zeta = \eta l_{k+1}/2$ and $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$. With boundary condition (2) at the entrance to the first section, we successively calculate the amplitudes at the entrance to each section. As a result, Eqs. (5) yield

$$E_t(L_{k+1}) = E_i,$$
$$E_d(L_{k+1}) = 0.5E_i\sum_{m=1}^{k} q_m l_m e^{j\zeta_m}\mathrm{sinc}(\zeta_m/\pi), \tag{6}$$

where $\zeta_m = \eta l_m/2$.

Thus, successively applying relationship (4) to each of the layers of the acoustic beam, we can calculate the distribution of the transmitted and diffracted light amplitudes at the exit from the acoustooptic cell. This distribution may be rather complex. By appropriately choosing the lengths l_k of the acoustic beam layers and the amplitudes q_k of the acoustic field, the side lobes of the transfer function may be significantly suppressed.

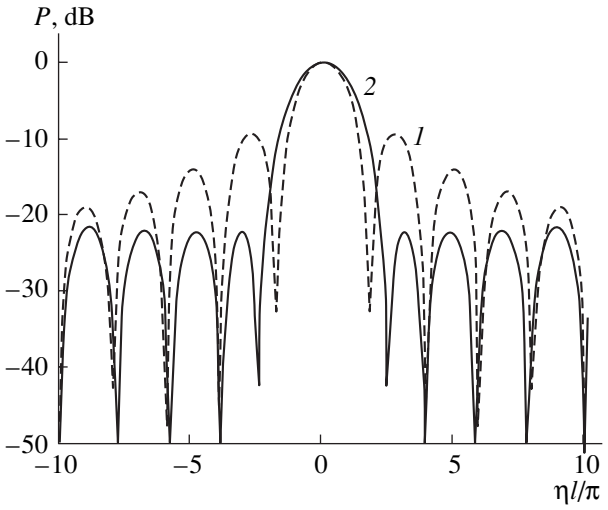


Fig. 1. Transfer function of the acoustooptic cell for strong interaction: (*1*) nonapodized cell and (*2*) piezoelectric transducer divided into nine sections of optimal lengths.

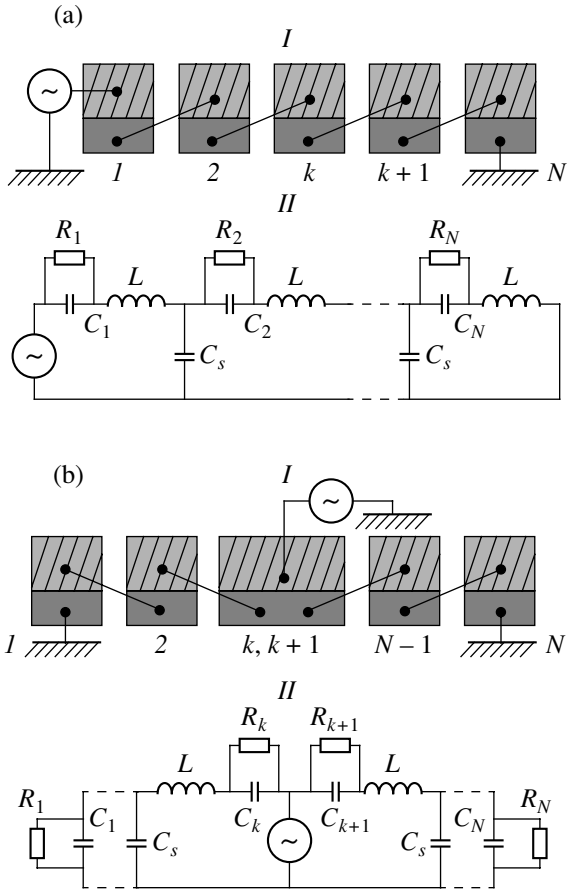


Fig. 2. Connection diagram for a multisection piezoelectric transducer: I, upper electrode; II, lower electrode. (a) The power supply is connected in series and (b) the power supply is connected to the central section of the transducer. $C_1, \ldots, C_N$ are the capacitances of the transducer sections; $R_1, \ldots, R_N$ are the radiation resistances of the sections; L is the spurious inductance of the wires; and C_s is the ground capacitance of a section.

THE EFFECT OF SPURIOUS ELEMENTS ON THE TRANSFER FUNCTION OF THE ACOUSTOOPTIC CELL

Curve *1* in Fig. 1 shows a typical transfer function of the acoustooptic cell with a nonapodized piezoelectric transducer. Here, P is the relative intensity of the diffracted light in decibels, $\eta l/\pi$ is the detuning parameter, and l is the length of acoustooptic interaction. The transfer function of this form has been widely discussed in the literature. A stepwise distribution of the acoustic field amplitude and a modified transfer function (curve *2* in Fig. 1) can be obtained by dividing the piezoelectric transducer into several electrically insulated series-connected sections of different lengths.

Figure 2 schematically illustrates a sectional piezoelectric transducer to which a power supply is connected (a) in series and (b) symmetrically. It should be noted that, in the symmetric connection (Fig. 2b), there is no need for dividing the central part of the transducer into two equal sections; therefore, in this case, the number of cuts and connections is less than in Fig. 2a by one. Each of the sections can be viewed as a capacitor whose value is proportional to the length of the section. The amplitudes of the acoustic field are proportional to the voltage applied to the sections of the transducer. Because of the series connection, the acoustic amplitude in each section is in inverse proportion to its capacitance and, consequently, to its length.

Such a transducer produces an acoustic beam consisting of several layers with different amplitudes. The longer the section, the lower the acoustic amplitude in it and conversely. Our calculations have shown that optimal proportions exist between the section lengths that minimize the side lobes of the transfer function. We calculated the parameters of optimal (in terms of side lobe suppression) piezoelectric transducers with various numbers of sections [11]. For example, curve *2* in Fig. 1 plots the transfer function for a piezoelectric transducer with strong acoustooptic interaction that is divided into nine sections ($l_1 = l_9 = 0.187l$, $l_2 = l_8 = 0.112l$, $l_3 = l_7 = 0.086l$, $l_4 = l_6 = 0.0785l$, and $l_5 = 0.073l$, where l_k is the length of the kth section).

This design reduces the side lobe level from –9.3 dB for an acoustooptic cell with a nonapodized transducer to –21.7 dB for an acoustooptic cell with an apodized transducer. Experimentally, the side lobe suppression with the optimally selected stepwise voltage distribution has been demonstrated in [12]. The larger the number of sections, the lower the side lobe level of the transfer function. However, a large number of sections requires a very high cutting accuracy, which may cause problems if the transducer is short. Moreover, a large number of electrical connections induces parasitics, which adversely affect the performance of the device.

The parasitics resulting in these circuits may substantially change the amplitudes and the phases of the voltages applied to the sections, because wires that connect adjacent sections always exhibit a certain induc-

tance. Also, each section exhibits a spurious ground capacitance. At low acoustic frequencies, the effect of these reactive components may be neglected. However, at high frequencies, they may significantly change the amplitude–phase distribution of the acoustic field. As a consequence, the condition for suppressing the side lobes will be violated. Figures 2a and 2b, respectively, show the equivalent circuits of piezoelectric transducers with the series and the symmetric connection of the power supply. Here, $C_1, \ldots, C_N$ are the capacitances of the sections; $R_1, \ldots, R_N$ are the radiation resistances of the sections, which describe the electric-to-acoustic energy conversion efficiency; L is the inductance of wires between the adjacent sections; and C_s is the ground capacitance of the sections.

It has been shown that, in these circuits, electric amplitude resonance occurs in the transducer sections at a frequency $f_0 = 1/2\pi\sqrt{LC}$, where $C = C_0/N$, N is the number of the sections, and $C_0 = \sum_{k=1}^{N} C_k$ is the total capacitance of the transducer. Figure 3 shows the voltage U_0 across the transducer providing a 100% diffraction efficiency versus ultrasonic frequency f at $N = 7$, $L = 15$ nH, $C_s = 10$ pF, and $C_0 =$ (*1*) 20 300, (*2*) 2100, and (*3*) 700 pF. These values are typical of modern acoustooptic piezoelectric transducers. The respective resonance frequencies are $f_0 = 24$, 75, and 130 MHz. It is seen that the voltage required for the 100% incident-to-diffracted light conversion is minimal at a frequency close to the resonance frequency for a particular C_0.

However, the higher C_0, the narrower the operating frequency band of the device, because at higher-than-resonance frequencies, the supply voltage necessary to obtain a high diffraction efficiency is large. The voltage amplitude distribution between the sections also depends on the sound frequency. Figures 4a and 4b illustrate this distribution for the sections connected as shown in Figs. 2a and 2b, respectively, with $N = 8$, $C_0 =$ 700 pF, $L = 15$ nH, and $C_s = 10$ pF. In Fig. 4, U_k is the voltage amplitude across the kth section. The amplitudes were calculated for 100% diffraction efficiency.

It should be noted that the curves in Fig. 4 are arbitrarily drawn to some extent, because the distributions are actually stepwise and the voltages indicated are constant within the respective section. The data points are connected only for the purpose of illustration. Curves *1* correspond to the optimal suppression of the transfer function side lobes at the resonance frequencies. At off-resonance frequencies, the optimal voltage amplitude distribution will be different. For the circuit shown in Fig. 2a, this distribution becomes asymmetric: at frequencies below or above the resonance frequency, the maximal amplitude tends to the left (curve *2*) or to the right (curve *3*). With the sections connected as shown in Fig. 2b, the voltage amplitude distribution among the sections remains symmetric but the amplitudes change.

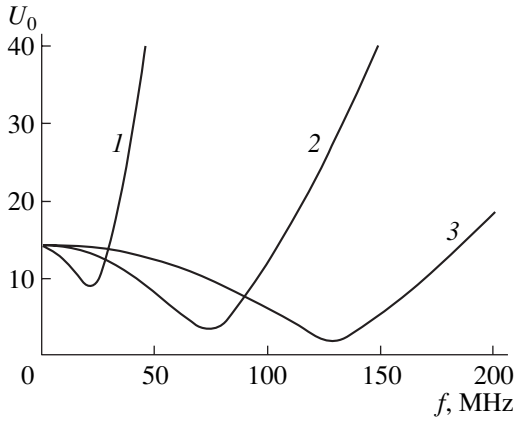


Fig. 3. Voltage U_0 across the piezoelectric transducer necessary a 100% diffraction efficiency versus acoustic frequency f.

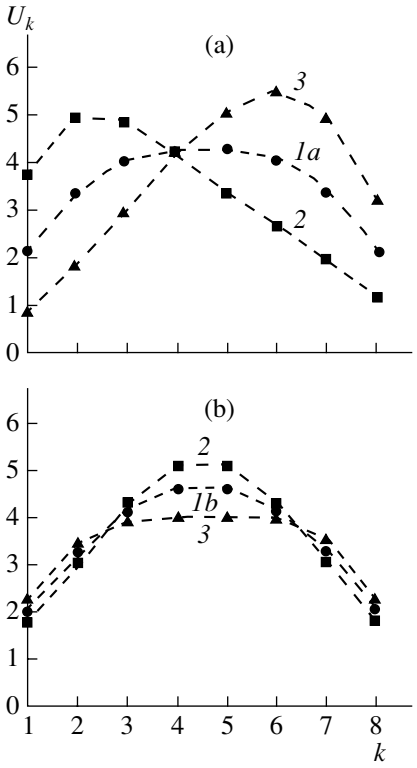


Fig. 4. Voltage amplitudes U_k on the sections of the transducer for the (a) series (Fig. 2a) and (b) symmetric (Fig. 2b) connections at $N = 8$, $C_0 = 700$ pF, $L = 15$ nH, $C_s = 10$ pF. $f_0 =$ (*1a*) 142.5 and (*1b*) 133 MHz; $f =$ (*2*) 100 and (*3*) 160 MHz.

The significant change in the voltage amplitude distribution increases the side lobe level, especially at high frequencies. Our calculations have shown that the effect of the spurious reactive components may be reduced by introducing additional trimming capacitors. These capacitors of value ΔC_k should be connected in parallel with C_k in Fig. 2. By appropriately selecting the

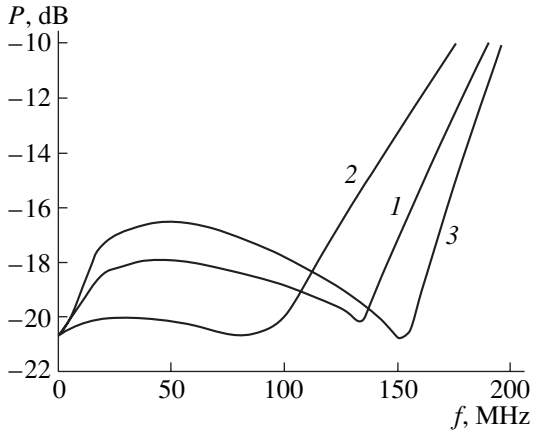


Fig. 5. Maximum side lobe level of the transfer function versus acoustic frequency f for the symmetric connection of the transducer sections at $N = 7$, $C_0 = 700$ pF, $L = 15$ nH, and $C_s = 10$ pF: (*1*) without parasitic suppression and (*2*, *3*) with parasitic suppression at 80 and 150 MHz, respectively.

values of ΔC_k, the section-to-section voltage amplitude distribution can be optimized at a particular frequency.

It should be noted that the trimming capacitors shift the characteristics shown in Fig. 3 towards lower frequencies. The best compensation has been shown to be achieved at the symmetric connection of the piezoelectric sections (Fig. 2b). For this case, the maximum side lobe level of the transfer function versus frequency is shown in Fig. 5 at $N = 7$, $C_0 = 700$ pF, $L = 15$ nH, and $C_s = 10$ pF. Here, curve *1* results in the absence of the trimming capacitors, while curves *2* and *3* show the suppression of the parasitics at 80 ($C_1 = C_7 = 2$ pF, $C_2 = C_6 = 4$ pF, $C_3 = C_5 = 10$ pF, and $C_4 = 30$ pF) and 150 MHz ($C_1 = C_7 = 60$ pF, $C_2 = C_6 = 25$ pF, $C_3 = C_5 = 6$ pF, and $C_4 = 0$), respectively.

From these data, we may conclude that, near the frequency for which the trimming capacitances are calculated, the side lobe level is lower than when $\Delta C_k = 0$. There exists a finite frequency band within which the side lobe level is sufficiently low, e.g., less than –20 dB. In particular, for the elimination of the parasitics at 80 MHz, this band is about 100 MHz wide, while at 150 MHz, its width is only about 20 MHz. Note also that at a high C_0, the elimination at above-resonance frequencies requires a high supply voltage (see Fig. 3). Nevertheless, it could be expected that, if the parameters of the spurious elements are known, for example, $C_0 = 700$ pF, one can find an operating frequency such that the side lobe level in its vicinity will be sufficiently low, e.g., below 20 dB.

CONCLUSION

Thus, the spurious elements arising in the apodized piezoelectric transducer may significantly affect the amplitude and phase distributions of the acoustic field on its sections, particularly at high frequencies, and, hence, violate the optimal condition for side lobe suppression. Such a situation should be avoided, because the use of the apodized transducer can significantly reduce the side lobes of the transfer function of the acoustooptic cell. However, the effect of the spurious elements can be eliminated by inserting additional capacitors into the circuit. We calculated their values at frequencies both below and above the resonance frequency. The symmetric connection of the supply voltage was shown to provide the most complete elimination of the spurious elements. For a seven-section piezoelectric transducer, the side lobe level as low as less than –20 dB within a certain frequency range can be reached.

REFERENCES

1. V. N. Parygin and V. I. Balakshiĭ, *Optical Information Processing* (Mosk. Gos. Univ., Moscow, 1987).
2. A. Korpel, *Acousto-optics* (Marcel Dekker, New York, 1988; Mir, Moscow, 1993).
3. L. N. Magdich, Izv. Akad. Nauk SSSR, Ser. Fiz. **44**, 1683 (1980).
4. V. I. Balakshiĭ, V. N. Parygin, and L. E. Chirkov, *Physical Principles of Acousto-optics* (Radio i Svyaz', Moscow, 1985).
5. M. M. Mazur, V. N. Shorin, *et al.*, Opt. Spektrosk. **81**, 521 (1996) [Opt. Spectrosc. **81**, 475 (1996)].
6. A. K. Zaĭtsev and V. V. Kludzin, Izv. Vyssh. Uchebn. Zaved., Élektron. **41** (10), 75 (1998).
7. V. Pustovoit and N. Gupta, EOS Topical Meetings Digest Series **24**, 35 (1999).
8. V. N. Parygin and A. V. Vershubskiĭ, Akust. Zh. **44**, 615 (1998) [Acoust. Phys. **44**, 529 (1998)].
9. V. N. Parygin and A. V. Vershubskiĭ, Radiotekh. Élektron. (Moscow) **43**, 1369 (1998).
10. V. N. Parygin, A. V. Vershubskiĭ, and K. A. Kholostov, Zh. Tekh. Fiz. **69** (12), 76 (1999) [Tech. Phys. **44**, 1467 (1999)].
11. V. N. Parygin, A. V. Vershoubskiy, and E. Yu. Filatova, J. Mod. Opt. **47**, 1501 (2000).
12. V. N. Parygin, V. Ya. Molchanov, and E. Yu. Filatova, EOS Topical Meetings Digest Series **24**, 45 (1999).

Translated by A. Khzmalyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1143–1150. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 79–87.

ELECTRON AND ION BEAMS, ACCELERATORS

RF Focusing of Ion Beams in the Axisymmetric Periodic Structure of a Linac

É. S. Masunov and N. E. Vinogradov

Moscow State Engineering Physics Institute (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia

e-mail: masunov@dinus.merphi.ru; masunov@edhem.merphi.msk.su

Received November 3, 2000

Abstract—A new approach to the analysis of charge particle motion in periodic resonant RF accelerators is suggested. A three-dimensional equation of motion in the Hamiltonian form is derived. This equation makes it possible to carefully study the relationship between the transverse and the longitudinal dynamics of the ion beams at low initial energies. General conditions for RF focusing in ion linacs are formulated. Basic results are compared with numerical simulation data for the beam dynamics in the polyharmonic field of the accelerating cavity. A version of an RF-focusing proton accelerator in which the current transmission coefficient is close to that in an accelerator with radio frequency quadrupole is described. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

It is known that the stable motion of an ion beam in a linac can be provided by using external focusing devices or specially configured accelerating fields (rf focusing). For low-energy ion accelerators, the latter approach seems to be more promising. Today, several approaches to RF focusing are known: alternating-phase focusing (APF), radio frequency quadrupole (RFQ), and undulator RF focusing (UrfF). The basic principles of APF in the single-wave approximation have been formulated in [1–3]. Later, APF has been considered in terms of a two-wave model where one wave is in synchronism with the beam [4, 5]. Subsequently, it has been shown [6] that the second nonsynchronous harmonic of the RF field should be taken into account in some cases. In [7], the situation with a large number of the spatial harmonics of a standing wave was treated in the short-gap approximation. The methods used in these works to describe APF suffer from disadvantages. The relationship between the longitudinal and transverse motions of the beam are considered incorrectly. The same is true for the effect of fast longitudinal oscillations on the beam dynamics. The averaging method used in [5] to analyze RF focusing is also not well-defined. Thus, the available APR theory incompletely discovers the RF focusing potentialities.

Currently, the problem of increasing the current and the current transmission coefficient in low-energy linacs is becoming more and more important. It cannot be tackled within conventional APF theory. Because of this, interest in axisymmetric RF focusing (ARF) has waned in recent years. In [8], RF focusing systems were considered with the method of averaging over fast oscillations (the so-called smooth approximation). It was shown that, in the smooth approximation, the analysis of three-dimensional beam dynamics is reduced to the analysis of the Hamiltonian function. The Hamiltonian of the system takes a simple form if some properties of solutions to the Maxwell equations are employed [8]. Later, such an approach was used to study beam acceleration and focusing in the polyharmonic field of the cavity in the one-particle approximation [9, 10].

This work extends the studies of ARF in periodic resonant structures [9, 10] with respect to the space charge field of the beam.

EQUATION OF MOTION IN SMOOTH APPROXIMATION

First we consider the equation of motion in the one-particle approximation without including the self-field of the beam. The external RF field of a periodic structure is represented as an expansion in spatial harmonics:

$$E_z = \sum_{n=0}^{\infty} E_n I_0(h_n r)\cos\left(\int h_n dz\right)\cos(\omega t),$$
$$E_r = \sum_{n=0}^{\infty} E_n I_1(h_n r)\sin\left(\int h_n dz\right)\cos(\omega t), \quad (1)$$

where E_n are the amplitudes of the harmonics at the axis, $h_n = h_0 + 2\pi n/D$, $h_0 = \mu/D$, μ is an oscillation phase advance, D is the period of the structure, and I_0 and I_1 are the modified Bessel functions of the zeroth and the first order.

In a polyharmonic field, the particle path can be represented as the sum of slow- and fast-varying compo-

nents. Averaging over fast oscillations, as was done in [8, 9], we come to the equation of motion

$$\frac{d^2\bar{\mathbf{R}}}{d\tau^2} = -\frac{\partial}{\partial\bar{\mathbf{R}}}\bar{U}_{\text{eff}}, \tag{2}$$

where $\bar{U}_{\text{eff}} = \bar{U}_0 + \bar{U}_1 + \bar{U}_2 + \bar{U}_3$ is the effective potential function. Here,

$$\bar{U}_0 = -\frac{1}{2}\bar{e}_s\beta_s[I_0(\bar{\eta}/\beta_s)\sin(\psi + \bar{\chi}/\beta_s) - (\bar{\chi}/\beta_s)\cos\psi],$$

$$\bar{U}_1 = \frac{1}{16}\sum_{n\neq s}\frac{\bar{e}_n^2}{(\Delta_{s,n}^-)^2}g_{s,n}(\bar{\eta}) + \frac{1}{16}\sum_{n}\frac{\bar{e}_n^2}{(\Delta_{s,n}^+)^2}g_{s,n}(\bar{\eta}),$$

$$\bar{U}_2 = \frac{1}{16}\sum_{\substack{n\neq s\\ n+p=2s}}\frac{\bar{e}_n\bar{e}_p}{(\Delta_{s,n}^-)^2}$$

$$\times[f_{s,n,p}^{(1)}(\bar{\eta})\cos(2\psi + 2\bar{\chi}/\beta_s) + 2(\bar{\chi}/\beta_s)\sin(2\psi)], \tag{3}$$

$$\bar{U}_3 = \frac{1}{8}\sum_{n\neq s}\frac{\bar{e}_n\bar{e}_p}{(\Delta_{s,n}^-)^2}$$

$$\times[f_{s,n,p}^{(2)}(\bar{\eta})\cos(2\psi + 2\bar{\chi}/\beta_s) + 2(\bar{\chi}/\beta_s)\sin(2\psi)],$$

$$h_n - h_p = 2h_s.$$

We use the following designations:

$$\bar{\mathbf{R}} = [\bar{\chi}, \bar{\eta}], \quad \bar{\chi} = 2\pi(z^{\text{slow}} - z_s^{\text{slow}})/\lambda,$$

$$\bar{\eta} = 2\pi r^{\text{slow}}/\lambda, \quad \tau = \omega t, \quad \Delta_{s,n}^{\pm} = (h_s \pm h_s)/h_s,$$

$$g_{s,n}(\bar{\eta}) = I_0^2\left(\frac{h_n}{h_s}\frac{\bar{\eta}}{\beta_s}\right) + I_1^2\left(\frac{h_n}{h_s}\frac{\bar{\eta}}{\beta_s}\right) - 1,$$

$$f_{s,n,p}^{(1)}(\bar{\eta}) = I_0\left(\frac{h_n}{h_s}\frac{\bar{\eta}}{\beta_s}\right)I_0\left(\frac{h_p}{h_s}\frac{\bar{\eta}}{\beta_s}\right) - I_1\left(\frac{h_n}{h_s}\frac{\bar{\eta}}{\beta_s}\right)I_1\left(\frac{h_p}{h_s}\frac{\bar{\eta}}{\beta_s}\right),$$

$$f_{s,n,p}^{(2)}(\bar{\eta}) = I_0\left(\frac{h_n}{h_s}\frac{\bar{\eta}}{\beta_s}\right)I_0\left(\frac{h_p}{h_s}\frac{\bar{\eta}}{\beta_s}\right) + I_1\left(\frac{h_n}{h_s}\frac{\bar{\eta}}{\beta_s}\right)I_1\left(\frac{h_p}{h_s}\frac{\bar{\eta}}{\beta_s}\right).$$

The variables ψ and $\beta_s = \omega/h_s c$ are the phase and the velocity of the synchronous particle, and the variables ($\bar{\chi}$ and $\bar{\eta}$) mean the transition to the coordinate system related to the particle.

The effective potential $\bar{U}_{\text{eff}}$ provides the full three-dimensional description of the particle dynamics in the smooth one-particle approximation. It is related to the Hamiltonian of the system as

$$\frac{1}{2}\left(\frac{d\bar{\mathbf{R}}}{d\tau}\right)^2 + \bar{U}_{\text{eff}} = \bar{H}. \tag{4}$$

Analysis of the effective potential makes it possible to set relations between the amplitudes of the harmonics at which the radial and phase stability of the beam is achieved, as well as to find restrictions imposed on the amplitudes.

GENERAL ANALYSIS OF THE EFFECTIVE POTENTIAL FUNCTION

Let us turn to expression (3) for $\bar{U}_{\text{eff}}$. The term $\bar{U}_0$ describes the interaction of the particle with the synchronous harmonic, which accelerates the beam and simultaneously defocuses it radially; that is, the extremum of $\bar{U}_0$ is a saddle point. The term $\bar{U}_1$ contributes only to the transverse motion, always focusing the beam in the radial direction. Its value depends neither on the amplitude nor on the phase of the synchronous wave. It will be shown that taking these two terms into account allows us to substantiate several types of APF mentioned above. With the condition $n + p = 2s$ (where s is also an integer) satisfied, the term $\bar{U}_2$ appears in the expression for $\bar{U}_{\text{eff}}$. This term affects both the longitudinal and the transverse motion of the particle and arises in the presence of two or more asynchronous field harmonics. Finally, if the condition $h_n - h_p = 2h_s$ ($n \neq s$) is met, the term $\bar{U}_3$ becomes nonzero. It also affects the radial and the longitudinal motion. Unlike $\bar{U}_2$, $\bar{U}_3$ cannot equal zero even in the two-wave approximation (one synchronous and one asynchronous waves). The extrema of $\bar{U}_2$ and $\bar{U}_3$, like that of $\bar{U}_0$, are also saddle points. Therefore, the necessary condition for the beam to be stable in both the longitudinal and transverse directions is the existence of the absolute minimum of $\bar{U}_{\text{eff}}$. In this case, $\bar{U}_{\text{eff}}$ is, in essence, a moving three-dimensional well at the bottom of which the synchronous particle is situated.

For generality, the synchronous velocity β_s is convenient to eliminate from all the expressions in (3). To do this, we pass to the new variables $\chi = \bar{\chi}/\beta_s$, $\eta = \bar{\eta}/\beta_s$, $e_n = \bar{e}_n/\beta_s$, and $\mathbf{R} = \bar{\mathbf{R}}/\beta_s$. With the new variables, the three-dimensional well $U_{\text{eff}} = \bar{U}_{\text{eff}}/\beta_s^2$ retains its shape.

We begin our analysis with the simplest case and expand U_{eff} in the vicinity of its minimum $\{\chi = 0, \eta = 0\}$:

$$U_{\text{eff}}(\chi, \eta) = U(0, 0) + \omega_\chi^2\chi^2/2 + \omega_\eta^2\eta^2/2 + \delta\chi\eta^2/2 + \gamma\chi^3/3 + \ldots. \tag{5}$$

The expansion coefficients here depend on the amplitudes e_n of the harmonics ($n = 0, 1, \ldots$). Obviously, radial and longitudinal phase focusing will be provided when the amplitudes are such that

$$\omega_\chi^2 > 0, \quad \omega_\eta^2 > 0. \tag{6}$$

Note that parametric coupled resonances, which disturb the stability of the beam, may occur for certain relationships between ω_χ and ω_η. To study these resonance effects, it is necessary to take into consideration the last two terms in (5).

In the general case, the analysis of the effective potential as a function of the harmonic composition of the field allows us to study the particle behavior not only near the minimum but also throughout the ranges of the coordinates χ and η and velocities. The first important restriction on the amplitudes of the spatial field harmonics can be derived from the condition of nonoverlapping resonances for various waves when the phase portrait of the beam–wave dynamic system is considered in the (β_χ, χ) plane. On the one hand, this condition specifies the applicability limits of the averaging method; on the other hand, it prevents possible disturbances of the longitudinal stability of the beam. Using Hamiltonian (4) and analyzing the shape of the four-dimensional phase space, one can easily find a relationship between a given longitudinal channel acceptance and the limit value of the transverse emittance in order to obtain the maximal current transmission coefficient. To illustrate the aforesaid, we will consider specific accelerating and focusing channels with different forms of U_{eff}.

IMPLEMENTATIONS OF RF-FOCUSING ACCELERATORS

The amplitudes of the spatial harmonics at the system axis are represented as $e_n = \alpha_n e_{\max}$ ($\alpha_n \le 1$). The value of $e_{\max}$ related to the breakdown amplitude of the accelerating field is usually defined by the cavity design and will be considered fixed. Thus, our goal is to find the weighting coefficients $\{\alpha_n\}$ for various focusings.

1. One Synchronous and One Asynchronous Field Harmonics

This system is the simplest. The addition of U_1, which provides radial focusing, to the saddle-like term U_0 yields, under certain conditions, a three-dimensional potential well. In this case, the weighting coefficients are $\{\alpha_s = \alpha, \alpha_n = 1\}$.

First we consider the field with the phase advance $\mu = 0$ and $s \ge 1$. The set $\{s = 1, n = 0\}$ corresponds to the usual Alvarez structure. The effect of focusing here is absent, since $U_1 = \text{const}$. The set $\{s = 2, n = 0\}$ corresponds to the same structure but with a doubled period. Of the systems with $\{s = 1, n = 2\}$ and $\{s = 2, n = 1\}$, which do not have the zero harmonic, the former can provide efficient focusing.

The implementation of ARF is much easier for the structures with the phase advance $\mu = \pi$ and $s \ge 0$. The sets $\{s = 0, n = 1\}$ and $\{s = 1, n = 0\}$ are Wideröe structures, where tubes of different lengths and inner diameters alternate. In the case $\{s = 0, n = 1\}$, U_3 appears but its effect is weak. This set of harmonics, as well as $\{s = 1, n = 2\}$, provides effective transverse focusing provided that the amplitudes are properly selected.

The condition for transverse focusing can be expressed in the form

$$\alpha \sin\psi < \frac{3}{8} e_{\max}\left(\frac{1}{(\Delta^-_{s,n})^2} + \frac{1}{(\Delta^+_{s,n})^2}\right)\left(\frac{h_n}{h_s}\right)^2 + \frac{3}{8} e_{\max}\frac{\alpha}{4}. \quad (7)$$

The energy increases with an acceleration gradient

$$\frac{dW_s}{dz} = \frac{1}{2}\alpha E_{\max}\cos\psi, \quad (8)$$

which is proportional to α; therefore, this parameter is bounded from below. In the simplest case of APF [2], the phase velocity of the focusing wave is lower than that of the accelerating velocity ($n > s$) and the amplitudes of the harmonics decrease with increasing harmonic number ($e_s > e_n$). Therefore, condition (7) can be satisfied only if the synchronous phase is small ($\sin\psi \ll 1$). This type of RF focusing has a small longitudinal acceptance and a high acceleration gradient.

In one of the cases considered, it is also assumed that $n > s$. Here, however, condition (7) can be met even at a large capture efficiency ($\sin\psi \approx 1$) if the accelerating harmonic amplitude e_s is smaller than the amplitude of the focusing harmonic e_n; that is, if $\alpha \ll 1$. This version of focusing can be implemented by selecting the structure period in such a way that it covers two or more accelerating gaps. From condition (7), it also follows that all versions with $n > s$ are inefficient for transverse focusing.

Another important restriction on the amplitudes follows from the condition that the resonances (separatrices) of adjacent waves do not overlap. The swing of the separatrices over longitudinal velocities grows with the harmonic amplitudes. When the amplitudes reach some values, the separatrices overlap. This disturbs the condition for the phase stability of the beam.

Except for the case $\{\mu = 0, s = 0, n = 1\}$, U_{eff} can be represented as

$$U_{\text{eff}} = -A[F_1(\eta)\sin(\psi + \chi) - \chi\cos\Psi)] + F_2(\eta). \quad (9)$$

The initial values of the Hamiltonian and U_{eff} strongly depends on the channel aperture and the synchronous velocity. Figure 1 shows sections obtained when U_{eff} is cut by the planes $\eta = 0$ and $\chi = 0$. To provide efficient particle capture by phase and efficient radial focusing, it is desirable that the cut depths in Fig. 1 be equal. Mathematically, this requirement is written as

$$A[\sin\psi - 2\psi\cos\psi] = -AF_1(a)\sin\psi + F_2(a), \quad (10)$$

where $a = 2\pi R/\lambda\beta_s$ is a dimensionless aperture.

For low-energy (100–300 keV) proton accelerators, requirement (10) is difficult to satisfy over the entire channel length. As follows from the results of numeri-

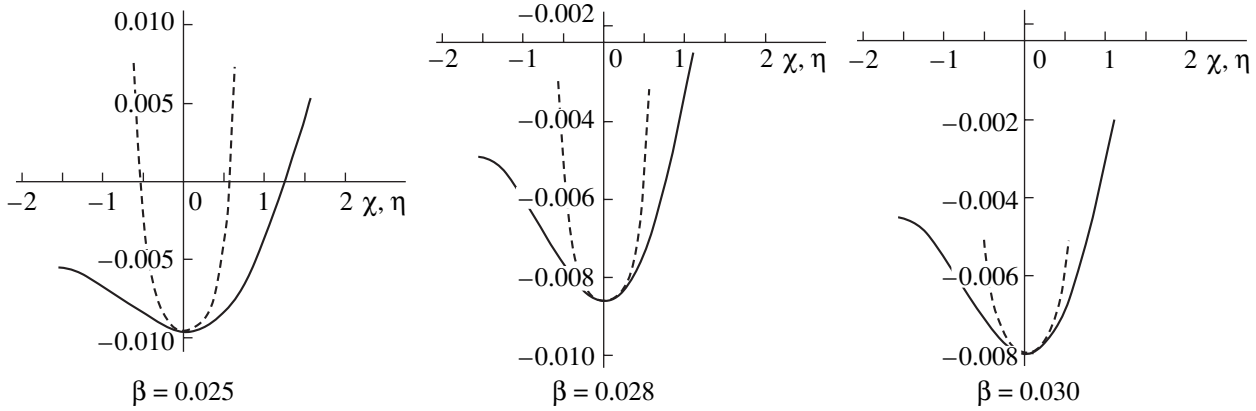


Fig. 1. Sections obtained by cutting U_{eff} by the planes $\eta = 0$ (continuous curve) and $\chi = 0$ (dashed curve) at different velocities of the synchronous particle. $\psi = 0.25\pi$.

cal simulation of the beam dynamics (see below), this is a reason for inefficient particle capture by the phase at the beginning of acceleration.

In the one-particle approximation, the choice of the RF structure and the optimization of the field composition are performed as follows. Phase advances μ, the synchronous phase ψ, the maximal field $e_{\max}$, and the parameter α satisfying condition (7) are selected. In this way, the longitudinal motion of the particles, including the longitudinal acceptance, are completely defined in the smooth approximation. The parameters α and ψ are selected such that the rate of acceleration and the longitudinal acceptance have reasonable values. The parameters $e_{\max}$ and μ depend on the design of the accelerator. Next, the transverse oscillation frequencies ω_η are calculated for various s and n. Systems with large values of ω_η are the best. It should be noted that, first, the field amplitude at which the separatrices overlap rapidly drops with increasing harmonic number. Second, the implementation of a field with high harmonic numbers ($n > 2$) at small β is difficult because of the need for complicating the structure period: a short period D must have two or more accelerating gaps. For example, the system with $\{s = 2, n = 3\}$, which has the maximum ω_η for any μ, cannot be implemented if the beam energy is small.

Table

Structure type	Composition of RF field	U_{eff}
$\mu = 0$, $s \geq 1$	$s = 1$, $n = 2$, $p = 0$	$U_0 + U_1 + U_2 + U_3$
	$s = 2$, $n = 4$, $p = 0$	$U_0 + U_1 + U_2 + U_3$
	$s = 2$, $n = 1$, $p = 3$	$U_0 + U_1 + U_2$
$\mu = \pi$, $s \geq 0$	$s = 0$, $n = 2$, $p = 1$	$U_0 + U_1 + U_3$
	$s = 1$, $n = 3$, $p = 0$	$U_0 + U_1 + U_3$
	$s = 1$, $n = 0$, $p = 2$	$U_0 + U_1 + U_2$
	$s = 2$, $n = 1$, $p = 3$	$U_0 + U_1 + U_2$

2. The Effect of the Second Asynchronous Harmonic

When implementing optimized accelerating resonators (even with a simple period structure), one must take into account the contribution of higher spatial harmonics whose amplitude usually rapidly drops as the harmonic number grows. The analysis of the U_{eff} behavior allows the elucidation of the effect of these harmonics on the particle dynamics. Consider the effect of the second nonsynchronous harmonic in the above-considered system with $\alpha \ll 1$. In this case, the set of the weighting coefficients $\{\alpha_s = \alpha, \alpha_n = 1)$ is supplemented by $\alpha_p \equiv \varepsilon < 1$. Two cases are possible depending on the harmonic number and the type of the structure. The first one is $U_{2,3} = 0$. This means that the second nonsynchronous harmonic adds only to U_1, i.e., enhances focusing, without affecting the longitudinal motion, the addition to the focusing power being proportional to ε^2. In the other case, the combination of the asynchronous waves generates the cross terms $U_{2,3}$ (see table) and the phase portrait of the dynamic system becomes qualitatively different. Figure 2 shows the section obtained by cutting the effective potential by the plane $\eta = 0$ at different ε. At $\varepsilon \longrightarrow 0$, U_{eff} is given by (9). As ε increases, the longitudinal capture width decreases and the rate of the acceleration grows in proportion with ε. The structure of the period D depends on the amplitude of the second harmonic. At some ε, one more minimum of U_{eff} appears, hence, the second bunch in the period. The general view of the effective potential and the map of its isolevels for this case are depicted in Fig. 3. The phase portrait upon forming the second region of stable motion is demonstrated in Fig. 4. Special investigations showed that the accelerating structure with the second large-amplitude ($\varepsilon \approx 1$) asynchronous wave is not efficient for obtaining the maximal current transmission coefficient. However, the approach suggested allows one to study corrections to the particle motion. These corrections are necessary for the analysis of the beam

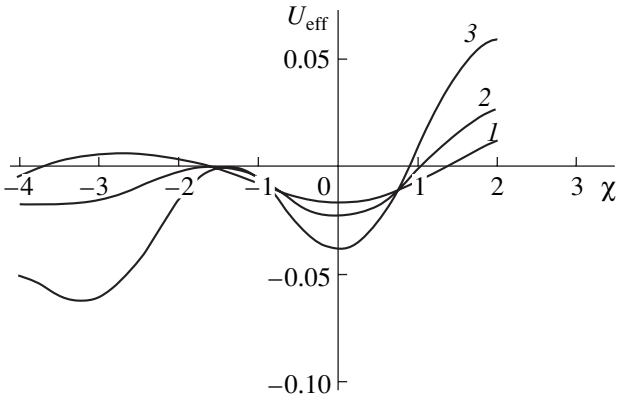


Fig. 2. Sections obtained by cutting U_{eff} by the plane $\eta = 0$ for $\varepsilon =$ (*1*) 0.05, (*2*) 0.3, and (*3*) 0.7.

dynamics. For example, one must take into consideration how the nonsynchronous harmonics affect the longitudinal motion equation, the period of the RF structure, and, hence, the velocity of the synchronous particle. In this respect, the APF modification suggested in [3] (AAPF) can be viewed as an attempt to improve the longitudinal channel acceptance through the effect of asynchronous waves on the phase motion of the beam.

3. Accelerator without Asynchronous Harmonic

Expression (3) implies that the particle can be accelerated and, at the same time, radial stability can be provided without the synchronous harmonic as well ($\alpha_s = 0$). The three-dimensional potential well can be

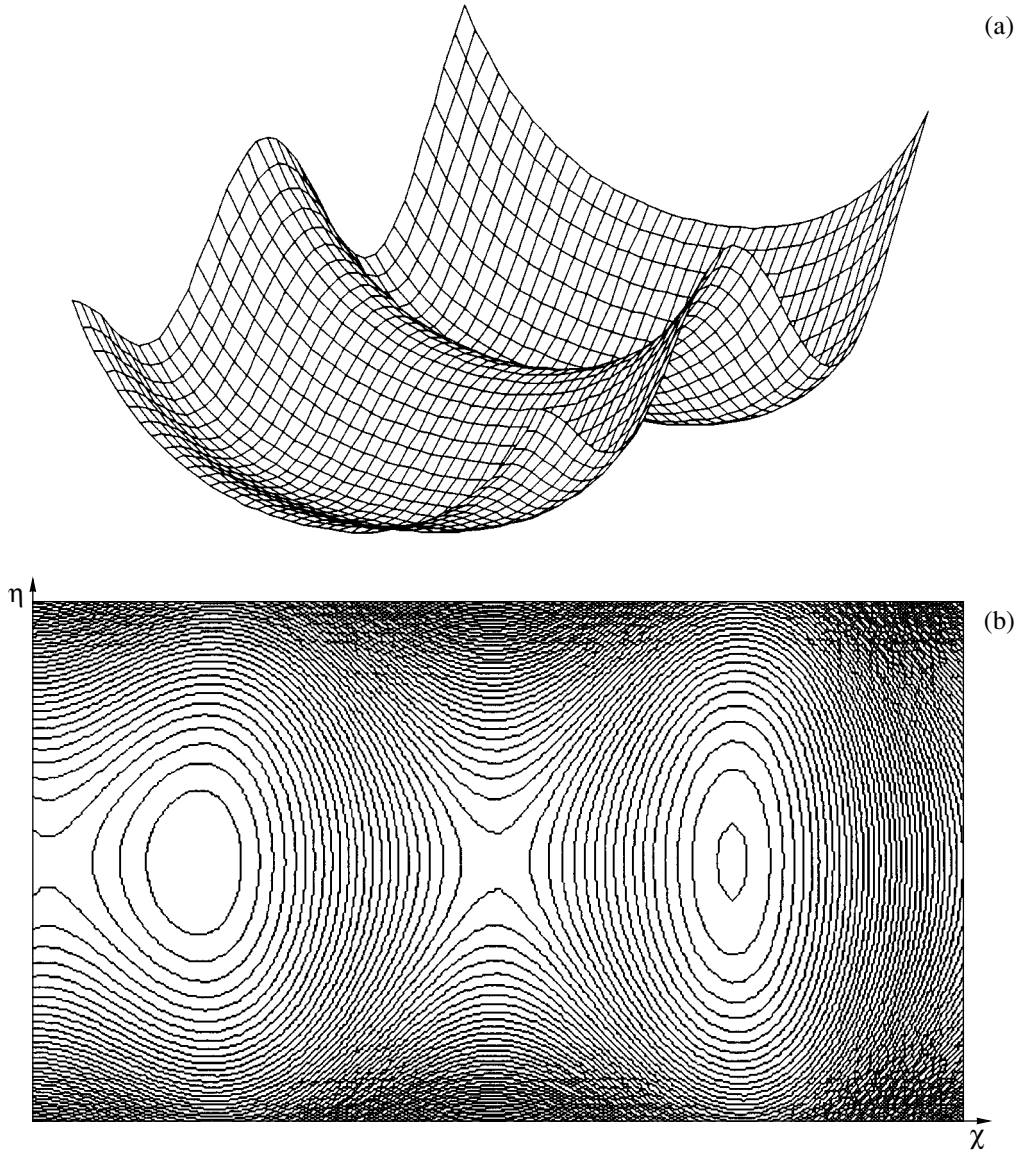


Fig. 3. Effective potential function with two minima: (a) general view and (b) isolevel map.

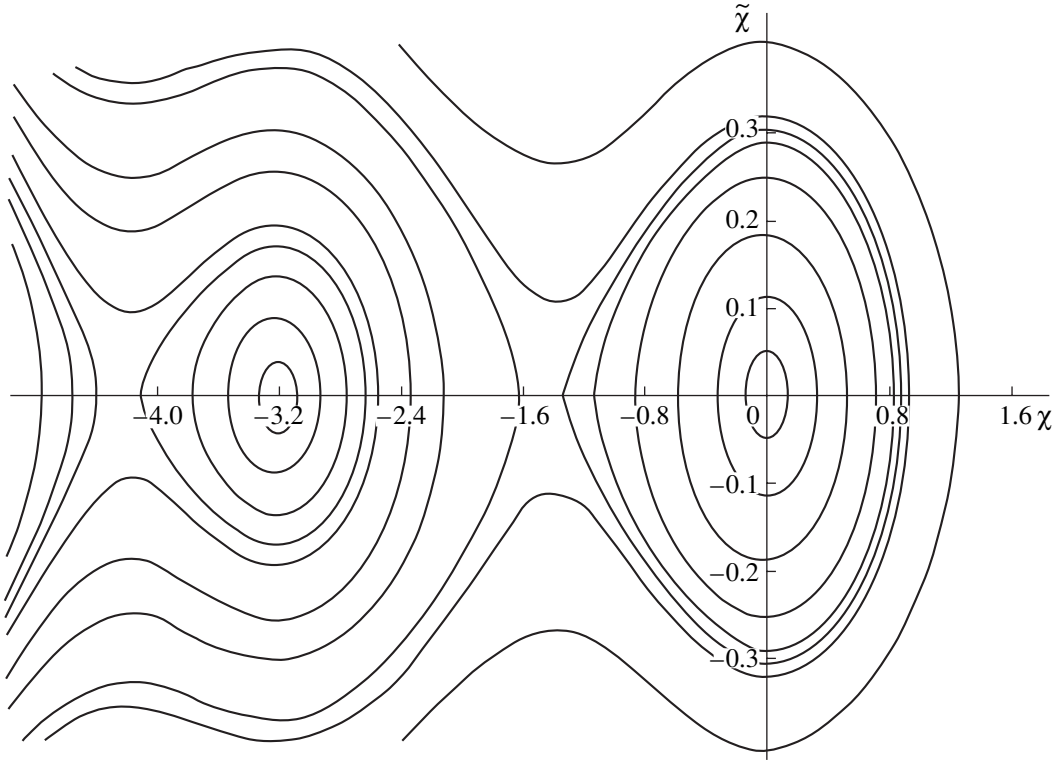


Fig. 4. Phase portrait of the system with $\{\mu = \pi, s = 1, n = 2, p = 0\}$ upon the formation of the second bunch.

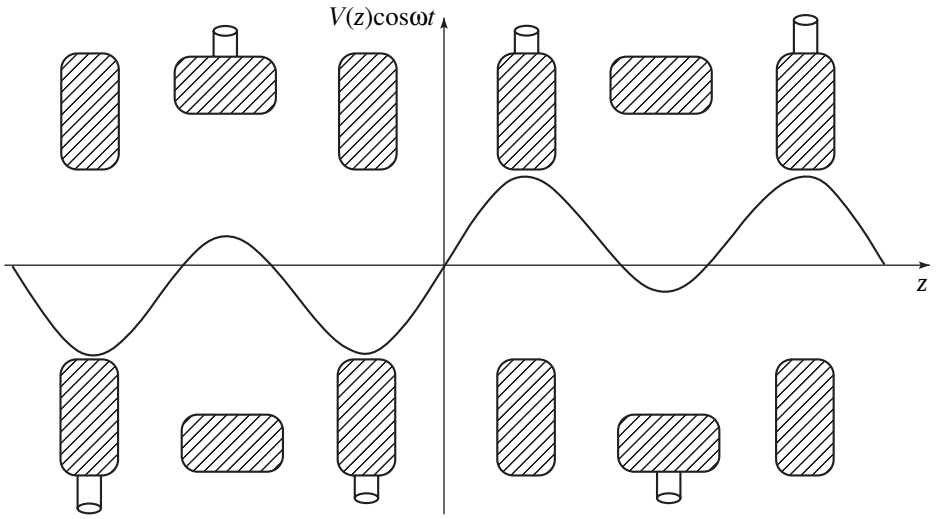


Fig. 5. Implementation of the structure.

formed by the terms $U_{2,3}$. In this case, acceleration takes place in the field of the composite wave arising from the combined action of two nonsynchronous harmonics (the undulator mechanism of acceleration [8]). This system has a number of intriguing features. First, modulation takes place at a doubled frequency. Second, the particle charge enters into expressions (2) and (3) in the quadratic form, which provides the chance to accelerate unlike ions in one bunch. In this case, the space charge can be neutralized and the beam current can markedly be increased.

Consider, as an example, the case $U_2 \neq 0$ and $U_3 = 0$ that is, the structures with $\{\mu = 0, s = 2, n = 1, p = 3\}$ and $\{\mu = \pi, s = 1, n = 0, p = 2\}$. The rate of the acceleration is given by

$$\frac{dW}{dz} = \frac{1}{4}\alpha_n\alpha_p\frac{e_{\max}E_{\max}}{(\Delta^-_{s,n})^2}\sin(2\psi). \qquad (11)$$

With the substitution $\hat{\chi} = 2\chi$ and $\hat{\psi} = 2\psi + \pi/2$, the effective potential takes the form of (9). Thus, all conclusions drawn for the Hamiltonian of the system with one synchronous and one asynchronous wave remain valid.

SELECTION OF ACCELERATING CHANNEL PARAMETERS

When intense beams are accelerated, the accelerating channel parameters must be selected with regard to the space charge field. Consider the structure with $\{\mu = \pi, s = 0, n = 1\}$ in the two-wave approximation. We subdivide the accelerating channel into the grouping and main (accelerating) sections. In the former, the synchronous wave linearly decreases from $\psi = \pi/2$ to some rated value and the amplitude of the RF field monotonically grows. In the accelerating section, these parameters remain fixed. Such a division allows one to substantially improve the phase capture. In the grouping section, the amplitude of the RF field as a function of the longitudinal coordinate must be defined in such a way that the coefficient of current transmission is as high as possible. Above, the amplitudes of the RF field harmonics have been defined as $E_n(z) = \alpha_n E_{\max}(z)$ ($\alpha_n \le 1$). In searching for the optimal function $E_{\max}(z)$ for the two-wave approximation, we start with the one-dimensional model. Let us represent the Hamiltonian of the system at the axis ($\eta = 0$) as

$$H_1 = \frac{P_\varphi^2}{2m_\varphi} + V_{\text{ext}}(\varphi) + V_c. \tag{12}$$

Here, $\{P_\varphi, \varphi\}$ are the canonically conjugate variables: the momentum $P_\varphi = W - W_s$, the coordinate

$$\varphi = \frac{2\pi}{\lambda}\int dz^{\text{slow}}(1/\beta_s - 1/\beta), \quad m_\varphi = mc^2\beta_s^2/\omega,$$

$$V_{\text{ext}} = \frac{2\pi mc^3\beta_s}{\lambda}\bar{U}_0(\varphi, \eta = 0),$$

and V_c is the part of the potential function that is related to the space charge. The distribution function can be taken in the form

$$f = f_0\sqrt{H_0 - H_1} \quad \text{for} \quad H_1 \le H_0, \qquad f = 0 \quad \text{for} \quad H_1 > H_0. \tag{13}$$

In this case, the indifferent equilibrium state is established at some value I_n of the current (we call I_n the limit current of the system). This parameter is linearly related to the longitudinal coordinate:

$$I_n(z) \propto \beta_c^2(z)E_{\max}(z)\Phi(\psi), \tag{14}$$

where $\Phi = \Delta(2\psi\cos\psi - \sin\psi) + \cos(\Delta - \psi) - \cos\psi + (\Delta^2/2 - 2\Delta\psi)\cos\psi$ and Δ is the phase width of the separatrix.

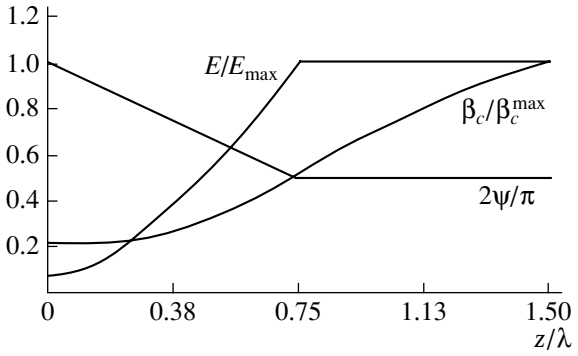


Fig. 6. Accelerator parameters vs. longitudinal coordinate.

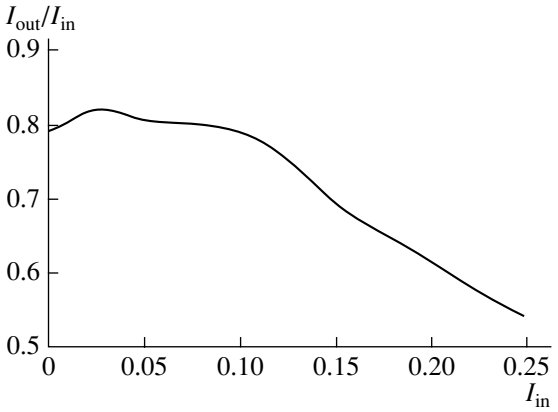


Fig. 7. Current transmission coefficient vs. input current.

It can be shown that the limit current I_n is roughly proportional to the longitudinal acceptance. For the current transmission coefficient to be high, the longitudinal acceptance must be a nondecreasing function of the longitudinal coordinate. From this condition, we can find the function $E_{\max}(z)$. Assume that, in the grouping section,

$$I_n(z) = I_n(0)F(z), \tag{15}$$

where $I_n(0)$ is the initial value of the limit current and $F(z)$ is some nondecreasing function of the longitudinal coordinate.

Relationships (2), (14), and (15) allow us to relate the accelerating channel parameters $E_{\max}(z)$, $\beta_c(z)$, and $\psi(z)$. This relation involves the unknown function $F(z)$. It must rapidly increase and, at the same time, be bounded from above. The function $E_{\max}(z)$ thus found can be refined by numerical optimization.

IMPLEMENTATION OF THE STRUCTURE

It is known that the amplitudes of spatial harmonics in resonant Wideröe structures rapidly drop with increasing harmonic number. As has been shown

above, RF focusing will be efficient and the beam will be accelerated if the amplitude of the first harmonic ($n = 1$) is much higher than that of the zero harmonic ($s = 0$). Such a situation is provided in the modified Wideröe structure, which has a complex period consisting of drift tubes of various lengths and inner diameters. Figure 5 demonstrates the implementation of such a structure and the potential distribution along the axis. The potentials applied to the adjacent electrodes are equal in amplitude and are phase-shifted by π. The field potential as a function of z (along the axis) has two equiamplitude antinodes within the period that are symmetric relative to the third one. Such a function was realized by properly selecting the inner diameter, length, and shape of the tubes. For example, the amplitude of the negative antinode is smaller than the amplitudes of the other two, since the inner diameter of the associated (middle) tube is larger. By properly selecting the lengths of the tubes and the shapes of the roundings, one can almost completely eliminate higher harmonics in the potential distribution. The thus-obtained transverse distribution of the RF field is consistent with its representation in terms of the modified Bessel functions. The length of the period is $\lambda\beta_c/2$, which makes the implementation of the structure with $n \geq 2$ difficult, since five or more tubes should be arranged within the period in this case. When heavy ions are accelerated, the synchronous velocity drastically drops and the period shrinks. This difficulty can be obviated by operating at lower frequencies, i.e., by increasing λ.

NUMERICAL SIMULATION

The dynamics of a high-current proton beam in the ARF accelerator described above was numerically simulated with the method of coarse particles using a specially developed computer program. The harmonic composition of the RF field was selected for the case $\{\mu = \pi, s = 0, n = 1\}$ by the above-mentioned technique. The accelerating channel parameters as functions of the longitudinal coordinate are represented in Fig. 6. The amplitude $E_{\max}(z)$ of the field in the grouping region was found with the above approach. This function can be refined by using special methods of numerical optimization. In the calculation, the parameters of the accelerator were the following: the operating frequency was 150 MHz, $\alpha_s = 0.1$, $\alpha_n = 1$, the maximal field amplitude was 300 kV/cm, the input/output energy was 0.1/1.22 MeV, the input/output beam current was 0.1/0.079 A, the current transmission coefficient was 0.79, the full channel length 3 m, the rate of the acceleration was 0.7 MeV/m, and the aperture radius was 0.6 cm. At the rate of the acceleration of 0.7 MeV/m, the current transmission coefficient is 0.79, which is a breakthrough for ARF accelerators. The current transmission coefficient vs. input current is plotted in Fig. 7. The averaging method was verified by numerical simulation in both the total and the averaged fields. The results obtained in both cases coincided up to 5–10%.

CONCLUSION

A new approach to describing RF focusing in the axisymmetric field of an ion linac is developed. The classification of RF focusing types based on the analysis of the harmonic composition of the RF field is suggested. Analyzing ARF, we selected the accelerating channel parameters for the Wideröe structures and obtained the high current transmission coefficient. Computer simulation of high-current ion beam dynamics in an ARF structure was performed. The validity of the averaging method was justified. In some cases, the performance of the accelerator under study is close to that of RFQ devices, which proves the ARF efficiency.

REFERENCES

1. M. L. Good, Phys. Rev. **92**, 538 (1953).
2. I. B. Faĭnberg, Zh. Tekh. Fiz. **29**, 568 (1959) [Sov. Phys. Tech. Phys. **4**, 506 (1959)].
3. V. V. Kushin, At. Énerg. **29** (3), 123 (1970).
4. V. S. Tkalich, Zh. Éksp. Teor. Fiz. **32**, 625 (1957) [Sov. Phys. JETP **5**, 518 (1957)].
5. V. K. Baev and S. A. Minaev, Zh. Tekh. Fiz. **51**, 2310 (1981) [Sov. Phys. Tech. Phys. **26**, 1360 (1981)].
6. V. D. Danilov and A. A. Il'in, in *Theoretical and Experimental Investigations of Charged-Particle Accelerations* (Énergoatomizdat, Moscow, 1985), pp. 93–97.
7. H. Okamoto, Nucl. Instrum. Methods Phys. Res. A **284**, 233 (1989).
8. É. S. Masunov, Zh. Tekh. Fiz. **60** (8), 152 (1990) [Sov. Phys. Tech. Phys. **35**, 962 (1990)].
9. E. S. Masunov, in *Undulator and RF-System for Ion Linear Accelerators: Proceedings of the 18th International Linac Conference* (CERN, Geneva, 1996), Vol. 2, p. 487.
10. E. S. Masunov and N. E. Vinogradov, in *Proceedings of the 1999 Particle Accelerator Conference, New York, 1999*, Vol. 4, p. 2855.

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1151–1160. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 88–96.

ELECTRON AND ION BEAMS, ACCELERATORS

A Planar Relativistic Electron Beam in a Plasma Channel Bounded by Conducting Walls

A. P. Kuryshev and V. D. Andreev

Baltic State Technical University, St. Petersburg, 198005 Russia

Received October 5, 2000

Abstract—The force interaction of a relativistic electron beam with a plasma in a channel bounded by plane-geometry highly conducting walls is studied. The steady-state interaction regime, $\omega = ku$, is analyzed using the model of a cold collisional electron plasma. The formulas for the transverse component of the force acting on the beam electrons are derived for an arbitrary deviation of the beam from the symmetry plane of the channel.

INTRODUCTION

The problems associated with the transport of relativistic electron beams (REBs) in plasma channels bounded by media with different conductivities (from vacuum to perfect conductors) have been widely discussed in the literature [1–10].

It is well known that the force interaction of an REB with a plasma in a channel bounded by a conducting wall induces positive charges and reverse currents on the wall surfaces; moreover, the force between the charges and the REB is attractive, whereas the force between the currents and the REB is repulsive. Clearly, the dominance of either an attractive or a repulsive force results in the preferential attraction or repulsion between the REB and the wall. If an REB is transported in a vacuum channel, then the induced charge dominates and the beam is attracted by the channel wall. In a plasma-filled transport channel, the beam charge is partially or wholly screened. As a result, the interaction between the currents may give rise to a force that acts to stabilize the deviation of the beam from the channel axis. This effect can be used to solve such problems as transverse stabilization of the trajectory of a beam transported through the plasma [2–6] and the formation of annular beams [7, 8]. The attraction of an REB to the wall of the channel filled with a weakly ionized plasma was analyzed by Vladyko and Dudyak [9], who applied the model in which the plasma properties are described in terms of a constant electrical conductivity σ. In our earlier paper [10], we solved the problem of the transport of a relatively thin annular electron beam in a cylindrical plasma channel.

A similar problem is treated in the present paper: we investigate the force interaction between an REB and a plasma in a channel bounded by plane-geometry highly conducting walls. We examine the steady-state interaction regime using the model of a cold collisional electron plasma and the model of a finite-duration beam with constant parameters.

The transverse component of the Lorentz force that acts on the beam electrons moving with velocity u along the z axis has the form

$$F = e(E_x - \beta B_y), \quad \beta = \frac{u}{c}. \tag{1}$$

The problem as formulated reduces to solving the corresponding Maxwell equations under the assumption that the linear properties of the medium are prescribed. In the linear model of a cold collisional electron plasma, the dielectric tensor is described by the familiar expression

$$\varepsilon_{ij}(\omega, \mathbf{k}) = \delta_{ij}\varepsilon(\omega) = \delta_{ij}\left(1 - \frac{\omega_p^2}{\omega(\omega + i\nu)}\right), \tag{2}$$

where ω_p is the Langmuir frequency of the plasma electrons and ν is the collision frequency of the plasma electrons.

We use Maxwell's equations

$$\begin{aligned} \mathrm{curl}\mathbf{B} &= \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c}\mathbf{j}_b, \quad \mathrm{div}\mathbf{B} = 0, \\ \mathrm{curl}\mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \quad \mathrm{div}\mathbf{D} = 4\pi\rho_b \end{aligned} \tag{3}$$

and the equation of state in which spatial dispersion is neglected:

$$D_i(t, \mathbf{r}) = \int_{-\infty}^{t} dt' \varepsilon_{ij}(t - t', \mathbf{r}) E_j(t', \mathbf{r}). \tag{4}$$

Taking into account the equation of state (4), we consider the projections of Eqs. (3) onto the axes of the Cartesian coordinate system. We take the Fourier–Laplace transformation (in the z coordinate and time t) of the electromagnetic field components, which deter-

mine the transverse component of the Lorentz force acting on the beam electrons, to obtain

$$F(\omega, k, x) = -e\frac{ik}{\kappa^2}(1-\beta^2\varepsilon)\frac{\partial E_z}{\partial x},$$
$$\kappa^2 = k^2 - \frac{\omega^2}{c^2}\varepsilon(\omega). \quad (5)$$

In order to analyze the steady-state interaction regime, we set $\omega = ku$ in expressions (5). As a result, the force component in question becomes

$$F = F(k, x) = -e\frac{i}{k}\frac{\partial E_z}{\partial x}. \quad (6)$$

We start by investigating the case of a perfectly conducting channel wall.

PERFECTLY CONDUCTING WALL

The geometry of the problem is shown in Fig. 1. The equations for the electric field component E_z in each of the regions of the system under investigation can readily be obtained from Maxwell's equations:

$$\mathcal{L}E_z \equiv \frac{\partial^2 E_z}{\partial x^2} - \kappa^2 E_z = H[\eta(x-a+h) - \eta(x-a-h)], \quad (7)$$
$$H = 4\pi\kappa^2 e n_b \frac{i}{k\varepsilon(ku)} j_z(k, x),$$

where $\eta(x)$ is the Heaviside step function, $en_b u j_z(k, x)$ is the Fourier-transformed (with respect to the z coordinate) beam current density, and $2R$ is the width of the plasma channel.

For definiteness, we assume that the x component of the beam current density is uniform. The boundary conditions for Eq. (7) have the form

$$E_z(x = \pm R) = 0, \quad \{E_z\}_{x=a\pm h},$$
$$\{B_y\}_{x=a\pm h} = 0 \Leftrightarrow \left\{\frac{\partial E_z}{\partial x}\right\}_{x=a\pm h} = 0, \quad (8)$$

$$\{f\} = f(x = a \pm h + 0) - f(x = a \pm h - 0). \quad (9)$$

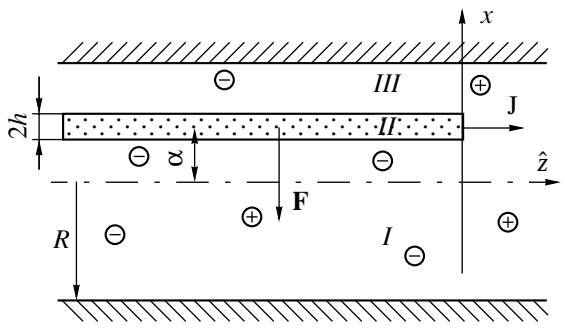


Fig. 1.

The solution to problem (7)–(9) is easy to construct:

$$E_z^{\mathrm{I}} = -\frac{2H}{\kappa^2}\frac{\sinh(\kappa h)\sinh(\kappa(R-a))}{\sinh(2\kappa R)}\sinh(\kappa(R+x)),$$

$$E_z^{\mathrm{II}} = \frac{H}{\kappa^2}\left[\frac{\cosh(\kappa a)\cosh(\kappa(R-h))}{\cosh(\kappa R)}\cosh(\kappa x) - \frac{\sinh(\kappa a)\sinh(\kappa(R-h))}{\sinh(\kappa R)}\sinh(\kappa x) - 1\right],$$

$$E_z^{\mathrm{III}} = -\frac{2H}{\kappa^2}\frac{\sinh(\kappa h)\sinh(\kappa(R+a))}{\sinh(2\kappa R)}\sinh(\kappa(R-x)),$$

$$\mathrm{Re}\,\kappa > 0.$$

We take the derivative of E_z with respect to x to obtain the desired expression for the Fourier component of the sought-for force acting on the beam electrons that occur at the symmetry plane of the channel

$$F(x = a) = -\frac{ei}{k}\frac{\partial E_z}{\partial x}\bigg|_{x=a}$$
$$= -\frac{4\pi e^2 n_b}{R}\frac{i}{k\varepsilon(ku)}\frac{iR\kappa\sinh(2\kappa a)\sinh(h\kappa)}{k\sinh(2\kappa R)}\tilde{j}(k). \quad (10)$$

In order to determine this force over the entire physical space, it is necessary to take the inverse Fourier transformation of expression (10) in the longitudinal coordinate $\hat{z}$.

In a collisionless plasma ($\nu = 0$), the desired expression for the force acting on the electrons of an ultrarelativistic ($\gamma \longrightarrow \infty$) beam has the form

$$F = -2\pi e^2 n_b \lambda \frac{\sinh\frac{2a}{\lambda}\sinh\frac{h}{\lambda}}{\sinh\frac{2R}{\lambda}}\left[1 - \cos\frac{\hat{z}}{\lambda}\right].$$

The finiteness of the relativistic factor and the collisions in the plasma can be incorporated in the same way as was done in [10]. As a result, expression (10) can be represented in the form

$$F(k) = -\frac{4\pi e^2 n_b}{R}G_1(k)G_2(k)j(k),\ G_1(k) = \frac{i}{k\varepsilon(ku)},$$
$$G_2 = i\frac{R\kappa\sinh(2\kappa a)\sinh(h\kappa)}{k\sinh(2\kappa R)}. \quad (11)$$

In order to extend this expression to the entire physical space, we first take the inverse Fourier transformation of the functions G_1 and G_2 and then successively find the two Fourier convolutions that involve the resulting functions. For $\nu < 2\omega_p$, the inverse Fourier

transform of the function $G_1(k)$ is determined by the contributions of the two poles

$$k_{1,2} = \pm k_0 - i\alpha \quad \text{and} \quad k_0 = \sqrt{\frac{1}{\lambda^2\beta^2} - \alpha^2}$$

$$\left(\alpha = \frac{\nu}{2u}, \quad \lambda = \frac{c}{\omega_p}\right)$$

that lie in the lower half-plane of the complex variable k and describe the excitation of the electron Langmuir oscillations (Fig. 2):

$$G_1(k) - G_1(\hat{z}) = \frac{1}{2\pi}\int G_1(k)\exp(ik\hat{z})dk$$

$$= e^{\alpha\hat{z}}\left[\cos k_0\hat{z} - \frac{\alpha}{k_0}\sin k_0\hat{z}\right]\eta(-\hat{z}).$$

The inverse Fourier transform of the function $G_2(k)$ is determined by the contributions of the zeros of the function

$$\sinh 2\kappa R = 0 \Leftrightarrow e^{4\kappa R} = 1 \Leftrightarrow R\kappa_n$$
$$= ij_n \Leftrightarrow (R\kappa_n)^2 = -j_n^2, \qquad (12)$$
$$j_n = \frac{\pi n}{2} \ (n = 1, 2, \ldots).$$

Equation (12) describes the spectrum of our problem and differs from the corresponding equation for an REB transported in a cylindrical channel [10] only in the numbers j_n. In our analysis, we can use the solution to Eq. (12) that was constructed in [10]. Equation (12), which can be reduced to a cubic equation, has three series of poles: one collisional series (the ν-series) $k_n^{(1)} = -iy_n^{(1)}$ and two relativistic series (the γ-series) $k_n^{(2,3)} = -iy_n^{(2,3)}$ (Fig. 2). The poles of the collisional series lie within the interval $(0, -i\nu/u)$ of the negative imaginary axis and are described by the formula

$$y_n^{(1)} = \frac{\nu}{u}\frac{1}{1+d_n^2}\left[1 - \frac{\lambda^2}{\gamma^2}\left(\frac{\nu}{u}\right)^2\frac{d_n^4}{(1+d_n^2)^3}\right], \quad d_n = \frac{R}{\lambda}\frac{1}{j_n}. \qquad (13)$$

The first relativistic series $y_n^{(2)}$ lies within the interval $(-i\gamma/\lambda, -i\infty)$ of the negative imaginary axis, and the second relativistic series $y_n^{(3)}$ lies within the interval $(i\gamma/\lambda, i\infty)$ of the positive imaginary axis. The relativistic series are determined by the formulas

$$y_n^{(2,3)} = \pm\frac{\gamma}{\lambda}\frac{\sqrt{1+d_n^2}}{d_n}\left[1 - \frac{1}{2}\frac{\nu}{u}\frac{\lambda}{\gamma}\frac{d_n(d_n^2+2)}{(1+d_n^2)^{3/2}}\right]. \qquad (14)$$

Formulas (13) and (14) imply that, in the collisionless limit, there is no ν series, while the ultrarelativistic limit ($\gamma \longrightarrow \infty$) is free of both of the γ series. In the case $\nu = 0$ and $\gamma \longrightarrow \infty$, we arrive at the above solution.

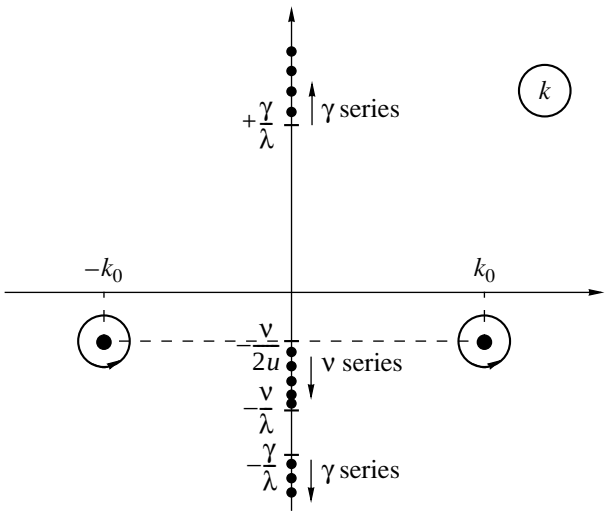


Fig. 2.

The contributions to the inverse Fourier transform $G_2(\hat{z})$ in the beam region ($\hat{z} < 0$) are determined by the collisional ν series and the first γ series, while the second γ series gives the corresponding contribution in the region ahead of the beam ($\hat{z} > 0$). It is easy to see that the singular point $k = 0$ is a removable singularity and does not contribute to the function $G_2(\hat{z})$, for which we successively obtain

$$G_2(\hat{z}) = \sum_{l=1}^{3}\sum_{n=1}^{\infty} C_{nl}e^{y_n^{(l)}\hat{z}} f_l(\hat{z}),$$

$$C_{nl} = \frac{(-1)^n j_n^2}{R^2 y_n^{(l)}}\frac{\sin\left(j_n\frac{2a}{R}\right)\sin\left(j_n\frac{h}{R}\right)}{\left[\frac{2y_n^{(l)}}{\gamma^2} + \frac{1}{\lambda^2}\frac{\nu}{u}\frac{1}{(y_n^{(l)} - \nu/u)^2}\right]},$$

$$f_{1,2}(\hat{z}) = \eta(-\hat{z}), \quad f_3(\hat{z}) = -\eta(\hat{z}).$$

For $\nu \neq 0$ and for finite values of γ, we calculate the above two Fourier convolutions to arrive at the final expression for the force acting on the beam electrons:

$$F = -\frac{4\pi e^2 n_b}{R}\left\{\sum_{n=1}^{\infty}\left(-\frac{C_{n1}}{y_n^{(1)}} - \frac{C_{n2}}{y_n^{(2)}} + \frac{C_{n3}}{y_n^{(3)}}\right)\right.$$
$$\times\frac{k_0^2 - \alpha^2}{\alpha^2 + k_0^2}\frac{1}{k_0}e^{\alpha\hat{z}}\sin k_0\hat{z} + \sum_{l=1}^{2}\sum_{n=1}^{\infty}\frac{C_{n1}}{y_n^{(l)}}\frac{1}{(-y_n^{(l)} + \alpha)^2 + k_0^2}$$
$$\times\left[y_n^{(l)}e^{y_n^{(l)}\hat{z}} + e^{\alpha\hat{z}}\left(-y_n^{(l)}\cos k_0\hat{z}\right.\right.$$
$$\left.\left.+\left(k_0 - \frac{\alpha}{k_0}(-y_n^{(l)} + \alpha)\right)\sin k_0\hat{z}\right)\right]$$

$$-\frac{C_{n3}}{y_n^{(3)}}\frac{1}{(-y_n^{(3)}+\alpha)^2+k_0^2}\left[y_n^{(3)}e^{-y_n^{(3)}(\hat{z}+u\tau)}\right.$$

$$\left.+\,e^{\alpha\hat{z}}\left(-y_n^{(3)}\cos k_0\hat{z}+\left(k_0-\frac{\alpha}{k_0}(-y_n^{(3)}+\alpha)\right)\sin k_0\hat{z}\right)\right]\Bigg\}.$$

Note that the total contribution of both of the relativistic series to the force F is proportional to $1/\gamma$.

PLASMA CHANNEL BOUNDED BY WALLS WITH FINITE CONDUCTIVITY

We consider an REB transported in a plasma channel bounded by walls with a finite conductivity. In this situation, the walls may be either metal conductors with a finite conductivity σ or conducting media (plasma). Let us denote the parameters of the plasma channel and of the conducting walls by the indices 1 and 2, respectively. Now, our task is to solve Eq. (7) for the electric field component E_z with the nonuniform boundary conditions

$$\{E_z\}_{x=\pm R} = 0, \quad \{B_y\} \equiv \left\{\frac{\varepsilon}{\kappa^2}\frac{\partial E_z}{\partial x}\right\}_{x=\pm R} = 0, \quad (15)$$

which imply that the z and y components of the electromagnetic field should be continuous at the boundary between the two media.

According to the superposition principle, we can represent the electric field component in question as

$$E_z = E_z^{(1)} + E_z^{(2)}.$$

Since the solution $E_z^{(1)}$ to the inhomogeneous equation (7) with uniform boundary conditions was constructed above, we only need to find the solution $E_z^{(2)}$ to the homogeneous equation (7),

$$\mathcal{L}E_z^{(2)} = 0,$$

with the nonuniform boundary conditions (15).

The solution to Eq. (7) with boundary conditions (15) can be constructed using a familiar approach. Inside the plasma channel bounded by highly conducting walls, we have

$$E_z^{(2)} = -8\pi e n_b \frac{i\kappa_b}{k\kappa_1\varepsilon_2}\frac{\sinh\kappa_1 h}{\sinh^2 2\kappa_1 R}$$

$$\times[\sinh\kappa_1(R+x)\sinh\kappa_1(R+a)$$

$$+\,\sinh\kappa_1(R-x)\sinh\kappa_1(R-a)]\tilde{j}(k).$$

The component $F_2(k, x)$ of the force acting on both the beam and plasma electrons has the form

$$F_2(k,x) = -8\pi e^2 n_b \frac{\kappa_2}{k^2\varepsilon_2}\frac{\sinh\kappa_1 h}{\sinh^2 2\kappa_1 R}[\cosh\kappa_1(R+x)$$

$$\times\sinh\kappa_1(R+a) - \cosh\kappa_1(R-x)\sinh\kappa_1(R-a)]\tilde{j}(k).$$

We thus arrive at the sought-for expression for the component $F_2(k, a)$ of the force acting on the beam electrons that occur at the symmetry plane of the channel:

$$F_2(k,a) = -\frac{8\pi e^2 n_b}{R}$$

$$\times R\frac{\sinh(\kappa_1 h)\sinh(2\kappa_1 a)\cosh(2\kappa_1 R)}{\sinh^2(2\kappa_1 R)}\frac{\kappa_2}{k^2\varepsilon_2}\tilde{j}(k) \quad (16)$$

$$= -\frac{8\pi e^2 n_b}{R}G_1(k)G_2(k)\tilde{j}(k).$$

One can readily see that the contribution of the plasma channel and the walls to expression (16) for the force component F_2 is multiplicative, which allows a separate treatment of media 1 and 2, whose properties are described in terms of the inverse Fourier transforms of the functions G_l. The desired force component can be solved for by successively calculating the convolutions of these inverse Fourier transforms, which, of course, are to be found in advance.

First, we consider the model of an ultrarelativistic electron beam ($\gamma \longrightarrow \infty$) transported in a channel filled with a collisionless plasma ($\nu_1 \longrightarrow 0$). In this model, expression (16) takes the form

$$F_2(k,a) = -\frac{8\pi e^2 n_b}{R}G_1(\infty)G_2(k)\tilde{j}(k),$$

$$G_1(\infty) = \frac{\sinh\dfrac{h}{\lambda_1}\sinh\dfrac{2a}{\lambda_1}\cosh\dfrac{2R}{\lambda_1}}{\sinh^2\dfrac{2R}{\lambda_1}}.$$

Thus, we have to calculate the inverse Fourier transform of the function $G_2(k) = \kappa_2/(k^2\varepsilon_2)$ with respect to the z coordinate.

The physical condition that the electric field should decrease at infinity,

$$\mathrm{Re}\,\kappa_2 > 0 \Leftrightarrow \begin{cases}\mathrm{Im}\,\kappa_2^2 = 0\\ \mathrm{Re}\,\kappa_2^2 < 0,\end{cases}$$

corresponds to the cut $(0, -\nu_2)$ (which will be denoted by Γ on the imaginary axis (Fig. 3), in which case we

have

$$\kappa_2^2\big|_\Gamma = \frac{1}{\lambda_2^2}\frac{\nu}{\nu+\bar\nu_2} < 0 \Leftrightarrow \kappa_2\big|_{\Gamma_\mp} = \pm\frac{i}{\lambda_2^2}\sqrt{-\frac{\nu}{\nu+\bar\nu_2}}.$$

Consequently, for $\omega_{p2} > \nu_2/2$, the function $G_2(k)$ has two poles in the lower half-plane of the complex variable $k = u + iv$,

$$k_{1,2} = -i\alpha_2 \pm k_0^{(2)}, \quad k_0^{(2)} = \sqrt{\frac{1}{\lambda_2^2} - \alpha_2^2},$$

$$\alpha_2 = \frac{\bar\nu_2}{2}, \quad \bar\nu_i = \frac{\nu_i}{c},$$

and is characterized by the cut Γ between the poles (Fig. 3). With these properties in mind, we obtain for $G_2(\hat z)$:

$$G_2 = -e^{\alpha_2\hat z}(\tilde A\cos k_0^{(2)}\hat z + \tilde B\sin k_0^{(2)}\hat z)\eta(-\hat z)$$
$$-\frac{1}{\pi}\mathcal{P}\int_0^{\tilde\lambda_2}\sqrt{\frac{\tilde\lambda_2 - x}{x}}\frac{e^{x\hat z/\lambda_2}}{x^2 - \tilde\lambda_2 x + 1}dx\eta(-\hat z),$$

$$\tilde A = a\bar\nu_2\lambda_2^2 - \frac{b}{k_0^{(2)}}(1 - \alpha_2\bar\nu_2\lambda_2^2),$$

$$\tilde B = \frac{a}{k_0^{(2)}}(1 - \alpha_2\bar\nu_2\lambda_2^2),$$

$$\kappa_2^2(k_{1,2}) = \left(\frac{1}{\lambda_2^2} - 2\alpha_2^2\right) \mp ik_0^{(2)}2\alpha_2$$

$$= \rho^2 e^{\mp\Phi i} \Rightarrow \kappa_2(k_{1,2}) = \rho e^{\mp\Phi i} = a \mp ib,$$

$$\Phi = \frac{1}{2}\arctan\frac{2k_0^{(2)}\alpha_2}{\lambda_2^{-2} - 2\alpha_2^2},$$

$$\tilde\lambda_2 = \frac{\lambda_2}{z_{\nu_2}}, \quad z_{\nu_2} = \frac{c}{\nu_2}, \quad \tilde\lambda_2 < \frac{1}{2}.$$

The Fourier convolution of $G_1(\hat z)$ with $\tilde j(\hat z)$ can be calculated in the same way as in section 2. As a result, under the conditions

$$\nu_1 = 0, \quad \gamma \longrightarrow \infty, \quad \frac{\nu_2}{2} < \omega_{p2},$$

we arrive at the following expression for the component F_2 of the sought-for force

$$F_2(\hat z, a) = 8\pi e^2 n_b\lambda_2\frac{G_1(\infty)}{R}\left\{\frac{1}{\pi}\mathcal{P}\int_0^{\tilde\lambda_2}\frac{1}{x}\sqrt{\frac{\tilde\lambda_2 - x}{x}}\right.$$
$$\times\frac{1 - e^{x\hat z/\lambda_2}}{x^2 - \tilde\lambda_2 x + 1}dx + \lambda_2[(\alpha_2\tilde A - k_0^{(2)}\tilde B)$$
$$- e^{\alpha_2\hat z}((\alpha_2\tilde A - k_0^{(2)}\tilde B)\cos k_0^{(2)}\hat z$$
$$\left. + (k_0^{(2)}\tilde A + \alpha_2\tilde B)\sin k_0^{(2)}\hat z)\right]\Big\}. \quad (17)$$

In the collisionless limit, the expression for the contribution of the component F_2 to the force F acting on the REB follows from formula (17):

$$F_2(x = a) = 8\pi e^2 n_b\lambda_2\left[1 - \cos\frac{\hat z}{\lambda_2}\right]\frac{G_1(\infty)}{R}.$$

Under the conditions $\gamma \longrightarrow \infty$, $\nu_1 = 0$, and $\omega_{p2} < \nu_2/2$, which imply that no oscillations are excited in the plasma channel, the poles of the function $G_1(k)$ (or, equivalently, the zeros of the function ε_2) lie on the cut Γ (Fig. 4) and do not contribute to the desired expression for the force component in question. Thus, we are left with the problem of determining the contri-

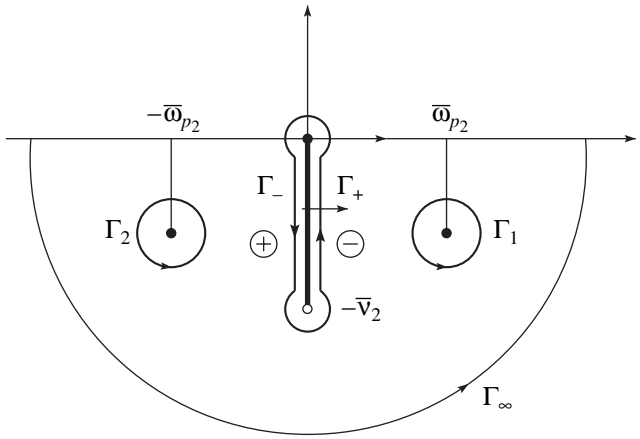


Fig. 3.

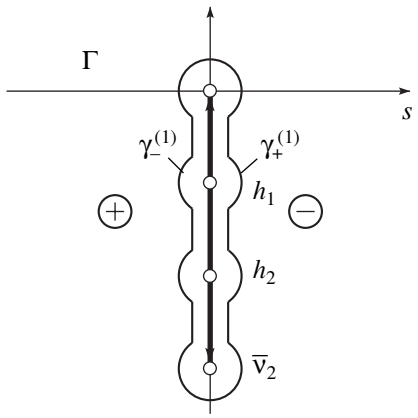


Fig. 4.

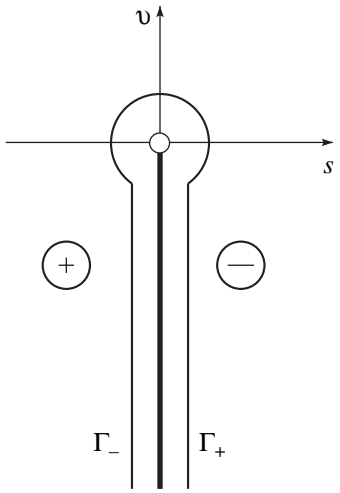


Fig. 5.

bution of the cut Γ, in which case the integral along the contour that envelopes the cut should be understood in terms of the Cauchy principal value. The expression for $F_2(\hat{z}, \nu_1 = 0)$ is governed by the first term in braces in formula (17).

First, we consider the so-called σ-model, in which the wall properties are described by Ohm's law:

$$\mathbf{j} = \sigma\mathbf{E}.$$

The physical condition that the electric field should decrease at infinity corresponds to the semi-infinite cut Γ on the imaginary axis (Fig. 5). On this cut, the function $G_2(k)$ has the form

$$G_2(k)|_\Gamma = \left.\frac{\kappa_2}{k^2\varepsilon_2}\right|_\Gamma = G_2(i\nu)|_{\Gamma_\mp} = \pm i\frac{\sqrt{-\nu}}{\nu\left(\nu + \dfrac{4\pi\sigma}{c}\right)}.$$

We successively calculate the inverse Fourier transform $G_2(\hat{z}) - G_2(k)$ and its convolution with the Fourier-transformed beam current density. As a result, we arrive at the following formula for the force component $F_2(\hat{z}, \nu_1 = 0)$:

$$F_2(\hat{z}, a) \equiv F_2(\hat{z}, \nu_1 = 0)$$
$$= \frac{8\pi e^2 n_b}{R} G_1(\infty)\frac{z_m}{\pi}\mathcal{P}\int_0^\infty \frac{dx\sqrt{x}}{x^2(1-x)}(1 - e^{x\hat{z}/z_m}),$$

where $z_m = \nu_m/c$ and $\nu_m = c^2/(4\pi\sigma)$ is the magnetic viscosity.

We consider an ultrarelativistic REB ($\gamma \longrightarrow \infty$) transported through a collisional plasma ($\nu_1 \neq 0$). For such a system, it is sufficient to determine the inverse Fourier transform $G_2(\hat{z})$ of the function $G_2(k)$ and then to calculate the convolution of this inverse Fourier transform with the corresponding function obtained in the collisionless limit $\nu_1 = 0$. In this case, the component F_2 of the desired force can be found from the formula

$$F_2(\hat{z}, a) = (F_2(\hat{z}, \nu_1 = 0)G_1(\hat{z})/G_1(\infty)), \qquad (18)$$

in which the component $F_2(\hat{z}, \nu_1 = 0)$ was determined above.

The inverse Fourier transform $G_1(\hat{z})$ should be calculated with allowance for the properties of the function $G_1(k, \nu_1)$ both at infinity and in the vicinity of the poles of the collisional series $k_n^{(1)} = -iy_n^{(1)}$. One can readily single out the singularities of the function $G_1(k, \nu_1)$ in the vicinities $k \longrightarrow k_n^{(1)}$ of the second-order poles $k_n^{(1)}$

$$G_1(k, \nu_1) \sim \frac{f(k_n^{(1)})}{(k - k_n^{(1)})^2},$$

where

$$f(k_n^{(1)}) = (-1)^{n+1}\left(\frac{\nu_1}{c}\right)^2\frac{R}{j_n^2}$$
$$\times \sin\left(j_n\frac{2a}{R}\right)\sin\left(j_n\frac{h}{R}\right)\frac{1}{(1 + d_n^2)^2}.$$

At infinity, the function $G_1(k, \nu_1)$ approaches a nonzero constant.

The contribution of the second-order poles is calculated from the familiar Cauchy formula:

$$G_1(\hat{z}, \nu_1) = \sum_n \frac{1}{2\pi}\int_{C_n}\frac{f(k_n^{(1)})e^{ik\hat{z}}}{(k - k_n^{(1)})^2}dk$$
$$= \sum_{n=1}^\infty f(k_n^{(1)})\hat{z}e^{y_n^{(1)}\hat{z}}\eta(-\hat{z}). \qquad (19)$$

Consequently, for $\nu_1 \neq 0$, the inverse Fourier transform $G_1(\hat{z})$ is described by the expression

$$G_1(\hat{z}) = G_1(\infty)\delta(\hat{z}) + G_1(\hat{z}, \nu_1). \qquad (20)$$

We can easily see that the contribution of the first term in expression (20) to the force component F_2 coincides with the contribution that was already obtained in the collisionless limit $\nu_1 = 0$. As a result, the expression for the force component F_2 contains two terms:

$$F_2(\hat{z}, a) = F_2(\hat{z}, \nu_1 = 0) + F_2(\hat{z}, \nu_1). \qquad (21)$$

The contribution $F_2(\hat{z}, \nu_1)$ of collisions in the plasma channel (medium 1) to expression (21) for the desired force component is determined by substituting the formulas for $F_2(\hat{z}, \nu_1 = 0)$ obtained in different physical models of the wall substance and calculating

Fourier convolutions. Thus, under the conditions $\gamma \longrightarrow \infty$, $\nu_1 \neq 0$, and $\omega_{p2} > \nu_2/2$, we arrive at the expression

$$F_2(\hat{z}, \nu_1) = \frac{8\pi e^2 n_b}{R}\sum_{n=1}^{\infty} f(k_n^{(1)})$$
$$\times\left\{\frac{\lambda_2^3}{\pi}\mathcal{P}\int_0^{\tilde{\lambda}_2}\left[\frac{e^{y_n^{(1)}\hat{z}}}{x-y_n^{(1)}\lambda_2}\left(\frac{e^{(x-y_n^{(1)}\lambda_2)\hat{z}}}{x-y_n^{(1)}\lambda_2}-\hat{z}\right)\right.\right.$$
$$\left.+\frac{1}{y_n^{(1)}\lambda_2}\left(\frac{e^{y_n^{(1)}\hat{z}}-1}{y_n^{(1)}\lambda_2}-\frac{\hat{z}}{\lambda_2}e^{y_n^{(1)}\hat{z}}\right)\right]\frac{1}{x}\sqrt{\frac{\tilde{\lambda}_2-x}{x}}$$
$$\times\frac{dx}{x^2-\tilde{\lambda}_2 x+1}+\lambda_2^2\left\{\frac{C}{y_n^{(1)}}\left[e^{y_n^{(1)}\hat{z}}\left(-\hat{z}+\frac{1}{y_n^{(1)}}\right)-\frac{1}{y_n^{(1)}}\right]\right.$$
$$+\frac{1}{(\alpha_2-y_n^{(1)})^2+k_0^{(2)^2}}\left[e^{y_n^{(1)}\hat{z}}\left[(k_0^2 D-(\alpha_2-y_n^{(1)})C)\hat{z}\right.\right.$$
$$-\frac{1}{(\alpha_2-y_n^{(1)})^2+k_0^{(2)^2}}(2D(\alpha_2-y_n^{(1)})k_0^{(2)} \qquad (22)$$
$$\left.\left.+C((\alpha_2-y_n^{(1)})^2+k_0^{(2)^2})\right)\right]$$
$$+\frac{e^{\alpha_2\hat{z}}}{(\alpha_2-y_n^{(1)})^2+k_0^{(2)^2}}[(2D(\alpha_2-y_n^{(1)})k_0^{(2)}$$
$$+C((\alpha_2-y_n^{(1)})^2+k_0^{(2)^2}))\cos k_0^{(2)}\hat{z}$$
$$+(D((\alpha_2-y_n^{(1)})^2-k_0^{(2)^2})$$
$$\left.\left.-2C(\alpha_2-y_n^{(1)})k_0^{(2)})\sin k_0^{(2)}\hat{z}]\right]\right\}\Bigg\},$$

where $C = \alpha_2\tilde{A} - k_0^{(2)}\tilde{B}$, $D = k_0^{(2)}\tilde{A} + \alpha_2\tilde{B}$, and the quantities $\tilde{A}$ and $\tilde{B}$ were defined above.

In the model in which plasma medium 2 is collisionless ($\nu_2 = 0$), expression (22) gives

$$F_2(\hat{z}, \nu_1) = \frac{8\pi e^2 n_b\lambda_2^3}{R}\sum_{n=1}^{\infty} f(k_n^{(1)})\left\{-\frac{1}{y_n^{(1)}\lambda_2}\right.$$
$$\times\left[e^{y_n^{(1)}\hat{z}}\left(-\frac{\hat{z}}{\lambda_2}+\frac{1}{y_n^{(1)}\lambda_2}\right)-\frac{1}{y_n^{(1)}\lambda_2}\right]$$
$$+\frac{1}{(y_n^{(1)}\lambda_2)^2+1}\left\{e^{y_n^{(1)}\hat{z}}\left[-y_n^{(1)}\hat{z}+\frac{(y_n^{(1)}\lambda_2)^2-1}{(y_n^{(1)}\lambda_2)^2+1}\right]\right.$$
$$\left.\left.-\frac{1}{(y_n^{(1)}\lambda_2)^2+1}\left[((y_n^{(1)}\lambda_2)^2-1)\cos\frac{\hat{z}}{\lambda_2}+2y_n^{(1)}\lambda_2\sin\frac{\hat{z}}{\lambda_2}\right]\right\}\right\},$$

so that, for $\gamma \longrightarrow \infty$, $\nu_1 \neq 0$ and $\omega_{p2} > \nu_2/2$, we obtain

$$F_2(\hat{z}, \nu_1) = \frac{8\pi e^2 n_b}{R}\sum_{n=1}^{\infty} f(k_n^{(1)})\frac{\lambda_2^3}{\pi}$$
$$\times\mathcal{P}\int_0^{\tilde{\lambda}_2}\left[\frac{e^{y_n^{(1)}\hat{z}}}{x-y_n^{(1)}z_m}\left(\frac{e^{(x-y_n^{(1)}z_m)\hat{z}}}{x-y_n^{(1)}z_m}-\hat{z}\right)\right.$$
$$\left.+\frac{1}{y_n^{(1)}\lambda_2}\left(\frac{e^{y_n^{(1)}\hat{z}}-1}{y_n^{(1)}\lambda_2}-\frac{\hat{z}}{\lambda_2}e^{y_n^{(1)}\hat{z}}\right)\right]\frac{1}{x}\sqrt{\frac{\tilde{\lambda}_2-x}{x}}\frac{dx}{x^2-\tilde{\lambda}_2 x+1}.$$

In the σ model, for $\gamma \longrightarrow \infty$ and $\nu_1 \neq 0$, we find

$$F_2(\hat{z}, \nu_1) = \frac{8\pi e^2 n_b}{R}\sum_{n=1}^{\infty} f(k_n^{(1)})\frac{\hat{z}_m^3}{\pi}$$
$$\times\mathcal{P}\int_0^{\infty}\left[\frac{e^{y_n^{(1)}\hat{z}}}{x-y_n^{(1)}z_m}\left(\frac{e^{(x-y_n^{(1)}z_m)\hat{z}}}{x-y_n^{(1)}z_m}-\hat{z}\right)\right.$$
$$\left.+\frac{1}{y_n^{(1)}z_m}\left(\frac{e^{y_n^{(1)}\hat{z}}-1}{y_n^{(1)}z_m}-\frac{\hat{z}}{z_m}e^{y_n^{(1)}\hat{z}}\right)\right]\frac{\sqrt{x}dx}{x^2(1-x)}.$$

Now, we proceed to a calculation of the inverse Fourier transforms of the function $G_2(k)$, which serves to describe the properties of the conducting wall in different models.

In the σ model of the wall substance, the finiteness of the relativistic factor is easy to take into account. The physical condition that the electric field should decrease at infinity corresponds to the two cuts Γ_i on the imaginary axis (Fig. 6):

$$\Gamma_i:\begin{cases}s = 0\\ \left[\begin{array}{l}v > v_0 = \dfrac{\beta\gamma^2}{z_m}\\ v < 0.\end{array}\right.\end{cases}$$

Also, the function

$$G_2(k) = \frac{\kappa_2}{k^2\varepsilon_2} = \frac{\kappa_2}{k\left(k+i\dfrac{1}{\beta z_m}\right)}$$

has the first-order pole

$$k_2 = -i\frac{1}{\beta z_m} = -i\nu_2.$$

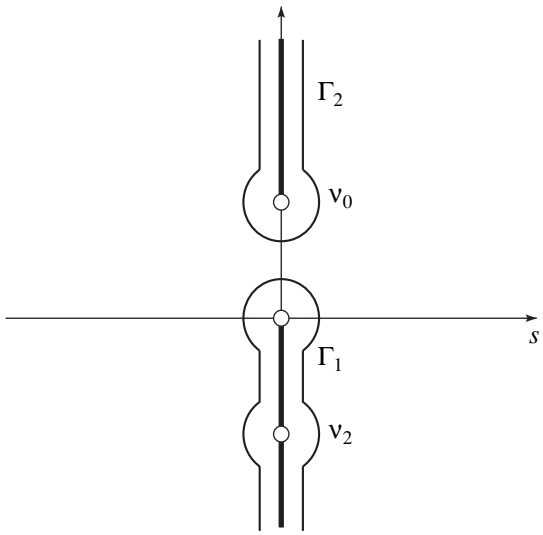


Fig. 6.

One can readily see that the pole k_2 does not contribute to the corresponding integral because the function κ_2 has different signs on different sides of the cut. In addition, the singular point $k_1 = 0$ is a removable singularity. As a result, the integral should again be understood in terms of the Cauchy principal value, so that we arrive at the following expression for the inverse Fourier transform $G_2(\hat{z})$ of the function $G_2(k)$:

$$G_2(\hat{z}) = -\frac{1}{\pi}\mathcal{P}\int_0^\infty \frac{\sqrt{x\left(\frac{x}{\gamma^2}+\beta\right)}}{(\beta^{-1}-x)x}e^{x\hat{z}/z_m}dx\eta(-\hat{z}) + \frac{1}{\pi}\int_{\beta\gamma^2}^\infty \frac{\sqrt{x\left(-\frac{x}{\gamma^2}+\beta\right)}}{(\beta^{-1}+x)x}e^{-x\hat{z}/z_m}dx\eta(\hat{z}). \quad (23)$$

The first and second terms on the right-hand side of this expression are the contributions of the cuts Γ_1 and Γ_2, respectively (Fig. 6). For $\gamma \longrightarrow \infty$, the second term, which describes the perturbation ahead of the beam front, vanishes. For finite values of γ, we first determine the inverse Fourier transform of the function $G_1(k)$ by using the above multiplicative property of expression (16) for the force component F_2. The singular points where the denominator of the function $G_1(k)$ vanishes are second-order poles; specifically, these are the poles of the ν series and the two γ series, which were already found in Section 2. In this case, the contribution of the second-order poles can be evaluated in the same way as was done when deriving formula (19). Thus, we successively find

$$G_1(k) - G_1(\hat{z}) = \sum_{l=1}^{3}\sum_{n=1}^{\infty} f(k_n^{(l)})\hat{z}e^{y_n^{(l)}\hat{z}}f_l(\hat{z}),$$

$$f(k_n^{(l)}) = (-1)^{n+1}\frac{j_n^2}{R^3}\frac{\sin\left(j_n\frac{2a}{R}\right)\sin\left(j_n\frac{h}{R}\right)}{\left[\frac{2y_n^{(l)}}{\gamma^2}+\frac{1}{\lambda_1^2}\frac{\nu_1}{u}\frac{1}{\left(y_n^{(l)}-\frac{\nu_1}{u}\right)^2}\right]^2}.$$

The expression for the force component $F_2(\hat{z}, a)$ can be obtained by calculating the corresponding two Fourier convolutions. Performing simple but rather laborious manipulations and taking into account finite γ values and collisions in the plasma channel, we arrive at the following formula for the force component F_2 in the σ model of the wall substance:

$$F_2(\hat{z}, a) = -\frac{8\pi e^2 n_b z_m^3}{R}\left\{\sum_n f_n^{(3)}e^{y_n^{(3)}(\hat{z}+u\tau)}\right.$$

$$\times\mathcal{P}\int_0^\infty dx\frac{\sqrt{x\left(\frac{x}{\gamma^2}+\beta\right)}}{(\beta^{-1}-x)x^2}\frac{(y_n^{(3)}z_m-x)\frac{\hat{z}+u\tau}{z_m}-1}{(y_n^{(3)}z_m-x)^2}(1-e^{-xu\tau/z_m})$$

$$-\sum_n f_n^{(3)}e^{y_n^{(3)}(\hat{z}+u\tau)}\int_0^\infty dx\frac{\sqrt{x\left(\frac{x}{\gamma^2}+\beta\right)}}{(\beta^{-1}-x)x^2}$$

$$\times\left[\frac{1}{(y_n^{(3)}z_m)^2}(1+e^{y_n^{(3)}(\hat{z}+u\tau)}(y_n^{(3)}(\hat{z}+u\tau)-1))\right.$$

$$-\frac{e^{x\hat{z}/z_m}}{(y_n^{(3)}z_m-x)^2}\left(1+e^{(y_n^{(3)}z_m-x)(\hat{z}+u\tau)/z_m}\right.$$

$$\left.\left.\times\left((y_n^{(3)}z_m-x)\frac{\hat{z}+u\tau}{z_m}-1\right)\right)\right] - \sum_n f_n^{(3)}\int_{\beta\gamma^2}^\infty dx\frac{\sqrt{x\left(-\frac{x}{\gamma^2}+\beta\right)}}{x^2(x+\beta^{-1})}$$

$$\times\left[\frac{1+e^{y_n^{(3)}(\hat{z}+u\tau)}}{(y_n^{(3)}z_m)^2}(y_n^{(3)}(\hat{z}+u\tau)-1)\right.$$

$$-\frac{e^{-x(\hat{z}+u\tau)/z_m}}{(y_n^{(3)}z_m+x)^2}(1+e^{y_n^{(3)}z_m+x)(\hat{z}+u\tau)/z_m}$$

$$\left.\times((y_n^{(3)}z_m+x)(\hat{z}+u\tau)-1)\right]$$

$$+\sum_{l=1}^{2}\sum_n f_n^{(l)}\mathcal{P}\int_0^\infty dx\frac{\sqrt{x\left(\frac{x}{\gamma^2}+\beta\right)}}{(\beta^{-1}-x)x^2}$$

$$\times\left[-\frac{1}{(y_n^{(l)}z_m)^2}(e^{y_n^{(l)}\hat{z}}(y_n^{(l)}\hat{z}-1)+1)+\frac{e^{x\hat{z}/z_m}}{(y_n^{(l)}z_m-x)^2}\right.$$

$$\left.\times\left(e^{(y_n^{(l)}z_m-x)/z_m}\left((y_n^{(l)}z_m-x)\frac{\hat{z}}{z_m}-1\right)+1\right)\right]$$

$$+\sum_{l=1}^{2}\sum_{n}f_n^{(l)}\int_{\beta\gamma^2}^{\infty}dx\frac{\sqrt{x\left(-\frac{x}{\gamma^2}+\beta\right)}}{(\beta^{-1}+x)x^2}$$

$$\times\left[-\frac{1}{(y_n^{(l)}z_m)^2}(e^{y_n^{(l)}\hat{z}}(y_n^{(l)}\hat{z}-1)+1)\right.$$

$$\left.+\frac{e^{-x(\hat{z}+u\tau)/z_m}}{(y_n^{(l)}z_m+x)^2}(e^{(y_n^{(l)}z_m+x)\hat{z}/z_m}((y_n^{(l)}z_m+x)\hat{z}/z_m-1)+1)\right]$$

$$+\sum_{l=1}^{2}\sum_{n}f_n^{(l)}\int_{\beta\gamma^2}^{\infty}dx\frac{\sqrt{x\left(-\frac{x}{\gamma^2}+\beta\right)}}{(\beta^{-1}+x)x^2}$$

$$\times\left\{(e^{-x\hat{z}/z_m}-e^{-x(\hat{z}+u\tau)/z_m})\right.$$

$$\left.\times e^{(y_n^{(l)}z_m+x)\hat{z}/z_m}\frac{(y_n^{(l)}z_m+x)\hat{z}/z_m-1}{(y_n^{(l)}z_m+x)^2}\right]\Bigg\}.$$

Finally, we consider the model in which the properties of the conducting wall are described by the dispersion law (2), the relativistic factor is finite, and the condition $\omega_{p2} > \nu/2$ holds. For this model, the singularities of the function $G_2(k)$ are illustrated in Fig. 7. These are three cuts Γ, which lie on the imaginary axis of the plane of the complex variable $k = s + i\nu$, and two poles $k_{1,2} = \pm k_0^{(2)} - i\alpha_2$, where $k_0^{(2)} = \sqrt{(\beta\lambda_2)^{-2}-\alpha_2^2}$, $\alpha_2 = \bar{\nu}_2/2$, and $\bar{\nu}_2 = \nu_2/u$:

$$\Gamma_i:\ \begin{cases} s = 0 \\ \dfrac{\nu(\nu-\nu_1)(\nu-\nu_2)}{\nu-\bar{\nu}_2} > 0, \end{cases} \quad i = 1, 2, 3$$

$$\nu_{1,2} = -\frac{1}{2}\bar{\nu}_2 \mp \sqrt{\frac{\bar{\nu}_2^2}{4}+\frac{\gamma^2}{\beta^2\lambda_2^2}}.$$

For $\omega_{p2} > \nu_2/2$, the contributions of the poles to the inverse Fourier transform $G_2(\hat{z})$ of the function $G_2(k)$ are evaluated in a way similar to what we did in the limit $\gamma \longrightarrow \infty$, in which case, however, the expressions

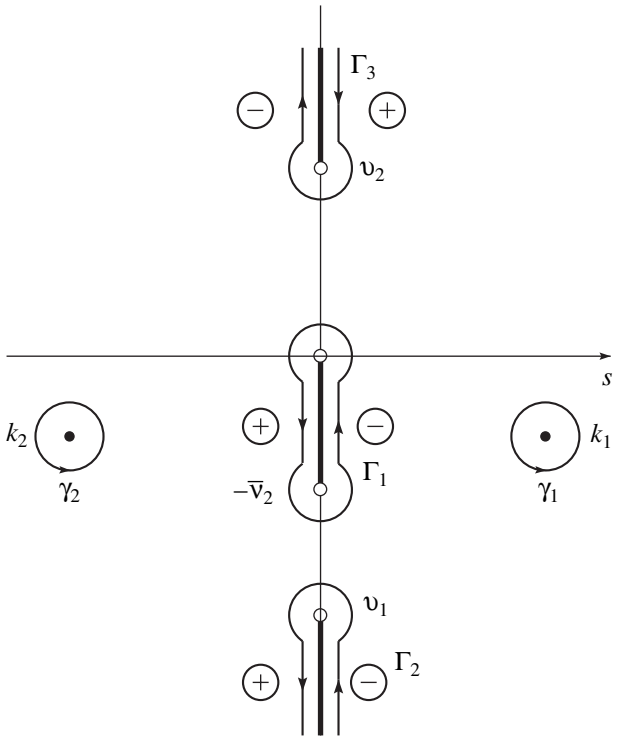


Fig. 7.

for $\kappa_2(k_{1,2})$, a, b, and Φ are somewhat different:

$$\kappa_2^2(k_{1,2}) = \left(\frac{1}{\lambda_2^2}-2\alpha_2^2\right)\left(1+\frac{1}{\gamma^2}\right) \mp i2\alpha_2 k_0^{(2)}\beta^2 = \rho^2 e^{\mp 2\Phi i},$$

$$\kappa_2(k_{1,2}) = \rho e^{\mp\Phi i} = a \mp bi,$$

$$\Phi = \frac{1}{2}\arctan\frac{2\alpha_2 k_0^{(2)}\beta^2}{(\lambda_2^{-2}-2\alpha_2^2)(1+1/\gamma^2)}.$$

As a result, we obtain

$$G_{1-2}(\hat{z}) \equiv G(\hat{z})|_{k_{1,2}}$$

$$= \frac{\lambda_2^2}{k_0^{(2)}}e^{\alpha_2\hat{z}}\{[2a\alpha_2 k_0^{(2)} - b(\lambda_2^{-2}-2\alpha_2^2)]$$

$$\times\cos k_0^{(2)}\hat{z} + [2b\alpha_2 k_0^{(2)} + a(\lambda_2^{-2}-2\alpha_2^2)]\sin k_0^{(2)}\hat{z}\}\eta(-\hat{z}).$$

The inverse Fourier transform $G_2(\hat{z})$ is determined by the additive contributions of the poles and the cuts Γ_i. The contributions of the cuts are as follows:

$$G(\hat{z})|_{\Gamma_1} = -\frac{1}{\pi}\mathcal{P}\int_0^{\tilde{\lambda}_2}\frac{dx}{x}\sqrt{\frac{\tilde{\lambda}_2-x}{x}+\frac{1}{\gamma^2}(\tilde{\lambda}_2-x)^2}$$

$$\times\frac{e^{x\frac{\hat{z}}{\tilde{\lambda}_2}}}{x^2-\tilde{\lambda}_2 x+1}\eta(-\hat{z}),$$

$$G(\hat{z})|_{\Gamma_2} = \frac{1}{\pi}\mathcal{P}\int_{x_1}^{\infty}\frac{dx}{x}\sqrt{\frac{\tilde{\lambda}_2 - x}{x} + \frac{1}{\gamma^2}(\tilde{\lambda}_2 - x)^2}$$
$$\times \frac{e^{x\frac{\hat{z}}{\tilde{\lambda}_2}}}{x^2 - \tilde{\lambda}_2 x + 1}\eta(-\hat{z}),$$
$$G(\hat{z})|_{\Gamma_3} = \frac{1}{\pi}\int_{x_2}^{\infty}\frac{dx}{x}\sqrt{-\frac{\tilde{\lambda}_2 + x}{x} + \frac{1}{\gamma^2}(\tilde{\lambda}_2 - x)^2}$$
$$\times \frac{e^{-x\frac{\hat{z}}{\tilde{\lambda}_2}}}{x^2 + \tilde{\lambda}_2 x + 1}\eta(\hat{z}),$$

where

$$x_{1,2} = \sqrt{\left(\frac{\nu_2}{2\beta\omega_{p2}}\right)^2 + \gamma^2} \pm \frac{\nu_2}{2\beta\omega_{p2}}.$$

For $\omega_{p2} < \nu_2/2$, the contribution of the poles equals zero; i.e., the term $G_2(\hat{z})$ drops out of the expression for the inverse Fourier transform $G_{1-2}(\hat{z})$. In other respects, the force component F_2 is calculated in the same way as in the σ model.

CONCLUSION

The results obtained in this paper make it possible to apply different models of the media of the plasma channel and the wall and take into account a wide variety of physical circumstances in order to describe the force interaction of an REB with a plasma in a channel bounded by highly conducting walls. We have completely solved the problem of the transport of a planar beam of arbitrary thickness at an arbitrary distance from the symmetry plane of the channel. In the case of a cylindrical channel, we succeeded in analyzing the problem at hand only for thin beams.

ACKNOWLEDGMENTS

We would like to acknowledge the financial support from the Integration program.

REFERENCES

1. A. A. Rukhadze, L. S. Bogdankevich, S. E. Rosinskiĭ, and V. G. Rukhlin, *Physics of High-Current Relativistic Electron Beams* (Atomizdat, Moscow, 1980).
2. W. E. Martin, Phys. Rev. Lett. **54**, 685 (1985).
3. P. S. Strelkov and D. K. Ul'yanov, Fiz. Plazmy **26**, 329 (2000) [Plasma Phys. Rep. **26**, 303 (2000)].
4. A. P. Kuryshev, V. B. Sokolyuk, and S. V. Chernov, Zh. Tekh. Fiz. **57**, 1292 (1987) [Sov. Phys. Tech. Phys. **32**, 764 (1987)].
5. A. P. Kuryshev and S. V. Chernov, Zh. Tekh. Fiz. **58**, 2106 (1988) [Sov. Phys. Tech. Phys. **33**, 1279 (1988)].
6. E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **60** (3), 40 (1990) [Sov. Phys. Tech. Phys. **35**, 298 (1990)].
7. A. G. Zelenskiĭ and E. K. Kolesnikov, Zh. Tekh. Fiz. **65** (5), 188 (1995) [Tech. Phys. **40**, 510 (1995)].
8. V. P. Grigor'ev, A. V. Didenko, and G. P. Isaev, Fiz. Plazmy **9**, 1254 (1983) [Sov. J. Plasma Phys. **9**, 723 (1983)].
9. V. B. Vladyko and Yu. V. Rudyak, Zh. Tekh. Fiz. **60** (8), 199 (1990) [Sov. Phys. Tech. Phys. **35**, 993 (1990)].
10. A. P. Kuryshev and V. D. Andreev, Zh. Tekh. Fiz. **66** (8), 143 (1996) [Tech. Phys. **41**, 821 (1996)].

Translated by O. Khadin

Technical Physics, Vol. 46, No. 9, 2001, pp. 1161–1167. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 97–104.

SURFACES, ELECTRON AND ION EMISSION

The Properties of the Nonequilibrium LaB_6 Surface Resulting from Field Evaporation

V. N. Gurin, M. M. Korsukova, M. V. Loginov, and V. N. Shrednik

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

Received January 18, 2000

Abstract—The nonequilibrium surface of single-crystal lanthanum hexaboride needles and its modifications are studied with a time-of-flight atomic probe. The surface is obtained by room-temperature field evaporation. The mass spectra of field evaporation shed light on the surface composition at the needle tip immediately after tip etching, corrosion in residual gases, intense cleaning by field evaporation, and the relaxation of the nonequilibrium surface by heating to 1250 K. Conditions for the breakdown of an oxide film on the tip surface and for obtaining the mass spectra of field evaporation for stoichiometric or lanthanum-enriched pure LaB_6 single crystals are discussed. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Ion flows due to the room-temperature field evaporation of LaB_6 have been studied in [1, 2]. Those works have been concerned with the stability of free ionized clusters of the LaB_n^{+m} type (where n = 1–6 and m = 1–4), which are present in abundance in the ion flow. It has been shown that the flow largely contains poorly stable and even metastable ions.

In this work, we use the same experimental technique as in [1, 2] but concentrate on studying the LaB_6 surface resulting from the field evaporation and serving as an active source of evaporating ions. In addition, modifications of this surface due to heating or interaction with adsorbed residual gases are of no less interest. Moreover, it is significant to know the properties of the LaB_6 surface initially kept under a vacuum after it has been etched in concentrated H_2SO_4 to form a fine tip and then kept in air. All these points have not been touched upon in [1, 2] and are dealt with in this article. Finally, we will try to find conditions under which the ion flow is enriched by lanthanum. The knowledge of these conditions would help to design point ion sources.

EXPERIMENT

Experiments were carried out with a time-of-flight atomic probe [3] with a mass resolution $M/\Delta M \approx 30$. Recently, its detecting system has been upgraded: the rate of data collection has been raised 30 times or more and the time resolution has been greatly improved. Accordingly, the reliability of the mass spectra has been improved and the time of recording the mass spectra has been cut. The tips were made of the needles grown from the solution in aluminum melt [4]. As in [1, 2], the temperature of the tip was kept at the room value. The pressure in the atomic probe chamber was maintained at a level of 10^{-9}–10^{-10} torr with an electrical discharge pump. The major residual gases were CO, H_2O, and CH_4. The crystal tips were attached to the support (anode) with either Aquadag or conducting epoxy resin containing fine-grain tantalum powder. In the former case, the chamber with the tip can be heated to 200 and 1300°C, respectively, and the water content in the residual gases was insignificant. In the latter case, the chamber and the tip were kept at room temperature and the amount of water was higher. The well-defined water peak was used to calibrate the mass spectra.

RESULTS AND DISCUSSION

1. Mass spectra of LaB_6 field evaporation. Figure 1 shows three mass spectra within one series of runs with the same LaB_6 tip. In all spectra, the base, V_b, and the pulse, V_p, voltages were 13 and 5.6 kV, respectively. The horizontal and vertical scales for the three spectra are the same. The spectra differ because they have the different histories. The first spectrum (Fig. 1a) was taken from the as-etched LaB_6 surface and contains a minor amount of the ions generated by 59 pulses of a total of 5000 pulses applied. The highest peak is that of CO^+ (CO is among the residual gases).

It is of interest that the water peak in Fig. 1a is absent, although the tip was fixed by epoxy resin and the instrument was not heated. Several weak peaks rising above unit ones (the latter contain also noise peaks) can be identified as corresponding to oxide ions, $B_2O_3^{++}$, B_2O^+, and LaO_2^{+++}, and to La_2B^{++} ions. Before this spectrum, we recorded four spectra with V_b = 8, 9, and 11 kV and with the same V_p = 5.6 kV. They all con-

1063-7842/01/4609-1161$21.00

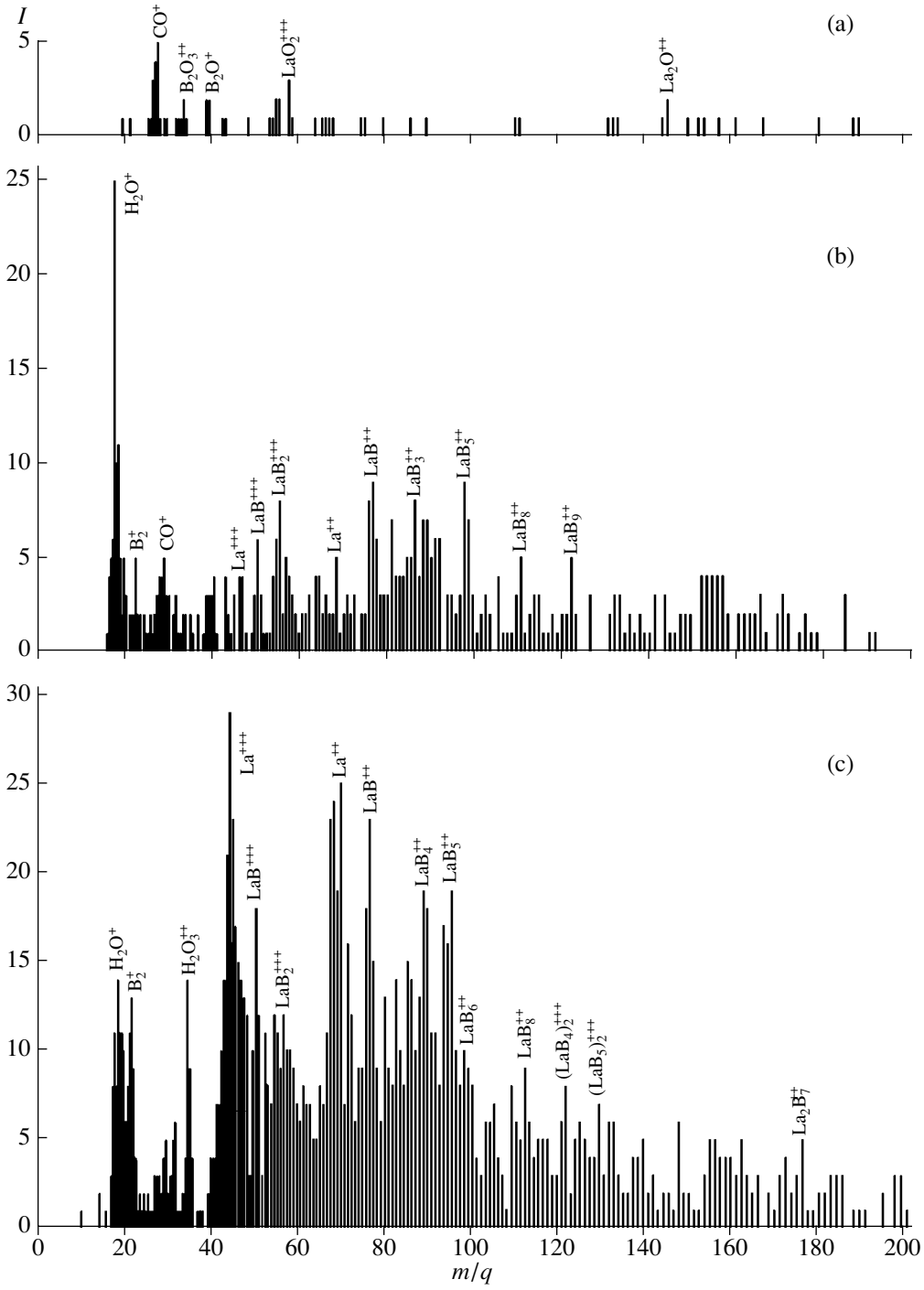


Fig. 1. Field evaporation mass spectra for the LaB_6 single crystal that were obtained with the atomic probe. The spectra were obtained for the sample after various treatments. m/q is the ion mass-to-charge ratio, and I is the number of ions. (a) The surface is covered by the oxide film after tip etching and exposure to air, n/N = 59/5000; (b) the surface after the removal of the film, n/N = 422/5000; and (c) the La-enriched surface, n/N = 1000/7586. N is the total number of pulses applied, n is the number of ion-generating pulses.

tained one well-defined CO^+ peak and the same weak peaks. None of the spectra contained water.

An increase in V_b to 15 kV resulted in a spectrum with a plenty of peaks. The spectrum was similar to that in Fig. 1a and also did not contain the H_2O^+ ion. This ion showed up as a high peak when V_b was decreased by 1 kV (at $V_b + V_p = 14 + 5.6$ kV). Then, V_b was decreased to 13 kV, and the spectrum depicted in Fig. 1b was obtained. The spectra in Figs. 1a and 1b radically differ: the latter contains the high well-defined

H_2O^+ peak and many cluster ions of the LaB_n^{+m} type. Of these peaks, those of the doubly charged ions were the highest. This spectrum does not contain excess lanthanum in large amounts, although an excess of lanthanum is typical of developed field evaporation of LaB_6 [1, 2]. It was found that the developed process is established if V_b is reduced for a while. In runs following those where the spectrum in Fig. 1b had been obtained, V_b was successively reduced to 11 and 9 kV and then was raised by steps to 11, 12, 12.5, and 13 kV. In the last case, an interesting spectrum (Fig. 1c) was obtained. It has intense La^{+++} and La^{++} peaks, La-enriched LaB^{+++} and LaB^{++} peaks, and a variety of other peaks of the LaB_n^{+m} type. The La^{+++} peak of the same height was present also at V_b = 12 and 12.5 kV. We, however, show the spectrum in Fig. 1c, because it was recorded for the same $V_b + V_p$ = 13 + 5.6 kV as those in Figs. 1a and 1b.

What happens when V_b grows starting from 8 kV and then varies back and forth? What is the reason that the three spectra for the same object differ so much at the same $V_b + V_p$?

The surface of an as-prepared air-exposed LaB_6 tip (to which the spectrum in Fig. 1a corresponds) is inert and hard. It does not adsorb water (there are no H_2O^+ ions in the spectrum), and the surface atoms almost do not evaporate at a pulse voltage of up to 18.6 kV. Near the surface, only CO^+ ions are generated. From those few ions that are still recorded, one can suggest that the surface is covered by boron and lanthanum oxides. The oxides are broken down at a pulse voltage of as high as 20.6 kV. Water is absent in the spectrum presumably because it is absent at the very end of the tip. At 15 + 5.6 kV, the strong base field V_b = 15 kV removes water at some distance away from the tip. When V_b is diminished by 1 kV, the adsorbed water reaches the tip and, subsequently, the H_2O^+ ion is observed at each of V_b from 14 to 9 kV.

The spectrum in Fig. 1b, like those for V_b = 15 and 14 kV, covers many peaks of the LaB_n^{+m} type, where n for the highest peaks lies between 1 and 5 and m = 2. Such products of field evaporation are typical of the case when the LaB_6 lattice is broken down at room temperature. Spectra like that depicted in Fig. 1b are invariably obtained when excess lanthanum is not delivered to the pure LaB_6 surface subjected to field evaporation.

The spectrum in Fig. 1b is the spectrum of ions generated by n = 422 pulses of N = 5000 pulses applied. Actually, however, the number of ions was greater than n (namely, about 600), since two ions were evaporated simultaneously in some cases.

For the spectrum in Fig. 1c, the number of ions is 1600 (n = 1000, N = 7586); it might be expected that here the number of peaks will be much greater than in Fig. 1b. If the former qualitatively copied the latter, the peaks would be 1600/600 = 2.67 times higher. This is observed, for example, for the B_2^+ peak, which did grow 2.6 times. At the same time, the La^{+++} and La^{++} peaks grew, respectively, seven- and fivefold. These estimates were made from the peak heights. A sharper estimate including ions in the adjacent channels and not only those in the channel of their maximal accumulation (i.e., an estimate of all ions of a given sort that is made from the whole area of the related peak) would most likely aggravate the differences mentioned above. However, such a sharp estimate is difficult to make because of the strong overlap of the peaks at the bottom. The selective growth is not observed for the peaks of LaB_n^{+m}-type ions with n = 1–5 and m = 2 or 3. In the spectrum in Fig. 1c, these peaks grew two or three times, which is not surprising. In Fig. 1c, the CO^+ and H_2O^+ peaks are noticeably weaker than in Fig. 1b because of a decrease in the residual pressure as the mass spectra are recorded.

2. The reason for excess lanthanum in the mass spectra in Fig. 1c. The spectrum in Fig. 1c was taken from the La-enriched surface. For this spectrum, the total content of La and B in the evaporated flow corresponds to the LaB_n formula (where n = 1–3) rather than to stoichiometric LaB_6. Excess lanthanum seems to be present also in the spectrum in Fig. 1b but to a lesser extent: here, possibly, $6 > n > 3$. The reason why the flow of LaB_6 evaporation is La-enriched in the latter case remains to be clarified. It is hardly probable that the LaB_6 single crystals are so nonuniform in composition that the spectra in Figs. 1b and 1c reflect the volume contents of the elements. The elemental composition of the LaB_6 needles was determined by precise gravimetric analysis, and the homogeneity of the needles follows from X-ray diffraction measurements [4].

The ratio $\eta = V_p/(V_b + V_p)$ equals 0.3. It is unlikely that steady-state field evaporation at V_b = 13 kV intensely and selectively removes boron, resulting in an excess of lanthanum (especially with regard for the fact that the previous spectra at V_b = 15 and 14 kV have already been taken). Note that the spectrum was La-enriched even at V_b = 11 kV ($\eta \approx 0.34$). Moreover, the practice of LaB_6 field evaporation suggests that selective removal is typical of lanthanum rather than boron. Hence, other reasons for excess lanthanum in the spectra should be looked for.

Once the oxide film has been removed from the end of the tip (at $V_b + V_p$ = 15 + 5.6 kV), the evaporation proceeds from the pure surface and generates various, including cluster, ions. Being strongly nonequilibrium, the surface thus produced contains a certain amount of La atoms. The field evaporation breaks down the LaB_6 lattice, which generates surface La atoms, among them, those capable of migrating over the surface at room temperature. These atoms are free to move over the surface for some time (that is, they do not evaporate or occupy lattice sites at once). Under the steady-state

conditions, the surface concentration of these atoms depends on field and temperature.

These atoms alone do not produce an excess of lanthanum. Basically, they can raise the local concentration at one site through a decrease in the concentration at another if the field and the temperature provide a considerable free path of these atoms on the surface. The free path must be no less than the diameter of the evaporating area of the tip. This area is roughly three orders of magnitude larger than the probing zone, which is projected onto a diaphragm through which ions to be analyzed escape into the drift space. In our experiments, the probing zone was at the end of the tip. The atomic probe spectrum detects excess lanthanum (exceeding the stoichiometric La : B = 1 : 6 proportion) if an excess concentration of La is produced in the probing zone by depleting the periphery of the evaporating area.

The permanently applied base field V_b produced a gradient of the field strength E at the curved surface. The gradient is directed toward the top of the tip. Under the action of this gradient, the migrating electropositive lanthanum atoms produce a gradient of the La concentration C with a maximum at the very end of the tip [5], i.e., in the probing zone. The field evaporation rate, as well as the rate of electromigration, depends on E, and the flow of evaporating La atoms creates an oppositely directed gradient of C. The balance between the inflow to the tip due to electromigration and the outflow from the tip due to evaporation specifies the resulting surface concentration of La. Here, we may face different situations.

(i) The field evaporation from the tip at a given base voltage exceeds the electromigration inflow to the tip. Then, a boundary where the two fluxes are equal to each other is set at some distance away from the tip. The region adjacent to the tip will be free of excess lanthanum, while the region on the other side of the boundary will be La-enriched. It seems that this situation takes place immediately after the breakdown of the oxide film at V_b = 15 kV and then at V_b = 14 kV. The associated spectrum corresponds to the volume (stoichiometric) composition.

(ii) As V_b is decreased or the tip becomes dull, this boundary approaches the tip. When it enters the probing zone, the spectrum will show excess lanthanum. The boundary may completely disappear (collapse) at the center of the probing zone, and the amount of excess lanthanum in the spectrum will rise with a further decrease in V_b.

(iii) The field evaporation at a base voltage is insignificant. In this case, the pulsed evaporation at $V_b + V_p$ governs the La concentration in (and, generally, beyond) the probing zone.

So, what is the reason for such a large difference between the spectra in Figs. 1b and 1c? The intense evaporation of La at V_b in Fig. 1b seems unlikely. If this were so, the water peak would be absent, since water evaporates at a smaller field than lanthanum. In all probability, the excess lanthanum was spent during the intense field evaporation at V_b = 14 and 15 kV (the previous spectra) and lanthanum had no time to be accumulated during recording of the spectrum in Fig. 1b. When V_b was reduced to 11, 10, and 9 kV, the evaporation of La was insignificant but the lanthanum was substantially redistributed by electromigration to provide subsequent enrichment of the probing zone.

The La concentration in this zone sharply decreased in one of the runs at V_b = 11 kV preceding the spectrum shown in Fig. 1c. For this run, the curve of ion accumulation (Fig. 2) dramatically increases after the application of 2800 pulses (at $V_b + V_p$). Partial spectra for this run before and after the kink in the curve are shown in Fig. 3. They conclusively show that the rise in the accumulation rate is associated largely with the evaporation of the excess lanthanum. In the spectrum before the kink (Fig. 3a), the peaks are weak and scarce. It is similar to the previous spectra at small V_b (10 and 9 kV). The spectrum after the kink is more intense. The peaks of pure lanthanum (La^{+++} and La^{++}), as well as the cluster peaks of the LaB_n^{+m} type (especially LaB^{+++} and LaB_2^{+++}), are substantially higher. Hence, the excess lanthanum is not only removed in the form of monatomic ions but also may capture one or more boron atoms.

As V_b increases to 12, 12.5, and then to 13 kV (the spectrum in Fig. 1c), the lanthanum inflow to the probing zone remains unchanged. In the course of evaporation, the tip might become blunt. Therefore, if the steady-state evaporation of La did not occur at V_b = 13 kV (Fig. 1b), it *a fortiori* could not take place at V_b = 11, 12, or 12.5 kV (Fig. 1c). Note, however, that the appreciable H_2O^+ peak was always present. In all of the cases, the lattice broken down over the evaporating area generated free lanthanum in large amounts and the sufficiently high V_b set the La concentration gradient, which made the probing zone La-enriched. At the same time, the pulse field had no time to counterbalance the electromigration flow while such a tendency was observed: the slope of the ion accumulation curve for the spectrum in Fig. 1c progressively declined.

3. Formation of the oxide film on the LaB_6 surface exposed to a vacuum. From subsections 1 and 2, it follows that the surface conditions of the same tip at the same $V_b + V_p$ may vary: It may be covered by an inert oxide film (the scarce spectrum in Fig. 1a), its surface may have the volume composition once the film has been broken down (various cluster peaks in the spectrum in Fig. 1b), and finally the surface may be La-enriched (the intense many-lined spectrum of field evaporation in Fig. 1c). Note also that, once the excess lanthanum had been accumulated on the top of the tip, spectra qualitatively similar to that in Fig. 1c (perhaps

less intense) were obtained even at $V_b + V_p = 11 + 5.6$ kV (Fig. 3), i.e., at a voltage markedly smaller than 13 + 5.6 kV, when the oxide film almost blocked the evaporation.

Such an inert film forms also upon long-term storage in the residual gases (at a pressure of about 10^{-7} torr), especially if the gases contain moisture. For example, after being exposed to such an environment, the tip (whose properties have been discussed in Sections 1 and 2) showed scarce spectra for V_b between 9 and 14 kV ($V_b = 5.6$ kV). These resembled the spectrum in Fig. 1a with the only difference: the moderate H_2O^+ peak was present even at $V_b = 12$ kV, becoming pronounced at $V_b = 13$ and 14 kV. The associated film resulting from the oxidation of boron and lanthanum on the surface in the residual gases seems to be somewhat thinner than the initial film, grown after etching and exposure in air (Fig. 1a). In spite of its smaller thickness, this film also prevents the evaporation of single-crystal LaB_6. However, it may have a better conductivity and another composition, which causes water adsorption and ionization. This surface film was broken down at $V_b + V_p = 15 + 5.6$ kV with the formation of the developed (many-lined) La-enriched spectrum like in Fig. 1c. Unlike the film associated with Fig. 1a, which was broken down during the rise in V_b from 13 to 15 kV, the film being discussed was broken down during recording the spectrum, as follows from the ion accumulation curve. Before the breakdown, the spectrum contained boron ions and boron oxides, as well as carbon ions and adsorbed CO. After the breakdown, the high La^{+++} peaks were observed not only at 15 + 5.6 kV and 15 + 6.2 kV but also at 13 + 6.2 kV and even 12 + 6.2 kV (which is much less than 14 + 5.6 kV, when the evaporation was blocked by the stable inert film). In all intense spectra taken after the breakdown at $V_b + V_p$ from 12 + 6.2 kV to 16 + 6.2 kV, high La^{+++}, La^{++}, and LaB_n^{+m} peaks were observed along with a more or less appreciable H_2O^+ peak. Thus, it could be concluded that La and LaB_n clusters were not evaporated in considerable amounts at the base voltage; hence, the spectra were valid.

The tip became blunt, and the constant field applied (V_b) was moderate; therefore, excess lanthanum was always present and spectra like in Fig. 1b (without excess lanthanum) were not observed. By selecting V_b not too high so as to prevent La evaporation but sufficient to draw free La into the probing zone, and V_p sufficient for the intense evaporation of La, one can provide conditions where the La^{+++} peak becomes the highest. This is of importance in designing point sources of La ions or La-enriched combined sources of boron and lanthanum.

Upon exposure to the residual gases, the inert film formed in many other cases as well. The film was thinner and easier to break down if the exposure was considerably smaller (hours instead of days) and the experimental conditions were cleaner (the absence of moisture in the setup heated). Carefully increasing V_b from spectrum to spectrum, we were able to detect a breakdown in a spectrum. Then, analyzing the partial spectra before and after the breakdown, we made certain that the former spectra contained boron oxide ions, carbon compounds, and, possibly, lanthanum oxides in greater amounts. The La spectra, however, were difficult to record, since they were superposed on the LaB_n spectra.

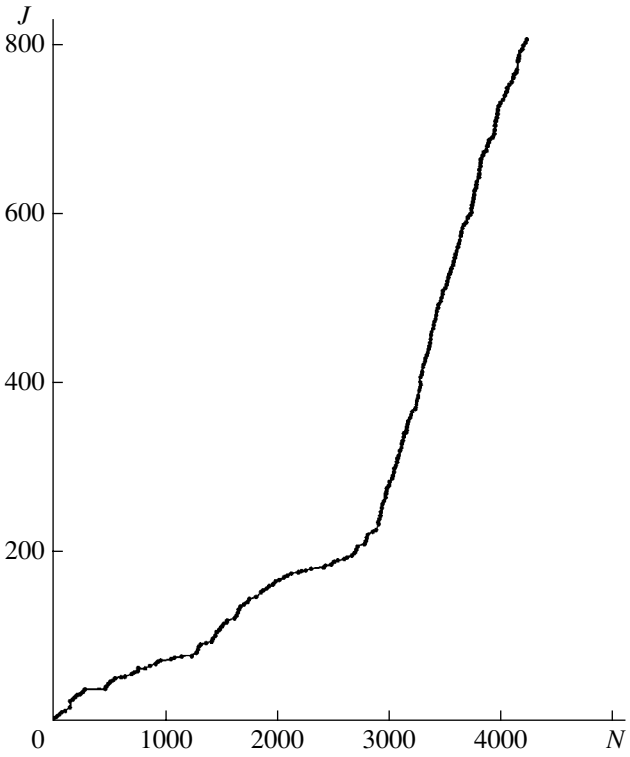


Fig. 2. Ion accumulation curve for the spectrum depicted in Fig. 3. $V_b + V_p = 11 + 5.6$ kV. For the complete spectrum, $n/N = 500/4250$. J is the number of ions collected.

4. Relaxation of the nonequilibrium LaB_6 surface by heating. In another series of experiments, a LaB_6 tip was fixed on a tantalum arc with Aquadag and could be heated by passing the current through the arc. In this series, the intense evaporation at $V_b + V_p = 15 + 6.2$ kV gave a spectrum like that in Fig. 1b with high and broad LaB_3^{+++}, LaB_4^{+++}, LaB_6^{+++}, La^{++}, and LaB^{++} peaks (with a minor amount of excess lanthanum). Then, V_b was decreased to 13 kV. With this voltage, the evaporation intensity was reduced but the spectrum did not change qualitatively. At $V_b + V_p = 13 + 6.2$ kV, the following procedure was repeated several times. The spectrum was obtained at room temperature, then the tip was heated to 1250 K in a high vacuum (10^{-10} torr) in the absence of the field, and subsequently the spectrum was taken again at room temperature. The spectra and the ion accumulation curves before and after the heating

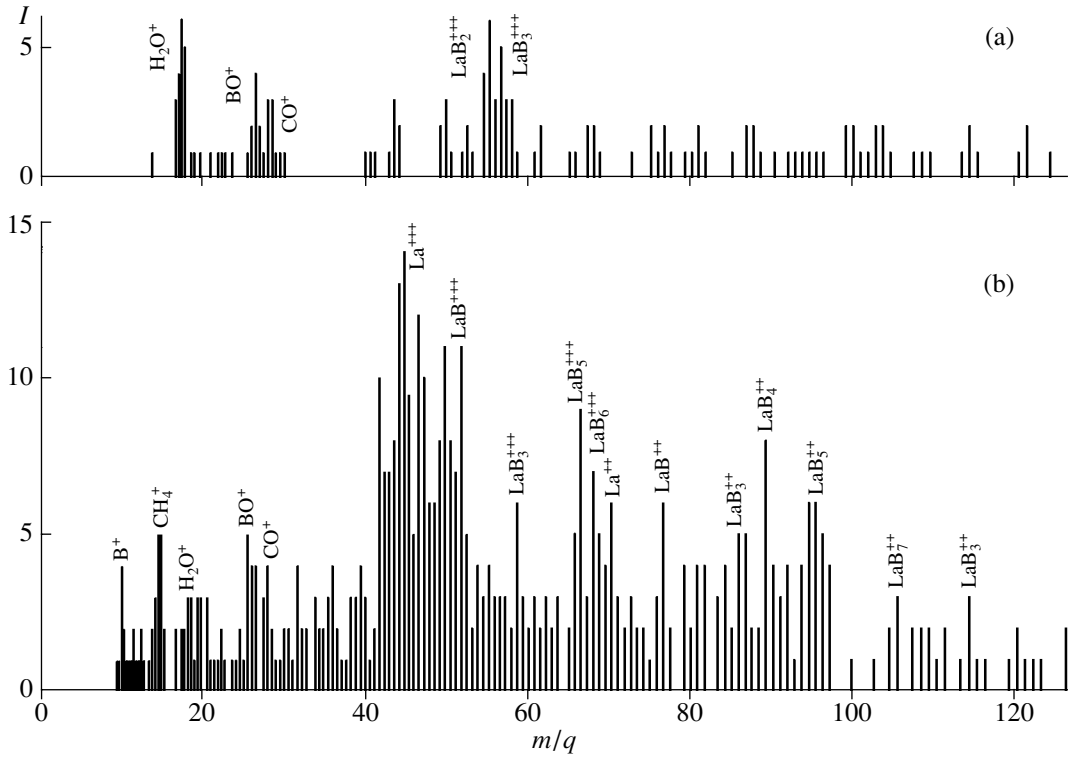


Fig. 3. LaB_6 field evaporation mass spectrum corresponding to the ion accumulation curve in Fig. 2 (a) Before (N = 1–2800) and (b) after (N = 2801–4250) the kink in the curve. For the complete parts of the spectrum, n/N = (a) 220/2800 and (b) 280/1450. V_b + V_p = 11 + 5.6 kV.

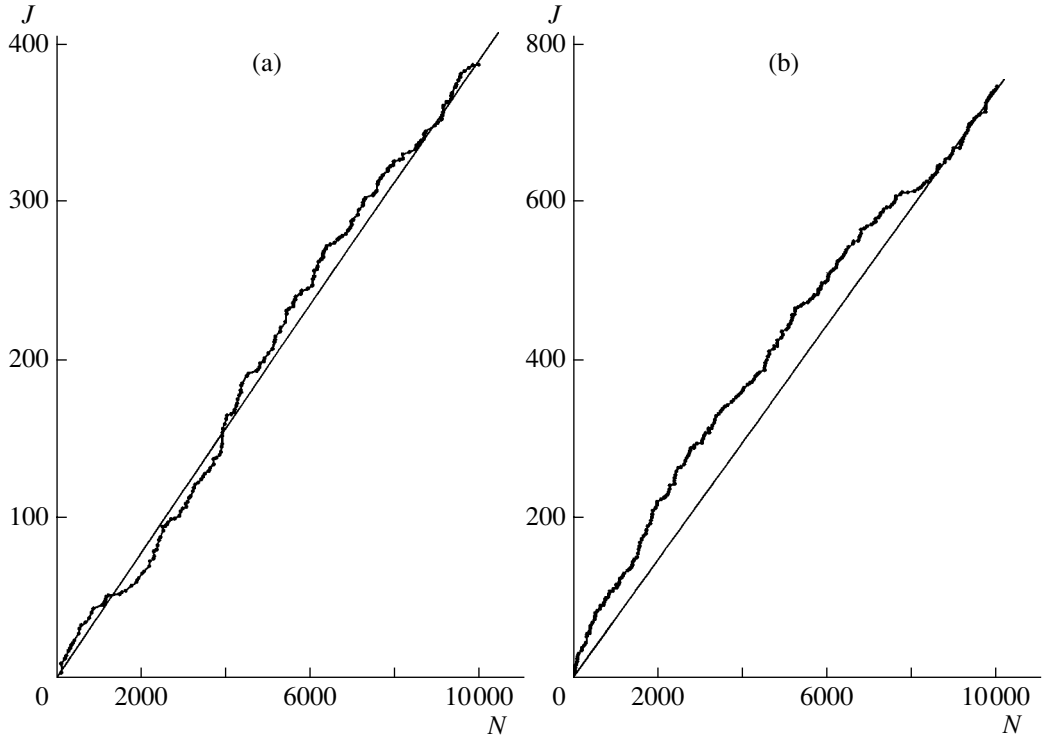


Fig. 4. Ion accumulation curves for the LaB_6 field evaporation spectra obtained at V_b + V_p = 13 + 6.2 kV from the same sample after (a) the intense evaporation at V_b = 13 and 15 kV, V_p = 6.2 kV, n/N = 202/10 000 and (b) heating to 1250 K (in the absence of the field) and subsequent cooling to room temperature, n/N = 379/10 000. Fine straight lines are drawn for comparison.

were compared. The basic difference was that the evaporation intensity increased roughly twice. The spectrum before the heating was obtained at the total number of pulses applied $N = 10\,000$ and the number of ion-generating pulses $n = 185$. For the post-heating spectrum, $n = 315$ at the same N. In another run of this series, $n = 202$ and 379, respectively, for the same N. Qualitatively, the pre- and postheating spectra were identical: all the peaks grew in proportion to n. In these experiments, the tip was pointed at the dark area (as demonstrated by the cold-electron imaging), which had a larger work function. Previously, the spectra with the excess lanthanum were obtained with the tip pointed at the readily emitting area. The peaks for the clusters from LaB_3^{++} to LaB_2^{++} were present in all these spectra. The ion accumulation curves (two of which are shown in Fig. 4) are nearly linear for the field-cleaned surface (Fig. 4a) and tend to decrease their slope with increasing N (Fig. 4b) for the evaporation from the heated and then cooled surface.

The explanation for the effect of enhanced ion collection after heating (with all other things being equal) is straightforward. After intense field evaporation, the surface becomes atomically smooth: all asperities disappear. The remaining structure is very stable against field evaporation. The resulting surface is strongly nonequilibrium and is retained only because it was formed at a sufficiently low (room) temperature. Heating allows the atoms to occupy more equilibrium sites; i.e., the surface relaxes. However, a number of atoms extending outward appear after relaxation. They provide the enhanced ion collection, which was detected in the experiment. Thus, our earlier assumptions that a part of the La atoms can migrate even at room temperature are confirmed by the above experiments: after room-temperature LaB_6 field evaporation, we do obtain the strongly nonequilibrium "frozen" surface.

CONCLUSIONS

(1) The surface of as-etched LaB_6 single crystals kept in a vacuum is covered by an oxide, presumably poorly conducting, film that is highly resistant to field evaporation.

(2) A sufficiently high electric field E breaks down the oxide film; subsequently, the intense field evaporation of LaB_6 crystals can proceed at appreciably smaller E.

(3) According to the relationship between the base, V_b, and pulse, V_p, voltages and also to the area of ion collection, basically two characteristic spectra of LaB_6 field evaporation are observed: without excess lanthanum (this spectrum roughly reflects the volume composition of the crystal) and with excess (as compared to the stoichiometric ratio 1 : 6) lanthanum. The latter spectra exhibit high La^{+++} and/or La^{++} peaks, high cluster LaB^{+++} and LaB^{++} peaks, and lower peaks of other cluster ions like LaB_n^{+m} (they are also typical of the former spectra).

(4) The presence of the excess lanthanum in the mass spectra in Figs. 1c and 3b is explained by its redistribution over the surface due to the constant electric voltage V_b. By properly selecting V_b and V_p, one can provide the field evaporation flows substantially enriched by lanthanum. This might serve as a basis for designing controllable point ion sources.

(5) The inert film like that mentioned in item 1 results during the interaction of the pure LaB_6 surface (after the field evaporation) with the residual gases in the atomic probe chamber. The longer the exposure time and the more H_2O in the residual gases, the more readily the film is produced. At a sufficiently high $V_b + V_p$, the film can be broken down and evaporated. Subsequently, the intense field evaporation of the crystal may take place at voltages lower than the breakdown voltage.

(6) The highly nonequilibrium final surface of LaB_6 relaxes under heating (in the absence of the field) to 1250 K. The relaxed surface evaporates atoms with an intensity twice as much as the initial nonequilibrium surface (all other things being equal).

ACKNOWLEDGMENTS

This work was supported by the program "Surface Atomic Structures" of the Ministry of Science of the Russian Federation (project no. 4.12.99).

REFERENCES

1. I. Bustani, R. Byunker, G. Khirsh, *et al.*, Pis'ma Zh. Tekh. Fiz. **25** (23), 43 (1999) [Tech. Phys. Lett. **25**, 944 (1999)]. (Names as in Tech. Phys. Lett.)
2. I. Boustani, R. Buenker, V. N. Shrednik, *et al.*, J. Chem. Phys. (2001) (in press).
3. M. V. Loginov, O. G. Savel'ev, and V. N. Shrednik, Zh. Tekh. Fiz. **64** (8), 123 (1994) [Tech. Phys. **39**, 811 (1994)].
4. M. M. Korsukova and V. N. Gurin, Usp. Khim. **56**, 3 (1987).
5. A. G. Naumovets, Ukr. Fiz. Zh. **9**, 223 (1964).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1168–1174. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 105–111.

SURFACES, ELECTRON AND ION EMISSION

Surface Photoemission from Ultrathin Potassium Films Adsorbed on Tungsten

M. V. Knat'ko, M. N. Lapushkin, and V. I. Paleev

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

e-mail: lapushkin@ms.ioffe.rssi.ru

Received November 23, 2000

Abstract—The initial stage of formation of ultrathin potassium films on W(100) is studied by threshold photoemission spectroscopy using *p*- and *s*-polarized light in a photon energy range of 1.6–3.5 eV. It is found that the photoemission current spectrum depends on the surface coverage by the alkaline atoms. Mathematically, this shows up as the dependence of the matrix elements responsible for photoemission excitation on surface coverage. The matrix elements vary because the photoelectron escape depth is small; hence, the emission comes from the surface layer under irradiation by both *p*- and *s*-polarized light. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Systems involving alkaline adatoms on metallic substrates have been the subject of much investigation over many decades. These systems are used as a model in studying the interaction of atoms with a solid adsorbent. In addition, alkaline adsorbates are finding wide application in various devices. Recent investigations in this field have been concerned with the generation and modification of surface states (SSs) induced by alkaline adsorbates on metallic adsorbents.

Experimental [1–3] and theoretical [4–7] studies suggest the following scenario of the SS density variation near the Fermi level E_F when an alkaline adsorbate is applied. The band of intrinsic SSs (IntSSs) of the adsorbent lies below E_F. At the initial stage of coating formation, adsorbate–substrate interaction shifts the IntSS band toward higher binding energies and causes the appearance of alkaline-atom-induced SSs (IndSSs). The IndSS band initially lies above E_F. As the surface coverage grows, both IntSS and IndSS bands move toward higher binding energies. For the coverage at which the work function of the adsystem is minimum, the IndSS band goes below E_F. With a further increase in the coverage, both bands experience modifications.

The electron configuration of an alkaline metal–metallic substrate system near its E_F is most conveniently studied by threshold photoemission spectroscopy (TPS). This technique is much more sensitive to surface states located near E_F than conventionally used UV spectroscopy [1, 8].

The TPS technique has been applied to studying Cs/W(100), Cs/W(111), Cs/W(110) [1], and Cs/Ag [9] systems. In all of them, near-E_F SS bands induced by alkaline adsorption and the modification of these bands in a submonolayer range of Cs films have been found. For the Cs/Au system [10], TPS has been applied to study the initial formation stage of the CsAu surface alloy.

The SS spectrum for potassium-on-metal adsorption has been taken for K/Al(111) [3, 11], K/Cu(110) [12], and K/Cu(100) [13] systems. Below, we report TPS results on the formation of the surface electron structure when ultrathin potassium films are applied on a W(100) substrate.

1. EXPERIMENT

Experiments were carried out at a pressure $P \approx 5 \times 10^{-10}$ torr. Single-crystal W(100) was used as a substrate. A total of three potassium monolayers were applied on the substrate at room temperature from an atomically clean potassium source. The coverage ϑ was found from the well-known coverage dependence of the work function φ. The minimal work function of the K/W(100) system was observed at $\vartheta = 0.6$ monolayer [7]. We recorded spectral dependences of the integral photoemission currents $I_p(h\nu)$ and $I_s(h\nu)$ induced by *p*- and *s*-polarized light, respectively, for photon energies between 1.6 and 3.5 eV and various ϑ. The angle of incidence of light was $\gamma = 45°$.

The spectral dependences of the photocurrents $I_p(h\nu)$ and $I_s(h\nu)$ for various K coatings are shown in Figs. 1 and 2. As for the systems Cs/W(100), Cs/W(111), Cs/W(110) [1], and Cs/Ag [9] studied earlier, the $I_p(h\nu)$ and $I_s(h\nu)$ curves differ in shape and intensity. In addition, when ϑ is changed, the $I_p(h\nu)$ and $I_s(h\nu)$ curves vary in a different way. The latter monotonically increase, except for $I_s(h\nu)$ at $\vartheta = 0.6$, which

1063-7842/01/4609-1168$21.00

has a convex form. The shape of the $I_p(h\nu)$ curves substantially changes with ϑ.

Obviously, the differences between the $I_p(h\nu)$ and $I_s(h\nu)$ curves for $\vartheta = 3.0$ and those obtained for "thick" films (100 nm thick) [14] are due to the fact that K films in the former case are extremely too thin. The shape of the $I_s(h\nu)$ curve for thick films is specified largely by the narrow conduction band of K, $E_{\text{band}} = 1.60$ eV [15].

Consider the curves obtained at the different light polarizations in greater detail. The photocurrent $I_s(h\nu)$ near the threshold obeys the Fowler law [16]

$$I_s = a(h\nu - \varphi)^2, \tag{1}$$

where a is a constant.

For metals, the threshold photon energy of photoemission equals the work function: $h\nu_0 = \varphi$. The energy range where the photocurrent obeys the Fowler law depends on ϑ; in other words, in general, the energy range where $I_s(h\nu)$ is a quadratic function varies with ϑ. However, there are two subranges, $\vartheta < 0.3$ and $1.1 < \vartheta < 2.8$, where this curve remains quadratic throughout the excitation energy range.

From $I_s(h\nu)$ curves [formula (1)] taken for various ϑ, one can find the shape of the $\varphi(\vartheta)$ curve (Fig. 3). The value of $\varphi = 2.30 \pm 0.03$ eV at $\vartheta = 3.0$ is somewhat greater than the reference data $\varphi = 2.22$ eV [17].

For *p*-polarized exciting light, the photocurrent near the threshold does not follow the Fowler law.

2. THEORETICAL CONSIDERATION AND DATA PROCESSING

In the theory of threshold photoemission [18], the photocurrents $I_p(h\nu)$ and $I_s(h\nu)$ are given by

$$I_p(h\nu) = |K|^2(h\nu - h\nu_0)^2\theta(h\nu - h\nu_0) \times \frac{|\mathscr{E}_0|^2\sin^2\gamma}{\left|\varepsilon_m\cos\gamma + \sqrt{\varepsilon_m - \sin^2\gamma}\right|^2} \times [|M_1|^2|\varepsilon_m|^2\sin^2\gamma + |M_2|^2|\varepsilon_m - \sin^2\gamma| + 2\sin\gamma\,\mathrm{Re}(M_1^* M_2 \varepsilon_m^* \sqrt{\varepsilon_m - \sin^2\gamma})], \tag{2}$$

$$I_s(h\nu) = |K|^2(h\nu - h\nu_0)^2\theta(h\nu - h\nu_0) \times \frac{|\mathscr{E}_0|^2\cos^2\gamma|M_3|^2}{\left|\cos\gamma + \sqrt{\varepsilon_m - \sin^2\gamma}\right|^2}. \tag{3}$$

Here, ν_0 is the threshold frequency; $\theta(h\nu - h\nu_0)$ is the Heaviside function; ε_m is the permittivity of the metal; $\mathscr{E}_0$ is the amplitude of the incident light wave; M_1 is the matrix element of transitions due to the electric vector component $\mathscr{E}_1$ that is perpendicular to the surface; M_2 and M_3 are the matrix elements of transitions due to the electric vector components $\mathscr{E}_2$ and $\mathscr{E}_3$ that are parallel to the surface; and K is a constant that takes into account the SS density features near E_F and the features of electron transitions at the metal–vacuum interface.

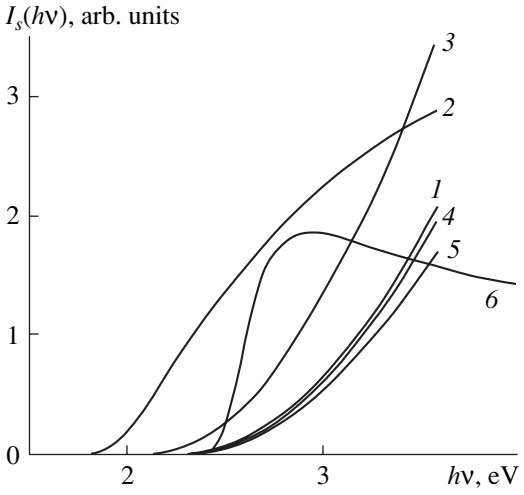


Fig. 1. Spectral dependences of the photocurrent $I_s(h\nu)$ for ϑ = (*1*) 0.30, (*2*) 0.60, (*3*) 1.0, (*4*) 2.0, and (*5*) 3.0. (*6*) Curve for the thick potassium film [14].

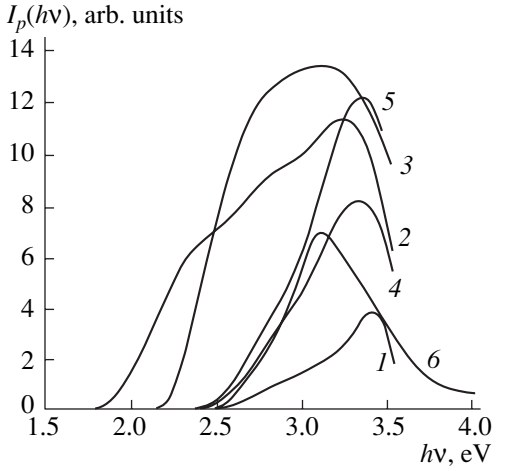


Fig. 2. Spectral dependences of the photocurrent $I_p(h\nu)$. (*1*)–(*6*) the same as in Fig. 1.

Note that the matrix elements M_2 and M_3 are theoretically responsible for photoemission from the interior of the metal, i.e., for bulk photoemission, while the element M_1, for photoemission excited at the metal–vacuum interface, i.e., for surface photoemission. Thus, *s*-polarized light causes only bulk photoemission, while *p*-polarized light must cause both. For isotropic materials, $M_2 = M_3$ [18].

From (3), we have

$$|K|^2|M_3|^2 = \frac{I_s\left|\cos\gamma + \sqrt{\varepsilon_m - \sin^2\gamma}\right|^2}{(h\nu - h\nu_0)^2\theta(h\nu - h\nu_0)|\mathscr{E}_0|^2\cos^2\gamma}. \tag{4}$$

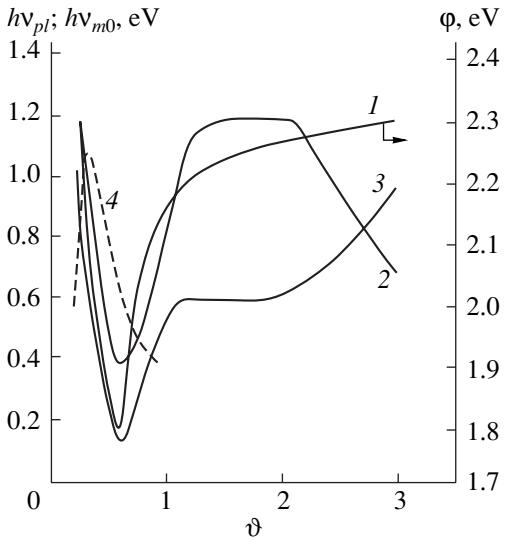


Fig. 3. (*1*) φ, (*2*) $h\nu_{pl}$ and (*3*) $h\nu_{m0}$ vs. ϑ. (*4*) $h\nu_{pl}$ vs. ϑ curve calculated for the Cs/W(100) system according to [1].

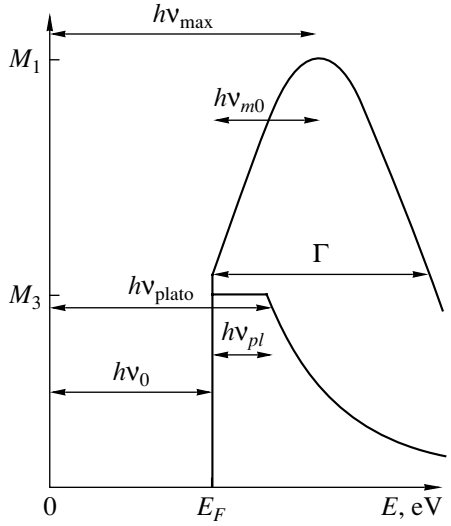


Fig. 4. Matrix elements $|K|^2|M_3|^2$ and $|K|^2|M_1|^2$.

Figure 4 shows the energy dependence of the matrix element $|K|^2|M_3|^2$ that was calculated by formula (4). From (2), the matrix element of surface photoemission cannot be found directly because of the interference term $M_1^* M_2$. In the optical frequency range, this term for metals cannot be ignored unless additional assumptions are made.

Theoretically [18], the elements M_i are represented as

$$M_i = |M_i| e^{i\delta_i},$$

where δ_i is the matrix element phase that is independent of $h\nu$.

Then, the interference term in (2) depends on the phase difference $\delta = \delta_1 - \delta_2$ alone. Hence, the interference term can be rejected when the factor $e^{-i\delta}\varepsilon_m^*\sqrt{\varepsilon_m - \sin^2\gamma}$ approaches zero. This procedure requires special analysis.

To find the value of $|K|^2|M_1|^2$, we must make assumptions on the form of this function. It is selected for each specific system. Parameters that specify the form of $|K|^2|M_1|^2$ are found by a method of minimum search. The minimum of the expression

$$\sum\Bigg(\frac{I_p^{\exp}(h\nu)}{(h\nu - h\nu_0)} - |K|^2\theta(h\nu - h\nu_0) \times \frac{|\mathscr{E}_0|^2\sin^2\gamma}{\left|\varepsilon_m\cos\gamma + \sqrt{\varepsilon_m - \sin^2\gamma}\right|^2} \times [|M_1|^2|\varepsilon_m|^2\sin^2\gamma + |M_2|^2|\varepsilon_m - \sin^2\gamma| + 2\sin\gamma\,\mathrm{Re}(M_1^* M_2\varepsilon_m^*\sqrt{\varepsilon_m - \sin^2\gamma})]^2\Bigg)^2, \quad (5)$$

where $I_p^{\exp}(h\nu)$ are the experimentally found photocurrent values, is sought.

Below, results of processing experimental data by formulas (4) and (5) are presented.

2.1. *s*-polarization. All functions $|K|^2|M_3|^2$ obtained by data processing (Fig. 2) are well described by the equation

$$KM_3 = Km_3(\theta(h\nu - h\nu_0) + \theta(h\nu - h\nu_{plato}) \times (h^n\nu^n h^{-n}\nu_{plato}^{-n} - 1)), \quad (6)$$

where Km_3 is the amplitude of the matrix element (its value is related to the density of states near E_F).

In the range $h\nu_0 \le h\nu \le h\nu_{plato}$, the current $I_s(h\nu)$ obeys the Fowler law; i.e., it is specified by electron states below E_F. For photon energies inducing the photoelectric effect, their density can be considered constant. The exponent n defines the rate of fall or rise of the density of states when photon energies induce photoemission ($h\nu > h\nu_{plato}$).

The values of $h\nu_{plato}$, n, and Km_s were found from the minimum of the expression

$$\sum\Bigg(\frac{I_s^{\exp}(h\nu)}{(h\nu - h\nu_0)^2} - |K|^2\theta(h\nu - h\nu_0) \times \frac{|\mathscr{E}_0|^2\cos^2\gamma|M_3|^2}{\left|\cos\gamma + \sqrt{\varepsilon_m - \sin^2\gamma}\right|^2}\Bigg)^2, \quad (7)$$

where $I_s^{\exp}(h\nu)$ are the experimentally found photocurrents.

Figure 3 plots $h\nu_{pl} = h\nu_{\text{plato}} - h\nu_0$ vs. ϑ.

When the coverage increases to $\vartheta_{\min}$, the plateau starts shrinking. As the potassium film thickens, the value of $h\nu_{pl}$ grows. In the range $1.2 < \vartheta < 2.0$, the plateau reaches its maximum length and then remains unchanged. For $\vartheta > 2.0$, $h\nu_{pl}$ decreases. For $\vartheta = 3.0$, $h\nu_{pl} = 0.85$ eV, which exceeds 0.40 eV for thick potassium layers according to [14]. Note that $h\nu_{pl}$ tends to decline with increasing ϑ.

In the range $\vartheta < 0.3$, $n = 0$. For ϑ between 0.3 and $\vartheta_{\min}$, n drops to –3 at $\vartheta_{\min}$. As the film grows further, n increases to 0 at $\vartheta = 1.0$. In the range $1 < \vartheta < 2.5$, n = const = 0. As the film grows further, n drops to –0.5 at $\vartheta = 3.0$. For thick potassium films, the value of n calculated according to [14] was found to be $n = -11$. This supports the tendency of n to decrease with growing adsorbate thickness.

Figure 5 plots $|K|^2|m_3|^2$ against ϑ. For $0.3 < \vartheta < \vartheta_{\min}$, the matrix element grows and attains a maximum at $\vartheta_{\min}$. At thicker films, $|K|^2|m_3|^2$ decreases. For $\vartheta > 2.0$, the value of $|K|^2|m_3|^2$ remains almost constant. We did not compare our data for $|K|^2|m_3|^2$ with its value for thick potassium layers, because the absolute values of the photocurrents were not measured.

It is seen that the experimentally found values of the parameters $h\nu_{\text{plato}}$, n, and $|K|^2|m_3|^2$ vary with ϑ like $\varphi(\vartheta)$.

2.2. *p*-polarization. For the K/W(100) system, the matrix element KM_1, which is responsible for surface photoemission, is best described by the Gaussian function

$$KM_1 = Km_1(\exp(-\ln(0.5) \times \sqrt{(h\nu - h\nu_{\max}) \div (\Gamma/2(1 + C_{\text{asym}}(h\nu_{\max} - h\nu)/\Gamma))})), \quad (8)$$

where $h\nu_{\max}$ is the position of the peak, Γ is the half-width of the peak, Km_1 is the amplitude of the matrix element, and C_{asym} is the asymmetry factor.

The matrix element M_1 is depicted in Fig. 4. The position of the peak, $h\nu_{\max}$, is slightly shifted toward higher energies relative to the peak of the SS band. In the threshold approximation, the half-width Γ of the peak of the matrix element responsible for the excitation of the surface band is much greater than the width of the SS band [1]. The sign of the factor C_{asym} shows the direction of broadening of the matrix element. At positive C_{asym}, the element broadens toward lower energies. The value of $|K|^2|m_1|^2$ is proportional to the density of states.

The phase difference between the matrix elements M_1 and M_3 is $\delta = 1.0 \pm 0.1$ irrespective of the film thickness. This value is somewhat larger than $\delta = 0.8$, which has been found for the Cs/W system [1].

Figure 3 shows the dependence of $h\nu_{m0}$ on ϑ, where $h\nu_{m0} = h\nu_{\max} - h\nu_0$ is the shift of $h\nu_{\max}$ relative to the photoemission threshold $h\nu_0$.

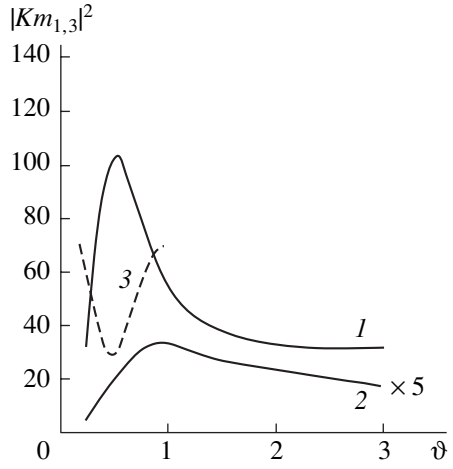


Fig. 5. (*1*) $|K|^2|m_3|^2$ and (*2*) $|K|^2|m_1|^2$ vs. ϑ. (*3*) $|K|^2|m_3|^2$ vs. ϑ curve calculated for the Cs/W(100) system according to [1].

For $0 < \vartheta < 2$, the shape of the curve $h\nu_{m0}(\vartheta)$ closely resembles that of the curve $h\nu_{pl}(\vartheta)$. At small ϑ ($\vartheta <$ 0.6), its increase causes a decrease in $h\nu_{m0}$, which drops to its minimal value at $\vartheta_{\min}$. A further increase in ϑ causes $h\nu_{m0}$ to increase. At $1.0 < \vartheta < 1.8$, $h\nu_{m0}$ remains unchanged. For $\vartheta > 1.8$, $h\nu_{m0}$ grows with ϑ.

The ϑ dependence of $|K|^2|m_1|^2$ is demonstrated in Fig. 5. $|K|^2|m_1|^2$ is much less than $|K|^2|m_3|^2$ throughout the range of potassium film thickness. A similar result has been obtained for photoemission from indium into an electrolyte [18]. Although KM_1 is much smaller than KM_3, the photoemission due to the normal component $\mathscr{E}_1$ of the electric vector of *p*-polarized light is stronger than the photoemission due to the parallel component $\mathscr{E}_2$ of *p*-polarized light. This is because in Eq. (2) for photocurrent $I_p(h\nu)$, the coefficient before M_1 is ε_m times greater than the coefficient before M_2. For tungsten, $\varepsilon_m \approx 20$ in the optical range [19]. The ϑ dependence of $|K|^2|m_1|^2$ is similar to the same dependence of $|K|^2|m_3|^2$. The value of $|K|^2|m_1|^2$ grows to $\vartheta = 0.80$ and then drops nearly to the initial value.

Figure 6 plots Γ against ϑ. Initially, Γ decreases and attains its minimum at the coverage $\vartheta_{\min}$. At a further rise in the coverage, Γ grows insignificantly; for $1.3 < \vartheta < 2.5$, this parameter remains constant. With the coverage increasing still further, Γ increases.

The curve $C_{\text{asym}}(\vartheta)$ is similar to the curve $\Gamma(\vartheta)$ (Fig. 6). The parameter $C_{\text{asym}}(\vartheta)$ is negative, except for the coverage ranges $0.4 < \vartheta$ and $\vartheta > 2.5$.

It is seen that the experimentally found values of the parameters $h\nu_{m0}$, Γ, $|Km_1|^2$, and C_{asym} vary with ϑ like $\varphi(\vartheta)$.

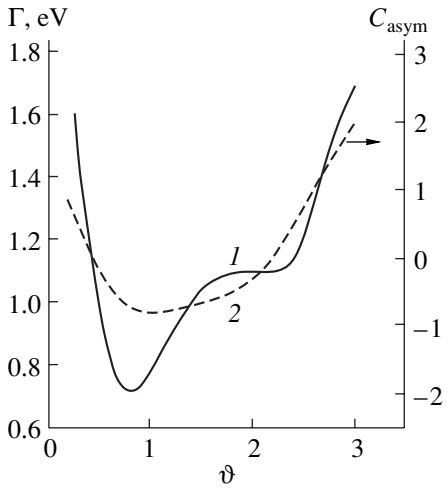


Fig. 6. (*1*) Γ and (*2*) C_{asym} vs. ϑ

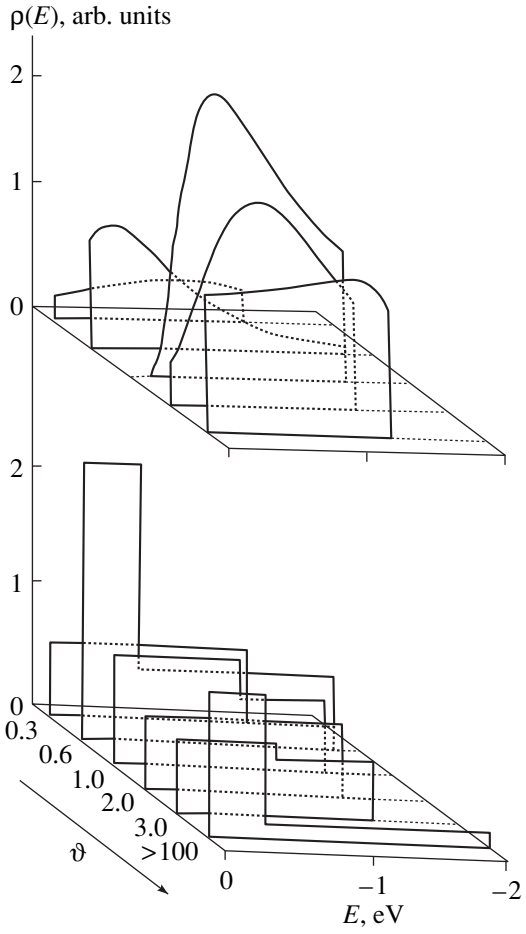


Fig. 7. Variation of the SS band for the K/W(100) system with the coverage ϑ.

3. DISCUSSION

The variations of the photoemission parameters, as well as of the functional shape and the value of the matrix elements M_1 and M_3, with ϑ allow one to qualitatively visualize the variation of the SS density near E_F. The general picture of SS modification is represented in Fig. 7.

3.1. Analysis of the matrix element M_1; SS band due to $\mathscr{E}_1$. Consider the variation of the $SS_\perp$ density (SSs in the direction normal to the surface) near E_F with the degree of coverage (the upper part of Fig. 7). Surface photoemission is known to proceed from the SS band. Ideally, we would have to calculate the matrix element responsible for surface photoemission with regard for the wave functions of the adsorbate and SSs of W, as well as their variation during K adsorption. Such a computational procedure is very tedious. Therefore, we will consider on a qualitative basis a relation between the parameters M_1 and the parameters of the SS band (Fig. 8). For threshold photoemission, where the transition from the surface band to the continuum of excited states takes place, the half-width Γ of the matrix element is several times larger than its associated SS band [1]. The parameter C_{asyn} indicates the SS band asymmetry. The position $h\nu_{m0}$ of the element maximum is 0.1 to 0.2 eV close to the Fermi level than the maximum of the SS band [1]. The position of the surface band maximum E_{max} relative to the Fermi level is given by $E_{max} \approx -h\nu_{m0}$. The value of $|K|^2|m_1|^2$ is proportional to the SS density ρ(*E*). We detect the integral photocurrent and, hence, obtain the smoothed pattern that shades drastic changes in ρ(*E*). The resulting variation of the SS band with increasing degree of coverage for the K/W(100) system is shown at the top of Fig. 7.

At ϑ = 0.3, the SS band is wide, as indicated by large Γ and by the low density of states (the small value of $|K|^2|m_1|^2$). This band peaks at high binding energies ($h\nu_{m0}$ is large). It is logical to assume that this surface band is the IntSS band of tungsten, as for the Cs/W(100) system [1].

As the coverage increases to ϑ_{min}, the SS band experiences great modifications. First, ρ(*E*) grows ($|KM_1|^2$ increases); second, the band shifts toward smaller binding energies (which the decrease in $h\nu_{m0}$ indicates); third, the band shrinks (Γ decreases). These modifications can be explained by shifting the band of potassium valence states below E_F, that is, by the formation of the IndSS band induced by potassium adsorption. Similar behavior has been observed for Cs adsorption on various W faces [1] and for K adsorption on Al(111) [3].

For the monolayer-thick adsorbate, the SS band shifts downward, i.e., toward higher binding energies ($h\nu_{m0}$ grows), and slightly broadens. The density of states increases ($|K|^2|m_1|^2$ rises). The SS band presumably consists of potassium valence electron levels on which the W substrate has a minor effect. Our results correlate with data in [20], where it has been shown that

the SS band width for an isolated K monolayer is ≈1.5 eV and the SS density $\rho(E)$ has no singularities.

The tendencies observed are retained at $1.0 < \vartheta < 2.2$. The surface band consists of potassium valence electron levels and is affected by the W substrate only insignificantly. Hence, the formation of the potassium band of surface states continues. Such a situation is observed to $\vartheta = 3.0$, which the variation of the parameters studied is an indication. The SS band broadens (Γ increases). The density of states in the peak declines ($|K|^2|m_1|^2$ decreases). The band peak shifts toward higher binding energies ($h\nu_{m0}$ increases). Thus, the IntSS band of potassium continues to form.

3.2. Analysis of the matrix element M_3; SS band due to $\mathscr{E}_3$. The theory states [18] that bulk photoemission from a metal is described by the matrix element M_3. Earlier, it was assumed that M_3 must remain unchanged during the application of the three-monolayer-thick potassium films, since the depth of light penetration into the metal is ≈100 nm, so that the escape depth of low-energy photoelectrons is too large (for the general energy dependence of the electron escape depth see [21]). It should be noted here that the escape depth of low-energy (less than 1 eV) photoelectrons from alkaline metals is very difficult to measure. Anyway, we have not found relevant data in the literature. As follows from Section 2.1, the parameters of the element M_3 vary during potassium deposition. This might indicate that the photoelectrons are emitted not from the bulk of tungsten, where $\rho(E)$ and M_3 cannot vary during potassium adsorption, but from a very narrow (several atomic layer thick) surface region. Such an assumption correlates with data in [22], where it has been shown that the escape depth of ≈4-eV photoelectrons from Cs is ≈1 nm. Consequently, one can suppose that photoelectrons come from the surface band that is excited by $\mathscr{E}_3$. In the dipole approximation, this means that the scalar product $((\boldsymbol{\mu}_\parallel \mathscr{E}_3) \neq 0$ ($\boldsymbol{\mu}_\parallel$ is the dipole moment parallel to the surface). This is possible if the $SS_\parallel$ band with the preferential direction parallel to the surface is formed in the surface layer. Such states in photoemission have been detected, for example, for oxygen adsorbed on Ni(100) (p_x and p_y states) [23] (x and y are coordinates in the surface plane). The processing of experimental data for the Cs/W(100) system [1] also indicates that M_3 depends on ϑ (see Figs. 3, 5).

In his classical work [16], Fowler considered photoemission in the flat-band approximation. Our results indicate that this approximation is valid. Below, we will relate the parameters of the element M_3 in formula (6) to the parameters of the band responsible for $\mathscr{E}_3$-induced photoemission. For threshold photoemission from the flat band (Fig. 8) in which the density of states goes to zero at E_{pl}, we have

$$\rho(E) = r(\theta(E - E_F) - \theta(E - E_{pl})),$$

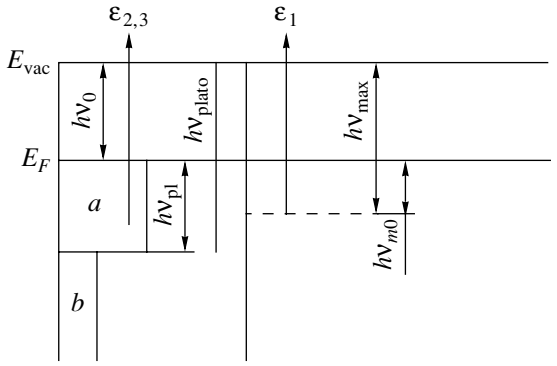


Fig. 8. SS band.

where r is a constant, $E_F = 0$, and $\theta(E - E_{F,pl})$ is the Heaviside function; here, the relationship $E_{pl} = h\nu_{pl}$ holds.

The parameter n shows how much the density of states in the band b is greater (smaller) than in band a (Fig. 8). The parameter $|K|^2|m_3|^2$ is proportional to the density of states in the band. The associated qualitative picture is depicted at the bottom of Fig. 7.

At $\vartheta = 0.3$, there exists the flat band ($h\nu_{pl} = 1.2$ eV) below E_F. This band is formed by the SSs of W(100), which are symmetric in the plane parallel to the surface. The presence of the $SS_\parallel$ on W(100) near E_F is supported by calculation [5].

When the coverage rises to ϑ_{min}, the SS band experiences sharp modifications due to the action of $\mathscr{E}_3$. The value of E_{pl} (or $h\nu_{pl}$) decreases. The density of states in the band grows ($|K|^2|m_3|^2$ increases). At $E < E_{pl}$, the density of states diminished ($n = -3$). The band seems likely to include the SS band of W(100) at binding energies above E_{pl} and the SS band of K near E_F. This result is consistent with the calculation of sodium adsorption on aluminum [4]. It has been shown in [4] that SSs with the preferential direction parallel to the surface exist even at $\vartheta = 0.2$.

When the coverage increases to $\vartheta = 1.0$, the $\mathscr{E}_3$-sensitive SS zone changes again. It becomes flat ($n = 0$), and the density of states drops ($|K|^2|m_3|^2$ decreases). Apparently, the photoemission from the SS zone of W(100) gives way to that from the forming $IndSS_\parallel$ band of potassium.

At $\vartheta = 2.0$, the density of states in the surface $SS_\parallel$ zone somewhat declines (as follows from decreasing $|K|^2|m_3|^2$). This means that the photoemission from W(100) is virtually absent.

At $\vartheta = 3.0$, the surface $SS_\parallel$ band changes once again. We observe the flat band near E_F up to $E > -0.7$ eV ($h\nu_{pl}$ decreases) and the slight decrease in the density of states at high binding energies ($n = -0.5$) The density of states near E_F remains the same ($|K|^2|m_3|^2$ is constant).

The SS band for the thick K films narrows near E_F, as follows from calculations according to [14], and the density of state in it sharply drops at high binding energies.

CONCLUSIONS

We have studied photoemission in the K/W(100) system subjected to visible polarized light at an adsorbate amount of $0.3 < \vartheta < 3.0$.

A method for separating the matrix elements M_1 and M_3, responsible for the photoemission, from the spectral dependences of the photocurrents under *s* and *p* excitations has been suggested. The parameters of the matrix elements have been shown to vary in accordance with the variation of φ.

The variations of the SS density (for both the perpendicular, $SS_\perp$, and the parallel, $SS_\parallel$, components) near and below E_F have been found. The variations of both components characterize the formation of the potassium SS band.

It has been established that, under the excitation by *s*-polarized light, the matrix element responsible for the photoemission changes even at a submonolayer coverage. This means that threshold photoemission is initiated at a very small distance (on the order of several atomic layers) from the surface.

Thus, both the normal, $\mathscr{E}_1$, component of the electric vector of *p*-polarized light and the tangential, $\mathscr{E}_\parallel$ components of *s*- and *p*-polarized light cause photoemission from the SSs of the alkaline metal–metallic substrate adsystem. The former component causes photoemission from the SSs of the adsystem that have the preferential direction normal to the surface, while the latter two cause photoemission from the SSs of the adsystem that have the preferential direction parallel to the surface.

ACKNOWLEDGMENTS

This work was supported by the State Research Program "Surface Atomic Structures" (project no. 3.14.99).

REFERENCES

1. G. V. Benemanskaya, M. N. Lapushkin, and M. I. Urbakh, Zh. Éksp. Teor. Fiz. **102**, 1664 (1992) [Sov. Phys. JETP **75**, 899 (1992)].
2. S. A. Lindgren and L. Wallden, Solid State Commun. **28**, 283 (1978).
3. K. H. Frank, H. J. Sagner, and D. Heskett, Phys. Rev. B **40**, 2767 (1989).
4. H. Ishida, Phys. Rev. B **38**, 8006 (1988).
5. E. Wimmer, A. J. Freeman, J. R. Hiskes, *et al.*, Phys. Rev. B **28**, 3074 (1983).
6. R. Q. Wu, K. L. Chen, D. S. Wang, *et al.*, Phys. Rev. B **38**, 3180 (1988).
7. J. Cousty, R. Riwan, and P. Soukiassian, J. Phys. (Paris) **46**, 1693 (1985).
8. B. Feurbacher and R. F. Wilis, J. Phys. C **9**, 169 (1976).
9. A. Libsch, G. V. Benemanskaya, and M. N. Lapushkin, Surf. Sci. **302**, 303 (1994).
10. M. V. Knat'ko, M. N. Lapushkin, and V. I. Paleev, Pis'ma Zh. Tekh. Fiz. **24** (10), 48 (1998) [Tech. Phys. Lett. **24**, 390 (1998)].
11. K. Horn, A. Hohfield, J. Somers, *et al.*, Phys. Rev. Lett. **61**, 2488 (1988).
12. B. Woratschek, W. Sesselmann, J. Kuppers, *et al.*, Phys. Rev. Lett. **55**, 1231 (1985).
13. T. Aruga, H. Toshihara, and Y. Murata, Phys. Rev. B **34**, 8237 (1986).
14. J. Monin and G. A. Boutry, Phys. Rev. B **9**, 1309 (1974).
15. B. S. Itchkawitz, I. W. Lyo, and E. W. Plummer, Phys. Rev. B **41**, 8075 (1990).
16. R. H. Fowler, Phys. Rev. **1**, 35 (1938).
17. V. S. Fomenko, *Emission Properties of Materials* (Kiev, Naukova Dumka, 1981).
18. A. M. Brodskiĭ and M. I. Urbakh, *Electrodynamics of Metal/Electrolyte Boundary* (Moscow, 1989).
19. J. H. Weaver, C. G. Olson, and D. V. Linch, Phys. Rev. B **12**, 1293 (1975).
20. E. Wimmer, J. Phys. F **13**, 2313 (1983).
21. M. P. Seah, Surf. Interface Anal. **1** (1), 1 (1979).
22. N. V. Smith and G. B. Fisher, Phys. Rev. B **3**, 3662 (1971).
23. G. J. Lapeyre, J. Vac. Sci. Technol. **14**, 384 (1977).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1175–1178. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 112–115.

SURFACES, ELECTRON AND ION EMISSION

Electrode Surface Erosion in a High-Frequency Discharge Plasma

G. P. Strokan'

Research Institute of Physics, Rostov State University, pr. Stachki 194, Rostov-on-Don, 344104 Russia

e-mail: int@ip.rsu.ru

Received January 9, 2001

Abstract—The surface erosion of electrodes made of different materials in the plasma of a high-frequency discharge, which is used for pumping ion lasers at a frequency of 10 MHz, is investigated. It is found that the erosion is due to blistering. The effect of the electrode temperature and material, as well as of the gas type, on the erosion evolution under typical operating conditions of a gas discharge tube is studied. The concentration of blistering products (dust particles) in the discharge is estimated in the framework of geometrical optics. Ways to prevent blistering in the discharge under such conditions are suggested. © *2001 MAIK "Nauka/Interperiodica"*.

INTRODUCTION

The characteristics of laser radiation are greatly affected by the electrode surface condition, since the pumping of transverse high-frequency discharge (THFD) lasers is affected by processes in the electrode region [1–5]. Moreover, during the operation, the electrode surface is exposed to intense ion irradiation with an ion energy comparable or equal to that of near-cathode acceleration. As a result, the cathode surface condition changes and the electrode material is sputtered both in the atomic form and as clusters into the discharge volume. Eventually, the gas discharge characteristics [6–10] and, hence, lasing parameters degrade.

Along with the well-known phenomenon of cathode sputtering [11], the electrodes in a gas discharge are eroded, for example, by blistering, which is well known in the technology of thermonuclear reactors [12]. However, this process in THFD lasers has not been studied. The purpose of this paper is to study blistering in a discharge used for ion laser pumping.

EXPERIMENTAL

Our experimental setup (Fig. 1) was used to observe macroparticles produced in the discharge tube and to investigate their effect on lasing parameters. In the first case, the test tube was placed in a cavity with the tube of an LG-75 helium–neon laser (Fig. 1a). The cavity incorporated mirrors with a reflection coefficient of 99.9%. Because of a high intensity inside the cavity, the birth of particle bunches and their subsequent localization could be observed. With this technique, we were able to observe the ejection of particles 10–30 μm in size. The drawback of this technique is that, as the particle concentration increases, the lasing power decreases and drops to zero at some threshold concentration.

In the second case (Fig. 1b), the test tube was placed outside the cavity and was subjected to a lower-intensity irradiation. In such a configuration, losses up to 100% due to light scattering by dust particles in the tube could be measured. The compensation measuring scheme was used to detect a small absorption (about 0.2%). Using a microscope, we studied the particle motion and measured the particle sizes. Selective mirror *7* prevented the spontaneous radiation of the test tube from reaching a photodetector. Before the measurements, a helium–neon laser was heated for 3–4 h to exclude the influence of intrinsic fluctuations.

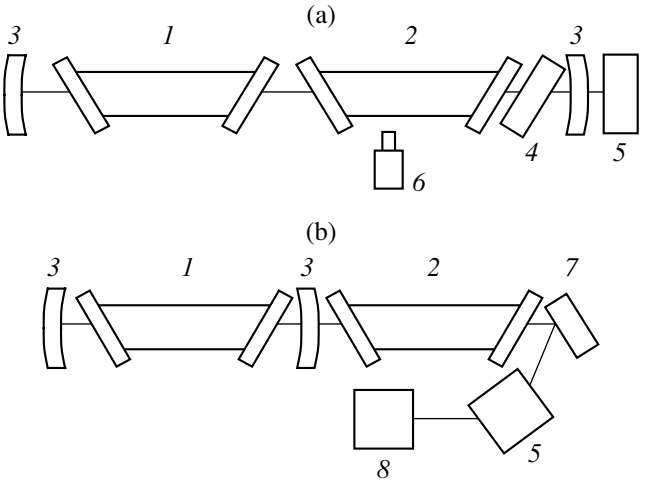


Fig. 1. Block diagram of the experimental setup for observation of blistering products. The test tube is (a) inside and (b) outside the cavity. *1*, LG-75 laser; *2*, test tube; *3*, cavity mirrors; *4*, deflection plate; *5*, power meter; *6*, microscope; *7*, selective mirror; and *8*, compensation signal source.

1063-7842/01/4609-1175$21.00

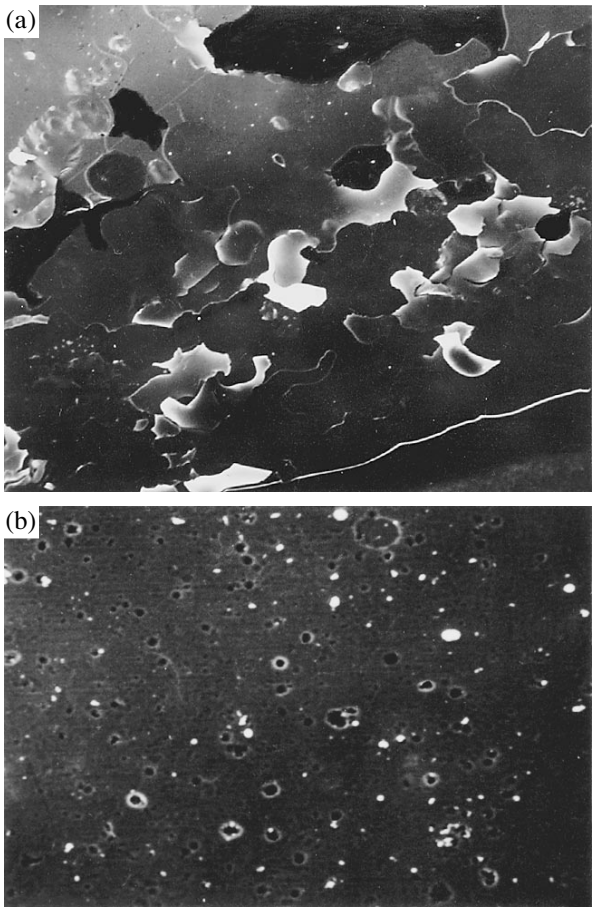


Fig. 2. Image of the stainless steel electrode surface in the case of (a) exfoliation and (b) blistering. Scale is 100 μm.

The test tube with an electrode inside was outgassed by evacuating to a pressure of about 10^{-3} Pa. Then, it was heated to a temperature of 300–400°C by an external resistance or induction-heated and evacuated again. If the current–voltage characteristics remained unchanged after several measurements, the electrode surface was assumed to be clean. It should be noted that the heating of the electrode in the absence of the discharge did not lead to the erosion of the electrode surface as in the case of blistering.

We used metallic (St3 steel, stainless steel, nickel, aluminum), nitride-coated (titanium nitride and chromium nitride), and insulating (quartz and ceramic) electrodes. The thickness of the nitride coatings was about 20 μm. An HF generator with a matched load provided an output of up to 1 kW at a frequency from 1 to 20 MHz.

RESULTS AND DISCUSSION

When studying blistering in the discharge tubes, we took into consideration that this phenomenon depends on temperature, type of the gas, and electrode material. The results obtained are as follows.

(1) Macroscopic particles result both in helium and in helium–cadmium discharges, the concentration of the particles being higher in the helium discharge. In discharges of other gases (air or krypton), the particles are virtually absent.

(2) At 300°C, the concentration of the particles is the highest in the discharge initiated by the stainless-steel and nickel electrodes. For the aluminum and nitride-coated electrodes, the particles are released in small amounts. The quartz and ceramic electrodes practically do not generate the particles.

(3) The temperature dependence of the particle generation was studied in the helium discharge initiated by the stainless steel electrodes. At low temperatures ($T <$ 150°C), dust formation is negligible. At $T = 280$°C, the generation of the macroparticles becomes intense and the particles settle on the walls of the discharge tube. A further increase in the temperature reduced the dust formation rate, which was minimal at $T = 700$°C.

It should be noted that the heating of the electrodes in the absence of the discharge did not cause appreciable powder formation. The electrodes were heated by the external resistance heater, as well as by the induction heater, to about 600°C.

These observations are explained by the properties of blistering described in [12, 13].

(1) Blistering should be most pronounced in a pure-helium discharge, since helium is not bonded to the electrode material and, filling available voids, causes blistering.

(2) The data obtained for the electrodes made of the different materials are explained as follows. When penetrating into the electrode, helium atoms are accumulated in voids, exerting a high pressure in them. As a result, blistering and then electrode erosion occur. If the surface is porous, the helium atoms emerge on the surface through pores without causing electrode erosion. This is seen most clearly in the case of the quartz electrode, which has a helium permeability as high as 7×10^{-3} Pa l/cm^2 at 300°C [14]. This comes into conflict with the Norton rule, according to which metals are impermeable to inert gases and must show blistering. The much lesser generation of the macroparticles in the discharge with the aluminum electrodes is due to the presence of an oxide layer whose thickness exceeds the penetration depth of helium atoms. Moreover, the aluminum electrode was investigated at a temperature of $0.7T_1$ (T_1 is the melting point), while at temperatures of 0.5–$0.6T_1$ or higher, blistering is highly improbable [12]. Similar results were obtained for the electrodes coated by ≈20-μm-thick titanium nitride.

(3) The temperature dependences of the macroparticle generation are in agreement with those obtained in [12, 13] for fusion reactors. The temperature dependences of blistering are in good agreement with those of particle generation that were observed in the helium discharge with the stainless steel electrodes. At temper-

atures less than 150°C, small (about several micrometers in diameter) blisters are produced on the electrode surface by the flow of helium ions. The detachment of small blister caps does not lead to an appreciable dust formation, because the surface erosion rate at this temperature is much smaller (by one or two orders of magnitude) than the rate of exfoliation [13]. At $T = 0.3T_1$, blistering gives way to exfoliation. In this case, the erosion rate and the particle sizes grow. The surface of the stainless steel electrode in the case of blistering and exfoliation is shown in Fig. 2. At an electrode temperature of about 700°C, small blisters appear again on the electrode surface as a result of helium ion bombardment; i.e., exfoliation gives way to blistering. Experimentally, this shows up as a substantial decrease in the particle concentration in the discharge volume.

ABSORPTION OF RADIATION IN LASER TUBES WITH NITRIDE-COATED ELECTRODES

It has been shown above that the nitride-coated electrodes generate a small amount of macroscopic particles. However, the radiation absorption in laser tubes with such electrodes remains to be estimated. In addition, to obtain the maximum lasing power, one should know losses in the tube, for example, for appropriately selecting the electrode material. Therefore, a more detailed study of dust formation and radiation absorption in the laser tube, as well as the determination of the particle concentration due to blistering, would be of much interest.

The measuring setup is shown in Fig. 1b. The measurements were carried out in an asymmetric tube with a 0.5-m-long electrode inside. The active channel has the form of a slot 3×2.5 mm in size. The inner surface of the channel was coated by titanium nitride.

After long-term heating and degassing of the test tube, the absorption coefficient was estimated at $K \approx 25\%$ and remained constant for a long time (Fig. 3). In this case, the discharge volume was free of dust clouds, which are typical of the stainless steel electrodes. The slight absorption can be accounted for by the presence of a small amount of the particles released from the electrode surface.

A decrease in the transmitted light intensity due to particle-related losses in the discharge can be estimated in the geometrical optics approximation:

$$\pi r^2 \alpha n = \frac{\ln(J_0/J)}{l}.$$

Here, n is the average particle density in the discharge; r is the particle diameter; l is the tube length; J_0 and J are the intensities at the tube inlet and outlet, respectively; and α is the damping parameter.

At a particle diameter of 3 μm, which corresponds to the typical particle diameter in Fig. 4 and also to the microscopic visualization parameter $\alpha = 2$ [15], we obtain with our method

$$n = 3.14 \times 10^4 \ln\left(\frac{100}{100 - K}\right).$$

In our case, the particle concentration is about 10^4 cm^{-3}.

Such estimations allow one to compare electrodes made of different materials in order to minimize radiation losses in the discharge tube of a THFD laser, as well as to reduce the noise level in the low-frequency spectral range.

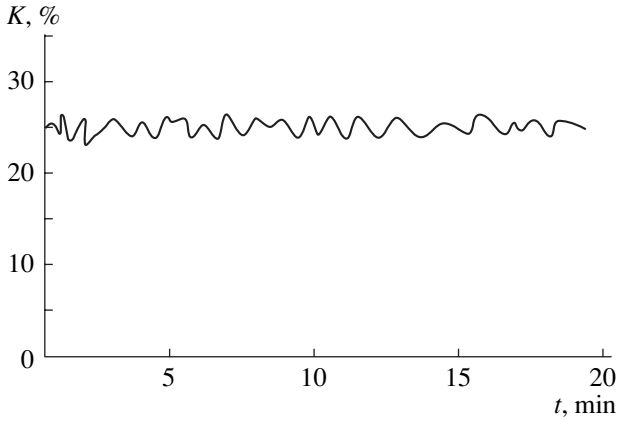


Fig. 3. Time dependence of the radiation absorption coefficient in the tube with the nitride-coated electrode. The electrode temperature is 300°C.

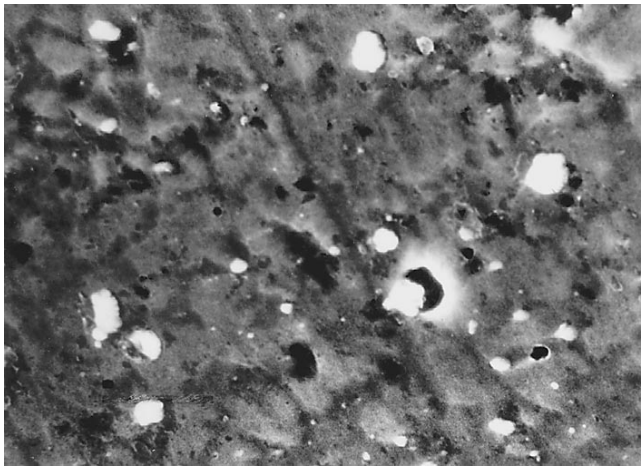


Fig. 4. Surface of the nitride-coated electrode after ten-hour operation under the THFD conditions at a temperature of 300°C. Scale is 10 μm.

CONCLUSION

Thus, in both HF and dc discharges, the surface of the metallic electrodes is eroded. Depending on the electrode material and temperature, erosion may show up as the blistering (the collapse of blisters with diam-

eters from units to tens of micrometers) or exfoliation. In the latter case, the erosion is one or two orders of magnitude faster and the erosion products are several hundreds of micrometers in size. Unlike the stainless steel electrodes, the aluminum and nitride-coated electrodes are not prone to exfoliation during the operation of a helium–cadmium laser, as indicated by the small concentration of dust particles.

To reduce the effect of blistering products on laser performance, we suggest using electrodes coated by nitrides or by a material the melting point of which exceeds the operating temperature by no more than a factor of 1.5.

REFERENCES

1. E. L. Latush, V. S. Mikhalevskiĭ, M. F. Sém, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **24** (2), 81 (1976) [JETP Lett. **24**, 69 (1976)].
2. S. V. Aleksandrov, V. V. Elagin, and A. É. Fotiadi, Pis'ma Zh. Tekh. Fiz. **6** (3), 160 (1980) [Sov. Tech. Phys. Lett. **6**, 70 (1980)].
3. M. G. Dyatlov, V. G. Kas'yan, and V. G. Levin, Pis'ma Zh. Tekh. Fiz. **3**, 644 (1977) [Sov. Tech. Phys. Lett. **3**, 264 (1977)].
4. A. N. Korol'kov, S. A. Rudelev, and V. A. Stepanov, Élektron. Tekh., Ser. 4: Élektrovak. Gazorazryadn. Prib. **9**, 12 (1977).
5. A. A. Kuzovnikov, V. P. Savinov, and V. G. Yakunin, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron. **21** (4), 75 (1980).
6. V. V. Savranskii and G. P. Strokan', J. Russ. Laser Res. **15** (1), 81 (1994).
7. V. S. Mikhalevskiĭ, G. P. Strokan', M. F. Sém, *et al.*, Kvantovaya Élektron. (Moscow) **16**, 37 (1989).
8. G. P. Strokan' and G. N. Tolmachev, Avtometriya, No. 1, 61 (1984).
9. A. V. Borodin, V. F. Kravchenko, and G. P. Strokan', Zh. Tekh. Fiz. **66** (8), 44 (1996) [Tech. Phys. **41**, 770 (1996)].
10. G. P. Strokan' and G. N. Tolmachev, in *TQE-85, Bucharest, 1985*, Vol. I, p. 89.
11. N. V. Pleshivtsev, *Cathode Sputtering* (Atomizdat, Moscow, 1968).
12. M. I. Guseva and Yu. V. Martynenko, Usp. Fiz. Nauk **135**, 671 (1981) [Sov. Phys. Usp. **24**, 996 (1981)].
13. K. L. Wilson, J. Plasma Phys. Thermonuclear Fusion, Special Issue, 85 (1984).
14. *Handbook of Thin-Film Technology*, Ed. by L. I. Maissel and R. Glang (McGraw-Hill, New York, 1970; Sov. Radio, Moscow, 1977), Vol. 1.
15. D. Deirmendjian, *Electromagnetic Scattering on Spherical Polydispersions* (Elsevier, New York, 1969; Mir, Moscow, 1971).

Translated by M. Astrov

Technical Physics, Vol. 46, No. 9, 2001, pp. 1179–1184. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 116–122.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

The Effect of Multiple Electron–Electron Scattering on the Energy Distribution of Electrons in a Correlated Pair

O. M. Artamonov and S. N. Samarin

Research Institute of Physics, St. Petersburg State University, Ul'yanovskaya ul. 1, Petrodvorets, St. Petersburg, 198906 Russia

Received May 30, 2000; in final form, December 26, 2000

Abstract—A unit event of electron–electron scattering in LiF layers is studied by correlation spectroscopy of scattered electrons. The energy distribution of electrons in a correlated pair when a 15- to 55-eV free electron is scattered by a valence electron of LiF is studied. It is shown that single electron–electron scattering prevails and the distribution is uniform when the energy of the primary electron is below 25 eV. As the energy of the primary electron increases, the formation of correlated pairs of electrons with equal energies becomes the most probable. With the energy of the primary electron above 40 eV, the pairs with substantially different electron energies dominate. Such evolution of the energy distribution of the electrons in the pair stems from the fact that first one and then the other electron of the pair successively takes part in electron–electron scattering. A phenomenological model for the single scattering and double scattering of primary electrons in LiF films is considered. Results obtained indicate that the strengths of single scattering and double scattering channels become comparable at electron energies above 25 eV. *© 2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Correlation spectroscopy of scattered electrons, or spectroscopy of electron–electron coincidences, is among the most promising methods for studying electron–electron scattering on and near the surface of solids. In correlation spectroscopy, a pair of correlated electrons that results from a unit event of scattering of a primary electron by a valence electron of the solid is detected. The idea of using correlation spectroscopy of secondary electrons for studying a solid surface was first suggested by us in [1]. Correlation spectroscopy of low-energy backscattered electrons was first applied to metal surfaces by a joint team of Russian and German researchers [2–7]. In recent works, the dynamics of slow electron surface scattering in the coincidence spectroscopy method has been treated theoretically [6, 8–10].

The energy distribution of electrons in a correlated pair is a basic parameter of electron–electron scattering. In the free-electron binary scattering approximation, this distribution reflects the dynamics of electron–electron scattering and is adequately described by the well-known Mott formula, which is the quantum-mechanical generalization of the Rutherford formula (see, e.g., [11]). The probability of generating a correlated pair as a function of the energy of one of the electrons has a minimum when the electrons have equal energies (this is valid if their spins are not taken into account, i.e., when the correlated electrons are undistinguishable). When a primary (especially low-energy) electron is scattered by a valence electron of the solid, the energy distribution of correlated electrons depends on many factors: (1) multiple electron–electron scattering in the solid, (2) the escape of the correlated pair from the solid to a vacuum, (3) the nonuniform distribution of vacant and occupied electron states in the solid, (4) the presence of surface electron states, (5) electron diffraction by the lattice, etc. It has been shown experimentally [2–7] that the energy distribution of electrons in a correlated pair may have both a minimum and a maximum at equal energies of scattered electrons when primary electrons of energy between 10 and 50 eV are scattered by the metal surface.

In this work, we study the effect of multiple electron–electron collisions on the energy distribution of the correlated carriers. Experiments were performed with insulating LiF films. Insulators feature a well-defined energy threshold for the scattering of a nonequilibrium electron by a valence one. This threshold equals the bandgap E_g (11.5–13 eV for LiF). Thus, we can easily determine the energy threshold for the band-to-band excitation of valence electrons [12, 13].

EXPERIMENT

1. Equipment. Experiments were performed with an (e, $2e$) electron spectrometer configured with two time-of-flight energy analyzers (Fig. 1) at a pressure on the order of 10^{-11} torr. A W(100) substrate was mounted on a rotatable support and was current-heated to

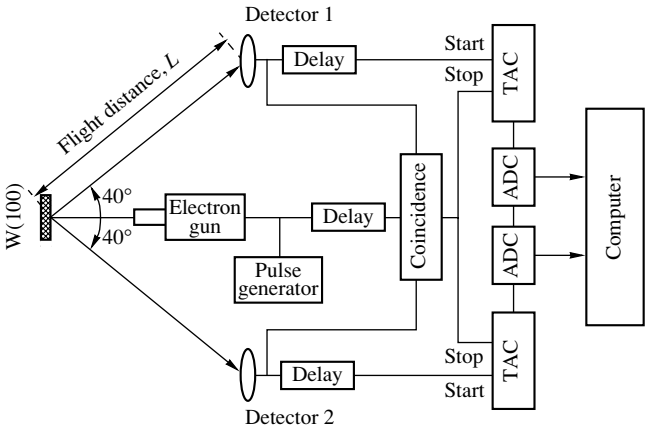


Fig. 1. (*e*, 2*e*) time-of-flight spectrometer.

2000°C at regular intervals for cleaning. The purity of the substrate was checked by Auger spectroscopy. LiF films were applied on the W(100) substrate by thermal evaporation from an electron-beam-heated molybdenum crucible. The thickness of the films was determined with a quartz microanalytical balance. An electron gun generated a pulsed electron beam with a mean current on the order of 10^{-14} A and a diameter of less than 3 mm. Two microchannel plates (MCPs) of diameter 30 mm were used as detectors. The detectors and the electron gun were arranged in the plane of the normal to the surface. The angular coordinates Θ_1 and Θ_2 of the detectors could be varied between 30° and 80° relative to the normal to the surface. Measurements were taken for $\Theta_1 = \Theta_2 = 50°$. The magnetic field of the Earth was attenuated approximately 100-fold by means of Helmholtz coils and a shield made of mu-metal, the shield being placed inside the vacuum chamber.

The energy of scattered electrons was measured from the time of flight of the electrons from the sample to the detector. The incident electron beam was modulated by pulses with a repetition rate of 2.5×10^6 pulse/s at a pulse width of no more than 5 ns. A modulating pulse from a pulse generator was used as the zero time of the time-of-flight scale and triggered two time-to-amplitude converters, TAC1 and TAC2. A correlated pair was detected as follows. An incident electron is scattered by a valence one, and both scattered electrons leave the surface. They are detected separately by the first MCP (MCP1) and the second MCP (MCP2) and generate pulses that shut down the TAC1 and the TAC2, respectively. A coincidence circuit (threshold generator) rejected all pulses that do not coincide in time (within 200 ns). The times of flight $T1$ and $T2$ are digitized with two ADCs and memorized. The result is a 2D distribution of the amount of electron pairs over the times of flight of both electrons of the pair.

Electron detection in the coincidence mode allows us to separate those electron pairs originating from inelastic scattering of one primary electron. The separation of energy-correlated electron pairs is accomplished by analyzing experimentally found distributions. For example, if the target is metallic and only those pairs of scattered electrons for which the total energy equals the energy of the primary electron minus the work function of the target are separated, it becomes clear that these pairs arose when the primary particles was scattered by valence electrons on the Fermi level.

2. Experimental results. Experimental 2D time-of-flight distributions were converted to 2D energy distributions of correlated electrons. Figure 2 shows three 2D energy distribution of the correlated electrons for incident electron energies of 23, 35, and 45 eV. The energy of the electron pair is plotted on the coordinate axes. The degree of blackening depends on the amount of pairs detected (the stronger the blackening, the larger the amount of the pairs). Any line parallel to the dashed line corresponds to pairs with a constant total energy of correlated electrons $E_{tot} = E_1 + E_2$ (where E_1 and E_2 are the energies of the first and the second electron, respectively) or to the constant binding energy of the valence electron E_b. The peaks of the 2D distributions lie on the dashed lines and correspond to certain values of the total energy of the pair. In Fig. 2a, the peak is at E_{tot} = 12 eV for E_p = 23 eV (E_p is the energy of a primary electron), which corresponds to the binding energy of the valence electron E_b = 11 eV excited from the valence band top. In Figs. 2b (E_{tot} = 22 eV at E_p = 35 eV) and 2c (E_{tot} = 29 eV at E_p = 45 eV), the binding energy is greater; here, the pairs are most likely to be excited from the center of the LiF filled gap.

Sections of the 2D distribution that are parallel to the dashed line in Fig. 2 yield energy distributions between the electrons in a pair at constant E_{tot}. Figure 3 depicts the energy distribution between the electrons of a pair for several E_p and E_{tot}. The selected values of E_{tot} correspond to the valence electron excited from the valence band top. The intensity axis shows the relative amount of correlated pairs with the electron energies E_1 and $E_2 = E_{tot} - E_1$. The energy distributions in Fig. 3 correspond to different energies of primary electrons and to the same energy of the valence electron. The values of E_p and the total energy E_{tot} are indicated by arrows. The spacing between E_p and E_{tot} in the energy axis gives the binding energy of the scattering valence electron (Fig. 3c).

All the distributions in Fig. 3 are symmetric about the midpoint $E_1 = E_2$, since the geometry of the experiment is symmetric and the detectors are undistinguishable for the scattered electrons. The energy of each of the electrons can vary between 0 and E_{tot}. However, the range of the energies measured is somewhat narrower, because only electrons of energies above 1.5–3.0 eV were detected (this range depends on the time interval that is measured by the time-of-flight analyzers). The distributions in Figs. 3a and 3b are nearly uniform within the statistical spread; that is, the probability of

detecting pairs with different combinations of the electron energies is constant within the available energy interval. The distribution in Fig. 3c is sharper, while that in Fig. 3d has distinct steps on both sides of the center. The height of the steps is roughly 0.5. In Figs. 3e and 3f, the distributions have a minimum at the center. In Fig. 3g, this minimum occupies the whole central region. The intensity in the minimum is 0.3. Later, we will show that the steps and the minima can be related to double scattering.

DISCUSSION

1. Model of electron–electron scattering on the surface. The elastic scattering of a primary electron by the ion core plays an essential part in the scattering kinematics and turns the total moment of a correlated pair toward a vacuum. Elastic scattering will be included in the scattering model under consideration. To be definite, we assume that elastic scattering precedes inelastic electron–electron scattering (the left of Fig. 4a), although the assumption that a primary electron interacts simultaneously with the core and with a valence electron would be more plausible. The energy balance for single electron–electron scattering is shown in the band diagram on the right of Fig. 4a. The lowest energy of the primary electron that is needed for the excitation of a correlated pair into a vacuum, E'_{th}, can be estimated from the law of conservation of energy. If the electron affinity of a material is close to zero (that is, the conduction band bottom roughly coincides with the position of the vacuum level), then $E'_{th} \cong E_g$, where E_g is the bandgap of the insulator. In this case, the scattered primary and valence electrons pass to the conduction band bottom and find themselves in a vacuum with the zero energy. With the potential barrier at the surface, E'_{th} grows. Its value also becomes larger than that estimated from the energy balance for momentum-correlated electrons.

The second event of electron–electron scattering may take place if the energy of at least one of the correlated electrons exceeds E'_{th}. This is possible when $E_p > E''_{th}$, where $E''_{th} \cong 2E'_{th}$ is the threshold of double scattering. With such an energy of the primary electrons, at least one of the correlated electrons has an energy that is sufficient for the excitation of an electron from the valence band. In this case, the initial correlated pair collapses. One would expect that the collapse is responsible for the symmetric steps on both sides of the major peak in Fig. 3d, where the energy of the primary electrons is 35 eV.

As the energy of the primary electron increases further, it may so happen that the energies of both electrons in correlated pairs exceed $2E'_{th}$; hence, both may take part in double scattering. The double scattering threshold for both electrons is $E'''_{th} \cong 3E'_{th}$. In this situation, the number of pairs with $(E_1, E_2) > E'_{th}$ decreases further. Initially, at $E_p \cong E''_{th}$, the number of the pairs

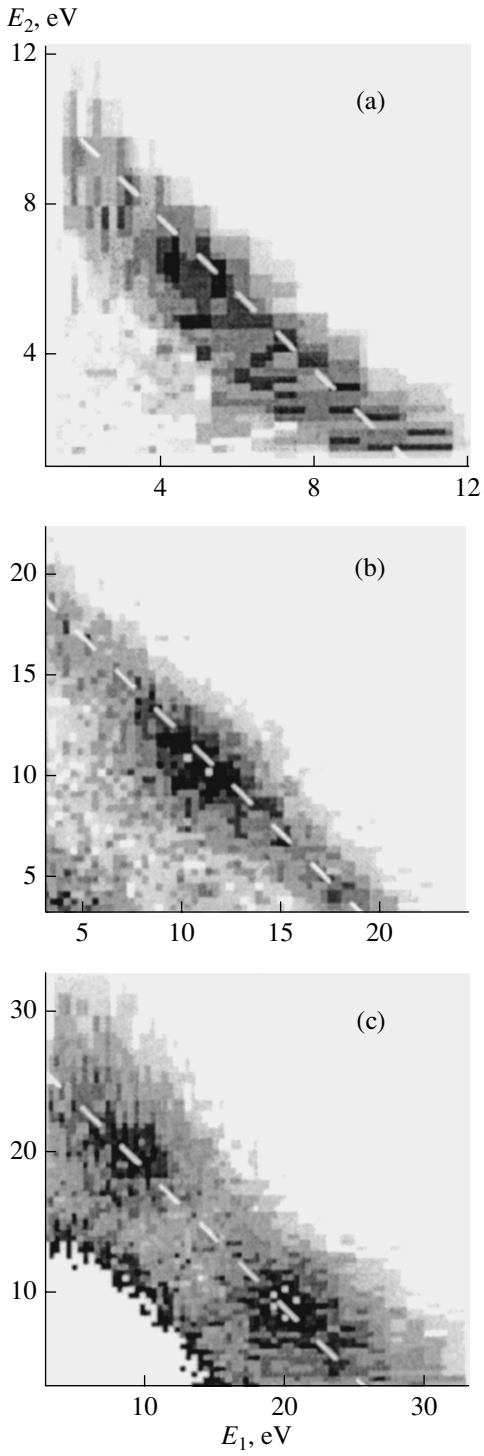


Fig. 2. 2D energy distribution of the correlated electrons that were scattered by the LiF/W(100) layer. $\Theta_1 = \Theta_2 = 45°$. The correlated pairs are near the dashed lines.

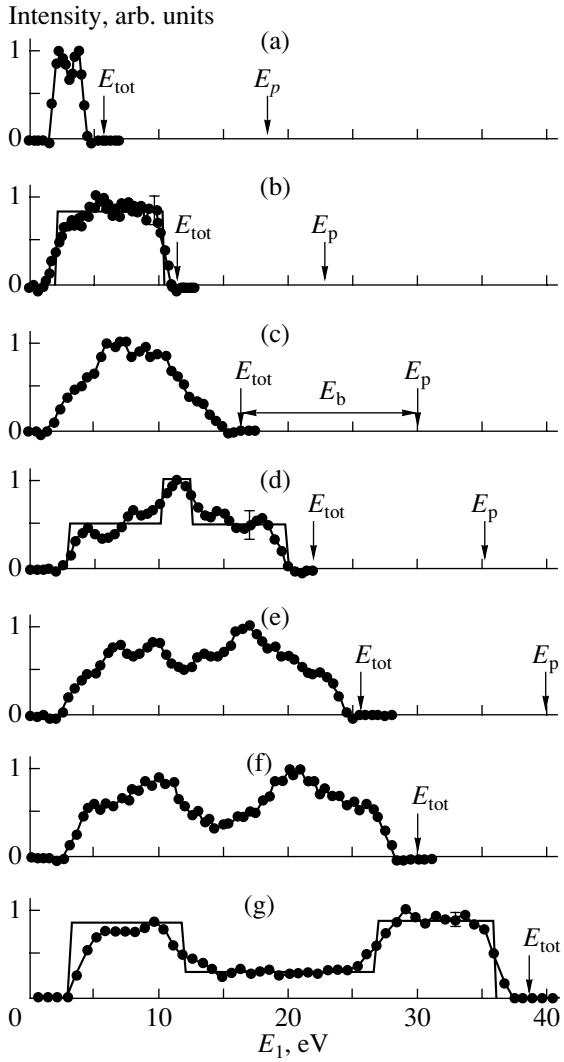


Fig. 3. Energy distributions of the correlated electrons with the total energy E_{tot} of the pair for E_p = (a) 18, (b) 23, (c) 30, (d) 35, (e) 40, (f) 45, and (g) 55 eV. E_{tot} = (a) 6, (b) 12, (c) 16, (d) 22, (e) 26, (f) 30, and (g) 39 eV. The energy of primary electrons and the total energy of the pair are indicated by arrows. Continuous lines in (b), (d), and (g) are model results.

starts decreasing from the center of the energy distribution; as the energy of the primary electron increases, this process spans the successively wider central region (Fig. 3g).

2. The effect of double scattering on the energy distribution of correlated pairs. To estimate the effect of multiple electron–electron scattering on the energy distribution of correlated electrons, one should make an assumption of this distribution without multiple scattering. The distributions in Figs. 3a and 3b refer to single electron–electron scattering, since the energy of the electrons in the pair is below E'_{th}. Within the statistical spread of data points, both distributions can be approximated by a constant value of the intensity in the energy range considered. Energies falling into this range are smaller than E_p, since the range of the measured electron energy is bounded from below by a value varying from 1.5 to 3.0 eV according to the energy of the primary electron. With these assumptions, a function that describes the energy distribution between the electrons in the pair can be represented as a plateau with steeply falling "side walls" at E_1 = 3 eV and $E_2 = E_{tot} - 3$ eV (Fig. 3b). This function can be extrapolated to the entire energy range of the primary electron. Actually, however, the shape of the side walls is of minor significance for reasoning that follows.

Consider the effect of inelastic electron–electron scattering on the energy distribution in a correlated pair immediately after its generation in a solid. We assume that an electron of the pair may either escape into a vacuum (and, hence, be detected as an electron of the pair) or take part in additional (second) electron–electron scattering, which results in the collapse of the pair. One can suppose that the energy dependence of the inelastic scattering probability, $\omega_{ee}(E_1)$, can roughly be approximated by the step function

$$\omega_{ee}(E_1 > E_g) = \omega_{ee}, \quad \omega_{ee}(E_1 \le E_g) = 0, \qquad (1)$$

where E_g is the threshold.

Function (1) characterizes the probability that an electron of a correlated pair loses its energy. If one of the electrons loses its energy, the pair disappears.

The probability $P_0(E_1,E_2)$ of detecting a pair of electrons is the product of the probabilities of detecting either electron, $P_1(E_1)$ and $P_2(E_2)$; that is,

$$P_0 = P_1(E_1)P_2(E_2) = P_1(E_1)P_2(E_{tot} - E_1). \qquad (2)$$

On the other hand,

$$\begin{aligned} P_1(E_1) &= A[1 - \omega_{ee}(E_1)], \\ P_2(E_{tot} - E_1) &= A[1 - \omega_{ee}(E_{tot} - E_1)], \end{aligned} \qquad (3)$$

where A is a constant specific of the given model.

Eventually,

$$P_0(E_1, E_{tot}) = A^2[1 - \omega_{ee}(E_1)][1 - \omega_{ee}(E_{tot} - E_1)]. \qquad (4)$$

Expression (4) is the probability of detecting a correlated pair of electrons where one of them has an energy E and the other, $E_{tot} - E_1$. If the initial energy distribution of the correlated electrons is assumed to be time-invariable, then function (4) describes the energy distribution of the electrons in a pair after some correlated pairs have disappeared as a result of inelastic electron–electron scattering.

3. Comparison of the model with experimental data. To compare experimental distributions with the model ones, we normalized the former to their maxima. As adjustable parameters, we used E_g (its value affects the position of the steps in the model curves) and ω_{ee} (the probability of electron–electron scattering affects the relative height of the steps). The best agreement was

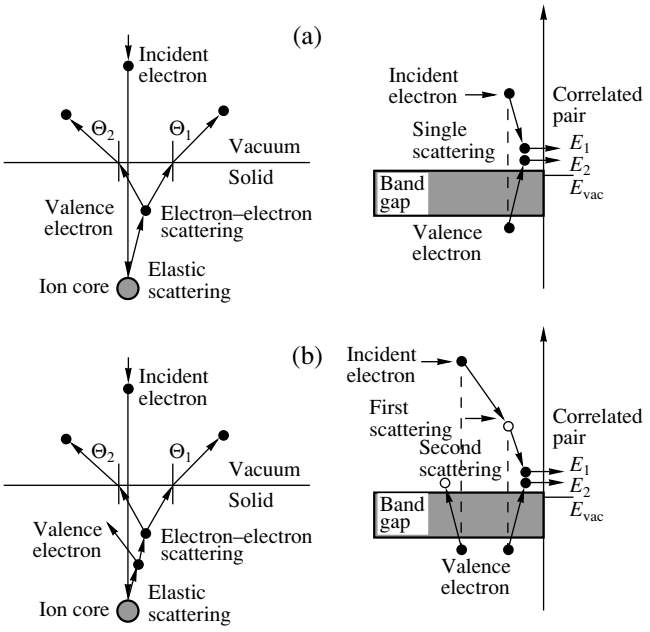


Fig. 4. Electron scattering model: (a) single and (b) double scattering. The left-hand and right-hand sides of the figure depict scattering geometry and band diagram, respectively.

found for E_g = 12.5 eV (for E_p = 35 eV) and 12 eV (for E_p = 55 eV) and ω_{ee} = 0.5 and 0.65, respectively (continuous lines in Figs. 3d, 3g). It should be noted that these values of E_g are close to the literature data [12, 13].

The results of fitting the model to the experimental data are in good agreement with the above qualitative explanation of the changes that are observed in the (*e*, 2*e*) spectrum as E_p grows. When the energy of the incident electrons rises, so does the total energy of the pairs at the ridge of the 2D distribution. For a given E_{tot}, the energy E_1 of one of the electrons takes the highest value if the energy E_2 of the other electron is the lowest and vice versa, since $E_2 = E_{tot} - E_1$. Pairs with such combinations of the electron energies lie at both edges of the function of energy distribution among the electrons. These pairs disappear as soon as the energy of one of the electrons exceeds the excitation threshold of the valence electron. In this case, the steps at the edges of the distribution function arise (Fig. 3d). When the energies of both electrons (E_1 and $E_2 = E_{tot} - E_1$) exceed E_g, the dip at the center of the distribution appears (Fig. 3g). This dip is due to the fact that both electrons have an energy that is sufficient for the excitation of the valence electron through the energy gap. The probability of this pair disappearing from the distribution may increase twice compared with the case when only one electron has a high energy. It is worth noting that the energy distribution functions for electrons in a pair have the fine structure, which is most pronounced in Fig. 3d. A reason for this structure may be that the dependence $\omega_{ee}(E)$ has a more complex form than a unit step, which was used in the calculation. Also, the fine structure may be attributed to the dynamics of scattering of a primary electron by a valence one.

CONCLUSION

It has been shown that, when the energy of incident electrons is below 26 eV, correlated pairs in LiF are excited with the highest probability via single electron–electron collisions. Thus, for such energies, the contribution of multiple collisions is negligible. For single electron–electron scattering, we can estimate the binding energy of the valence electron involved in the process. Accordingly, it has been demonstrated that the valence electrons are excited largely from the valence band top. At energies between 26 and 40 eV, one of the electrons in the pair has an energy sufficient for the valence electron to be excited. At energies above 40 eV, both correlated electrons acquire such a capability. Because of the additional scattering of the correlated electrons, the number of pairs generated by single scattering events decreases. A comparison between the model and the experimental data indicates that the probability of a correlated pair escaping from a solid and the probability that the electrons of the pair will take part in further electron–electron scattering events may equal to each other. In the latter case, the function of energy distribution between the correlated electrons may change. The modified function reflects both the primary scattering of the primary electron by the valence one and subsequent scattering of the correlated electrons in the resulting pair.

The method of (*e*, 2*e*) spectroscopy turns out to be an efficient tool for studying the scattering of slow electrons by insulating surfaces. It is believed that double (or multiple) scattering of incident electrons is also essential for semiconductors and metals. Thus, multiple scattering plays a major part in (*e*, 2*e*) spectroscopy of solid surface.

ACKNOWLEDGMENTS

This work was supported by the State Program “Integration” [project no. A 0151 (0326.37)] and by the Russian Foundation for Basic Research (project no. 99-02-16769).

The authors thank the administration of the Max Planck Institute (Halle, Germany) for permission to publish experimental results obtained at the institute.

REFERENCES

1. O. M. Artamonov, Zh. Tekh. Fiz. **55**, 1190 (1985) [Sov. Phys. Tech. Phys. **30**, 681 (1985)].
2. J. Kirschner, O. M. Artamonov, and S. N. Samarin, Phys. Rev. Lett. **75**, 2424 (1995).

3. O. M. Artamonov, S. N. Samarin, and J. Kirschner, Phys. Rev. B **51**, 2491 (1995).
4. O. M. Artamonov, S. N. Samarin, and J. Kirschner, Appl. Phys. A: Mater. Sci. Process. **A65**, 535 (1997).
5. S. Samarin, G. Herrmann, H. Schwabe, and O. Artamonov, J. Electron Spectrosc. Relat. Phenom. **96**, 61 (1998).
6. R. Feder, H. Gollisch, D. Meinert, *et al.*, Phys. Rev. B **58**, 16 418 (1998).
7. S. Samarin, J. Berakdar, R. Herrmann, *et al.*, J. Phys. IV **9** (Pr6), 137 (1999).
8. H. Gollisch, D. Meinert, Yi. Xiao, and R. Feder, Solid State Commun. **102**, 317 (1997).
9. J. Berakdar and M. P. Das, Phys. Rev. A **56**, 1403 (1997).
10. J. Berakdar, S. Samarin, H. Herrmann, and J. Kirschner, Phys. Rev. Lett. **81**, 3535 (1998).
11. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Fizmatgiz, Moscow, 1963; Pergamon, New York, 1977).
12. S. C. Erwin and C. C. Lin, J. Phys. C **21**, 4285 (1988).
13. W. Pong, D. Pondyal, and D. Brandt, J. Electron Spectrosc. Relat. Phenom. **21**, 261 (1980).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1185–1189. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 123–127.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

High Electrical Conductivity of Trotyl Detonation Products

S. D. Gilev and A. M. Trubachev

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

e-mail: Gilev@hydro.nsc.ru

Received January 9, 2001

Abstract—A new measurement scheme makes it possible to study the conductivity of detonation products of condensed explosives with a time resolution of about 10 ns. Experiments with cast trotyl show that conduction under detonation is a complex phenomenon associated with the chemical reaction zone and the expansion of the reaction products. The time variation of the electrical conductivity has a sharp peak ($\approx 250\ \Omega^{-1}\ cm^{-1}$) and a plateau ($\approx 35\ \Omega^{-1}\ cm^{-1}$). The peak corresponds to the highest conductivity value that has been ever observed for the products of chemical explosive detonation. The results support the validity of the contact method for measuring the detonation conductivity of trotyl. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The electrical conductivity of detonation products of condensed explosives has been studied over the past 40 years [1–13]. Interest in detonation conduction is dictated by the fact that the material in a detonation wave takes a specific state, which is a strongly imperfect low-temperature plasma (pressure ≈20 GPa, mass velocity ≈2 km/s, temperature $\approx 3 \times 10^3$ K, density ≈2 g/cm^3, and molecule concentration $\approx 3 \times 10^{22}$ cm^{-3}). Available theoretical approaches to describing such a complex object are inadequate; therefore, experimental techniques prevail. It has been shown in many experiments that the conductivity of detonation products depends on the nature of an explosive and its parameters (density, etc.). For many explosives, the conductivity typically equals $\sigma \approx 1\ \Omega^{-1}\ cm^{-1}$. A value of σ as high as $100\ \Omega^{-1}\ cm^{-1}$ has been obtained only once for liquid trotyl [4]. This result has not been confirmed in later publications, and the values of σ for solid trotyl have turned out to be lower by more than one order [1, 2, 8, 10–12]. Subsequently, however, it has been found that early experiments gave conservative conductivity values for trotyl detonation products and $\sigma \approx 25\ \Omega^{-1}\ cm^{-1}$ has been obtained [13]. In [13], the time resolution was poor, about 0.5 μs, and the reaction zone was not resolved (the accuracy dropped with increasing conductivity). Interest in detonation conduction has sharpened over the last decade because of the discovery of detonation-produced diamonds [14, 15]. The electrical conductivity method was applied for testing the diamond phase [11, 13].

In the conventional contact method for measuring high ($\sigma > 1\ \Omega^{-1}\ cm^{-1}$) electrical conductivity [1], the dc mode is employed. Direct current passes in a circuit consisting of a current source, a shunt, and an explosive parallel-connected to the shunt. A detonation wave propagating through the explosive reaches the electrical contacts and connects the conducting area. Such measurements pose some difficulties: (1) The time resolution of the measuring system is $\tau \approx L/R$ (where L is the inductance and R is the resistance of the shunt–explosive circuit), which makes the measurement of very high conductivity difficult [16]; (2) at the initial time instant, the extended conducting area with the *a priori* unknown current field configuration is connected to the contacts; and (3) the contact resistances and the metal–plasma transition layers may affect the measurements.

It was the aim of our work to study the conductivity of trotyl detonation products using a scheme with a high time resolution. We used a measuring cell [17], which to a great extent is free of the disadvantages listed above.

With this cell (Fig. 1), we detected insulator(semiconductor)–metal transitions in shock waves [17–19]. The shunt, a thin metal foil, was applied on the sample. The shock enters the sample through the shunt. The voltage is detected by electrodes connected to the plane of the shunt. Such a measuring scheme provides the highest time resolution and offers a well-defined current line configuration. In addition, the effect of the

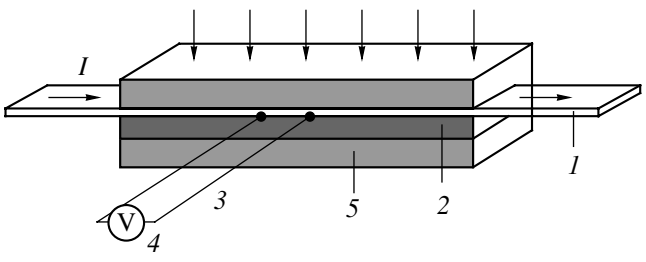


Fig. 1. Cell for measuring the conductivity of detonation products of condensed explosives: *1*, shunt (metal foil); *2*, explosive; *3*, voltage electrodes; *4*, oscilloscope; and *5*, insulator.

1063-7842/01/4609-1185$21.00

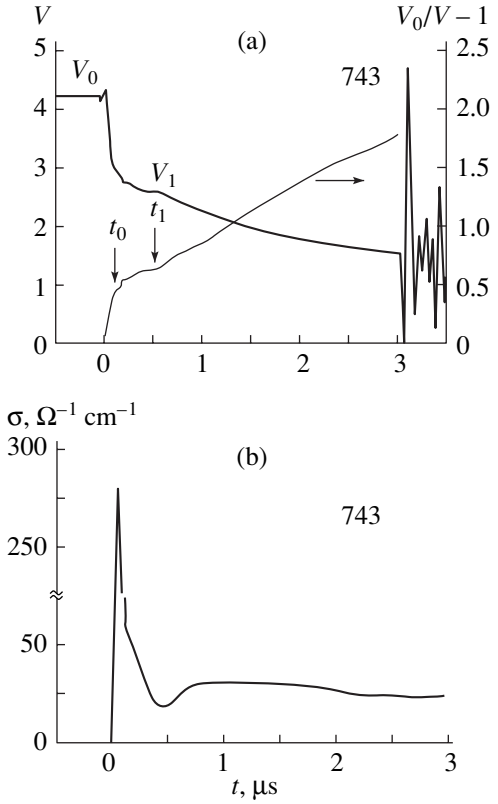


Fig. 2. (a) Experimental voltage vs. time curve for trotyl (thick line, left ordinate) and the same curve after integral processing (thin line, right ordinate); (b) conductivity profile.

contact resistances in the electrode region is weak. The time resolution of the circuit is restricted by electromagnetic transients in the shunt–conducting medium system. This restriction, however, can be considerably smoothed by analyzing the dynamic skin effect in the shock [20, 21] and using the procedure of restoring the conductivity of the conducting region based on information coming from its boundary [19]. In this work, the conductivity measurement method previously used for shock waves is applied to detonation processes.

EXPERIMENT

Trotyl filled a groove (width 20 mm, depth about 20 mm, and length 80 mm) in a textolite plate. A 0.1-mm thick constantan foil was used as a shunt. Its width varied from run to run. An initiating shock wave was generated with a plane wave generator of diameter 75 mm. Its blasting cartridge made of cast trotyl was 60 mm thick. For this generator, the time spread of wave appearance over a circle of diameter 50 mm was no more than 40 ns. The plane wave generated propagated through the 5-mm-thick insulating plate and entered the trotyl. The electrode spacing was 10 mm. The current value was up to 400 A. The signals were recorded by an S9-27 oscilloscope with a sampling interval of 10 ns.

Figure 2a shows oscillograms for the experiments with cast trotyl. The instant the shock enters the trotyl is marked by a small peak due to electromagnetic transients in the foil. This peak is a reliable temporal marker and specifies the zero time in the oscillograms. It was found in several experiments that a shock-induced change in the constantan conductivity is negligible; hence, it was not taken into account. Estimations of the shock parameters made from the shock adiabat [22] yielded the pressure $P = 22.5$ GPa. The pressure of stationary detonation in the trotyl measured with a manganin sensor in special experiments was found to be $P \approx 18$ GPa. Thus, when the shock wave entered the explosive, it was in the somewhat overcompressed state compared with stationary detonation.

As follows from the oscillograms, at the instant the shock enters the trotyl, the voltage V starts to decrease. This means that the conducting state appears in the sample without any delay. After the early sharp decrease, the rate of fall of the voltage changes and the new rate is retained until the shock wave reaches the insulating plate (high-intensity oscillations in the oscillogram). Figure 2a also shows the processed curve in the $(V_0/V - 1, t)$ coordinates (V_0 is the initial voltage). If the conductivity of the material behind the detonation front is constant and the electromagnetic transients are absent, the curve processed must represent a straight line [17]. However, the conductivity of the detonation products is seen to be variable. The early drop of the voltage correlates with the onset of the high conductivity state, which exists for the time t_0. Then, the conductivity noticeably decreases. At $t > t_1$, the curve processed is linear; hence, the conductivity remains constant in this time region. Generally, two basic portions where the conductivities greatly differ can be distinguished in the curve.

Experimental results on the conductivity of the cast trotyl detonation products are summarized in the table.

Table

Run no.	a_s, mm	D, 10^3, m/s	t_0, μs	t_1, μs	σ_1, Ω^{-1} cm^{-1}	σ_2, Ω^{-1} cm^{-1}
705	5.5	6.55 ± 0.1	0.14	0.52	180	35
726	2.85	6.69 ± 0.12	≈0.1	0.42	≈300	33
727	2.3	6.59 ± 0.1	0.30	0.67	200	56
743	2.9	6.64 ± 0.1	0.15	0.54	240	31

Here, a_s is the shunt width, D is the detonation rate, t_0 is the instant of sharp drop of the voltage, t_1 is the full time of recording the initial portion of the oscillogram (Fig. 2a), σ_1 is the mean conductivity in the range 0–t_1, and σ_2 is the mean conductivity for $t > t_1$.

The detonation rate D was found from the instants the initiating shock entered the explosive (the first peak in the oscillogram) and left it, i.e., touched the insulating plate (deviation from the straight line, the onset of the oscillations). For a trotyl density of 1.58 g/cm^3, the rate of stationary detonation has been estimated at ≈6.87 km/s [22]. The measured value of D was close to this figure. The discrepancy between the experimentally found and predicted values was the least for run no. 726 (–2.6%) and the greatest for run no. 705 (–4.7%). The delay times of initiation that were determined from the above rates are 70 and 120 ns, respectively. Since the time delay is small, we can assume that near-stationary detonation starts in the trotyl when the initiating shock enters the explosive.

The conductivity of the detonation products was found under the assumption that the expanding conducting zone is connected in parallel to the shunt. In the electrotechnical approximation, the mean conductivities for the two time ranges can be determined from the formulas

$$\sigma_1 = \frac{a_s}{a}\frac{\delta_s}{\rho_s(D-u)}\frac{1}{t}\left(\frac{V_0}{V}-1\right),$$
$$\sigma_2 = \frac{a_s}{a}\frac{\delta_s}{\rho_s(D-u)}\frac{1}{t}\left(\frac{V_0}{V}-\frac{V_0}{V_1}\right), \quad (1)$$

where a is the width of the explosive, δ_s is the shunt thickness, ρ_s is the shunt resistivity, and u is the mass velocity at the Chapman–Jouguet point. In (1), V_1 is the voltage at the time instant t_1, from which the second portion of the conductivity curve starts. It is assumed that the resistance of the first conducting zone remains unchanged (which holds for the stationary conductivity profile).

In run no. 727, the pressure of the initiating shock in the trotyl was 16.3 GPa, which is lower than the pressure at the Chapman–Jouguet point for stationary detonation. At the initial time instant, such a detonation is undercompressed. As follows from the table, the values of t_0 and t_1 are larger than in the other runs, while D, σ_1, and σ_2 are close to the associated values for the other runs. This implies that the detonation wave approaches the stationary mode.

The conductivity profile for the detonation products is obtained by using the differential processing technique [23]; in our case, it yields

$$\sigma(t) = -\frac{a_s}{a}\frac{\delta_s}{\rho_s(D-u)}\frac{V_0}{V^2}\frac{dV}{dt}. \quad (2)$$

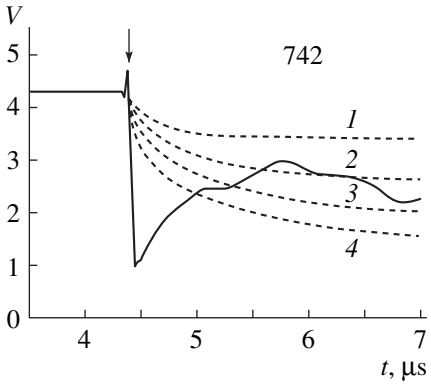


Fig. 3. Time variation of the voltage for the oppositely directed detonation wave (continuous line) and the results of electromagnetic simulation for a system having two regions of various constant conductivities (dashed lines). The conductivity of the detonation products is (*1*) 10, (*2*) 25, (*3*) 50, and (*4*) 100 Ω^{-1} cm^{-1}.

The graph of (2) is depicted in Fig. 2b. The conductivity profile is intricate: initially, the conductivity is very high; then, it sharply drops and remains constant until the detonation front comes to the insulating wall. Typical values for the two portions are $\sigma_1 \approx 2.5 \times 10^2$ and $\sigma_2 \approx 30$ Ω^{-1} cm^{-1}, respectively.

Formulas (1) and (2) are valid if two conditions are met: (1) the skin effect in the detonation products is negligible and (2) the conductivity profile is stationary. The role of the skin effect can be estimated from the ratio of the magnetic field diffusion time in a conductor to the time of wave propagation. For shock waves, the parameter $R = \mu_0\sigma(D-u)^2t$ is estimated (where σ is the electrical conductivity of the conductor) [20]. If $R < 1$, the magnetic field changes due to diffusion and the skin effect is insignificant. Such an approach can naturally be extended for detonation waves (if the mass velocity in a Taylor wave is not too high). In our runs, $R < 0.4$; therefore, we can ignore the electromagnetic nonuniformity in the first approximation and use the electrotechnical model.

A number of factors indicate that the conductivity profile is stationary. First, the experimentally found detonation rate is close to the expected one. Second, the voltage taken from the electrodes monotonically decreases as the detonation wave propagates. If the peak conductivity were due to transients during the formation of the detonation wave, the voltage records would have had a number of singularities. For example, the disappearance of the high-conductivity portion would have led to an increase in the voltage. Third, experiments with various strengths of the initiating shock wave (the pressure is smaller or greater than the Chapman–Jouguet value) give similar results. The conductivity profile for run no. 727 is similar to that shown in Fig. 2b (the only difference is the more extended region of high conductivity in the former case). Fourth,

the constancy of the conductivity in the second time interval agrees with the results obtained for greater-diameter coaxial cells [13]. Our experiments with the 38-mm-diameter coaxial (the total diameter of the explosive charge is 50 mm) have shown that, in the absence of side unloading, the conductivity of the trotyl detonation product is constant and close to the maximal value obtained in [13].

Certainly, the above considerations are indirect. Experimental methods for profiling the trotyl conductivity in a strictly stationary detonation wave are lacking. Here, difficulties are of a fundamental character, because, in experiments, the requirement of detonation stationarity comes in conflict with the requirement that edge effects of current spread be absent. For detonation to be strictly stationary, the explosive charge must be very long. As such a wave propagates, the extended conducting zone is connected to the contacts at some time instant. The current field configuration depends on the conductivity profile, which is *a priori* unknown.

Yet, we gained qualitative information on the conductivity profile in the case of stationary detonation. To this end, we used a measuring cell similar to that depicted in Fig. 1 where a detonation wave moved in the opposite direction (see the oscillogram in Fig. 3). The wave propagated in the trotyl and reached the shunt contacting with the plastic insulator (Getinaks). At this time instant, the current starts to diffuse from the shunt to the conducting region of a finite thickness. The transient process takes some time that depends on the thickness of this region and the conductivity of the detonation products. If the conductivity of the material behind the detonation front is constant, the voltage must monotonically decrease with time. Figure 3 shows a family of model curves obtained from the analysis of electromagnetic diffusion in a shunt–detonation product system having two regions of different conductivities. The curves were taken under the assumption that the conductivity behind the detonation front is constant. As soon as the detonation wave leaves the shunt (arrow in Fig. 3), the voltage is much smaller than the expected one but then its variation qualitatively agrees with the model dependence. Such a behavior can be explained as follows. First, the narrow high-conductivity zone related to the detonation front is connected to the shunt. Subsequently, this zone disappears and the voltage is controlled by current diffusion into the extended constant-conductivity region. The truth of the detonation product conductivity value thus obtained is open to question (the physical state of the products reflected from the wall may change, and the accuracy of conductivity recovery is poor). Yet, the shape of the curves points to the presence of the high-conductivity zone that spatially contacts with the detonation front. Thus, the conductivity peak moves in space with the detonation front rather than resulting from transients at the insulator–explosive interface.

DISCUSSION

The use of the new measuring scheme to study detonation processes has lead to several intriguing results. First, the detonation conductivity in the trotyl exhibits complex behavior. The initial sharp peak changes to the lower plateau. Then, the detonation conductivity turned out to be considerably higher than that obtained previously. The value $\sigma_1 \approx 2.5 \times 10^2\ \Omega^{-1}\ \mathrm{cm}^{-1}$ appears to be the highest among those which have ever been obtained for explosive detonation products. Our data agree with early results for liquid trotyl [4], where the conductivity was an increasing function ($\sigma_{max} \approx 10^2\ \Omega^{-1}\ \mathrm{cm}^{-1}$) and the measurement period was no more than 20 ns. In our experiments, the observation period was much longer, which allowed us to record the entire conductivity profile. The conductivity value on the plateau, σ_2, exceeds data in [1, 2, 8, 10–12] roughly by one order of magnitude and coincides (by one order of magnitude) with the maximal value in [13]. The cell employed in [13] did not allow the detection of the conductivity peak near the detonation front because of the worse time resolution and the influence of the edge effects.

The reason for the high conductivity of trotyl remains to be clarified. First of all, the fact stands out that the time of peak existence (≈150 ns) roughly equals the time of chemical reaction in trotyl (100–200 ns) [22]. The high conductivity in the peak seems to reflect physical processes taking place in the chemical reaction zone. In particular, it indicates that the reaction produces highly conducting particles that become bonded or spatially separated after a short time. These might be particles of carbon heated to high temperatures that is released during an exothermal reaction. Thermodynamic analysis of the detonation products has shown that the amount of free carbon in them is about 0.5 g/cm^3 [24]. The carbon particles grow in the reaction zone, forming porous fractal structures. The mechanism of macroconduction may be contact [4]. The resulting network of conducting carbon particles is responsible for the high conductivity in the peak. The expansion of the detonation products in the reaction zone (which is accompanied by additional transformations and cooling of the particles) causes the conductivity to drop. Other mechanisms that have been invoked to explain detonation conduction (thermal ionization, chemical ionization, thermionic emission, and dissociation of resulting water [1, 7–9]) cannot explain such high values of the conductivity.

The total resistance of the trotyl detonation products is ≈0.01 Ω, which makes possible the use of detonation conduction in high-current physical experiments. Because of their high density, the detonation products offer good dielectric strength, which is retained during a time specified by rarefaction gas dynamics.

CONCLUSION

A new conductivity measuring scheme improves the time resolution roughly by one order of magnitude, thus enabling the detection of new phenomena related to detonation conduction in the widely used explosive. The conductivity of the detonation products is a complex function of time, which reflects the combination of various processes taking place in the chemical reaction zone. The high sensitivity of the new method can be useful in studying fast physicochemical conversions. The effect of high conductivity concentrated in a narrow layer and transferred with a detonation rate may have a number of high-energy applications.

ACKNOWLEDGMENTS

The authors thank A.P. Ershov for the valuable discussion.

This work was supported by the Russian Foundation for Basic Research (grant no. 99-02-16807).

REFERENCES

1. A. A. Brish, M. S. Tarasov, and V. A. Tsukerman, Zh. Éksp. Teor. Fiz. **37**, 1543 (1960) [Sov. Phys. JETP **10**, 1095 (1960)].
2. R. Shall and K. Vollrath, in *Les ondes de d'etonation* (Centre National de la Recherche Scientifique, Paris, 1962), pp. 127–136.
3. R. L. Jameson, S. J. Lukasik, and B. J. Pernick, J. Appl. Phys., Part 1 **35**, 714 (1964).
4. B. Hayes, in *Proceedings of the 4th International Symposium on Detonation* (Office of Naval Research, Washington, 1967), ACR-126, p. 595.
5. A. D. Zinchenko, V. N. Smirnov, and A. A. Chvileva, Fiz. Goreniya Vzryva **7**, 422 (1971).
6. A. P. Ershov, P. I. Zubkov, and L. A. Luk'yanchikov, Fiz. Goreniya Vzryva **10**, 864 (1974).
7. V. V. Yakushev and A. N. Dremin, Dokl. Akad. Nauk SSSR **221**, 1143 (1975).
8. A. G. Antipenko, A. N. Dremin, and V. V. Yakushev, Dokl. Akad. Nauk SSSR **225**, 1086 (1975).
9. A. P. Ershov, Fiz. Goreniya Vzryva **11**, 938 (1975).
10. K. Tanaka, *Report on 5th International Colloquium of Gasdynamics of Explosions and Reactive Systems* (Bourges, France, 1975).
11. A. M. Staver, A. P. Ershov, and A. I. Lyamkin, Fiz. Goreniya Vzryva **20** (3), 79 (1984).
12. A. I. El'kind and F. N. Gusar, Fiz. Goreniya Vzryva **22** (5), 144 (1986).
13. A. P. Ershov, N. P. Satonkina, O. A. Dibirov, *et al.*, Fiz. Goreniya Vzryva **36** (5), 97 (2000).
14. A. I. Lyamkin, E. A. Petrov, A. P. Ershov, *et al.*, Dokl. Akad. Nauk SSSR **302**, 611 (1988) [Sov. Phys. Dokl. **33**, 705 (1988)].
15. Greiner N. Roy, D. S. Phillips, J. D. Johnson, *et al.*, Nature **333**, 440 (1988).
16. V. V. Yakushev, Fiz. Goreniya Vzryva **14** (2), 3 (1978).
17. S. D. Gilev and A. M. Trubachev, Prikl. Mekh. Tekh. Fiz., No. 6, 61 (1988).
18. S. D. Gilev and A. M. Trubachev, Phys. Status Solidi B **211**, 379 (1999).
19. S. D. Gilev and T. Yu. Mikhaĭlova, J. Phys. IV **5**, C3-211 (1997).
20. S. D. Gilev and T. Yu. Mikhaĭlova, Zh. Tekh. Fiz. **66** (5), 1 (1996) [Tech. Phys. **41**, 407 (1996)].
21. S. D. Gilev and T. Yu. Mikhaĭlova, Zh. Tekh. Fiz. **66** (10), 109 (1996) [Tech. Phys. **41**, 1029 (1996)].
22. L. V. Al'tshuler, G. S. Doronin, and V. S. Zhuchenko, Fiz. Goreniya Vzryva **25** (2), 84 (1989).
23. L. V. Kuleshova, Fiz. Tverd. Tela (Leningrad) **11**, 1085 (1969) [Sov. Phys. Solid State **11**, 886 (1969)].
24. Ch. L. Mader, *Numerical Modeling of Detonations* (Univ. of California Press, Berkeley, 1979; Mir, Moscow, 1985).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1190–1195. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 128–133.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Device Structures Based on Resonant Tunneling Diodes: A Theoretical Consideration

I. I. Abramov and A. V. Korolev

Belarussian State University of Informatics and Radioelectronics, ul. Brovki 6, Minsk, 220027 Belarus

e-mail: nanodev@bsuir.edu.by

Received January 9, 2001

Abstract—Simple device structures incorporating resonant tunneling diodes (RTDs) are considered in terms of electrical models and the EC–RTS–NANODEV software suite. It is shown that the structures can be used in multilevel logic, frequency converters, and generators of harmonic, relaxation, and chaotic signals. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Advanced nanodevices, offering extended functionality, have been reduced to the point where quantum-size effects become essential [1–4]. To date, a number of device structures that integrate several resonant tunneling diodes (RTDs) [5], several resonant tunneling transistors [6], RTDs and heterojunction bipolar transistors [7], RTDs and selectively doped heterojunction FETs [8], RTDs and CMOS elements [9], etc. have been created. These functionally integrated structures feature unique properties; however, their physics is much more complicated compared with the physics of the individual constituents. This is because various regions and elements may interact with each other, qualitatively changing the principle of operation. The theoretical analysis of such structures, while difficult, is necessary, since the industry is now facing the onset of nanoelectronic quantum-effect ICs.

The aim of this paper is to theoretically treat several simple RTD-based structures and to illustrate their wide functionality.

MODELS

Experience shows that integrated structures including quantum-effect devices are today the most convenient to study in terms of electrical models. More vigorous models, for example, those based on a self-consistent numerical solution of the Schrödinger and Poisson equations, as well as of the kinetic equation for the Wigner function [10], are still under development. Moreover, they are, as a rule, complex and can be applied only to simple quantum-effect devices. Such a situation radically differs from integrated devices whose operation can be described in terms of the classical diffusion–drift models [11]. Therefore, we dwelled on electrical models in our analysis.

Electrical models for resonant tunneling structures have been considered in detail elsewhere [12–16]; we will only briefly outline them. Within these models, the *I–V* characteristic of an RTD is approximated as

$$\begin{aligned} I(V) &= I_P(V/V_P)\exp[a_0 + a_1(V/V_P) + a_2(V/V_P)^2] \\ &+ a_3 I_P \exp(a_4)\{\exp[a_5(V/V_P) + a_6(V/V_P)^2] \\ &- \exp[a_7(V/V_P)]\}, \end{aligned} \quad (1)$$

where I is current, V is voltage, I_P is the RTD maximal tunnel current, V_P is the external bias at which the tunnel current is maximal, and $a_0 - a_7$ are coefficients.

Each of the terms in (1) contributes to the current through the RTD; hence, they can be considered as voltage-controlled current sources. Thus, the equivalent circuit for the RTD steady-state operation involves three diodes (Fig. 1a). The parameter R_N includes the resistance of the RTD passive parts, as well as other parasitic (under the steady-state conditions) components. The RTD equivalent circuit to study transients is given in Fig. 1b. Expression (1) is valid throughout the range of the applied voltage, while R_P, L_P, and C_P describe the RTD dynamic properties at large signals. Note that R_P may differ from R_N, although this difference is ignored for the small-signal case [15]. The parameter C_P includes the overall capacitance of the RTD active part and the parasitic components. The inductance L_P describes dynamic processes in the quantum well and also includes the parasitic components.

These electrical models are based on present-day physical concepts of RTD operation [13, 15, 16]. For simple devices considered in our work, such an approach seems to be more appropriate than using macromodels. It has been shown [17] that the latter are the most useful in simulating complex quantum-effect structures, where physical processes taking place in the individual element are of lesser concern. That the elec-

1063-7842/01/4609-1190$21.00

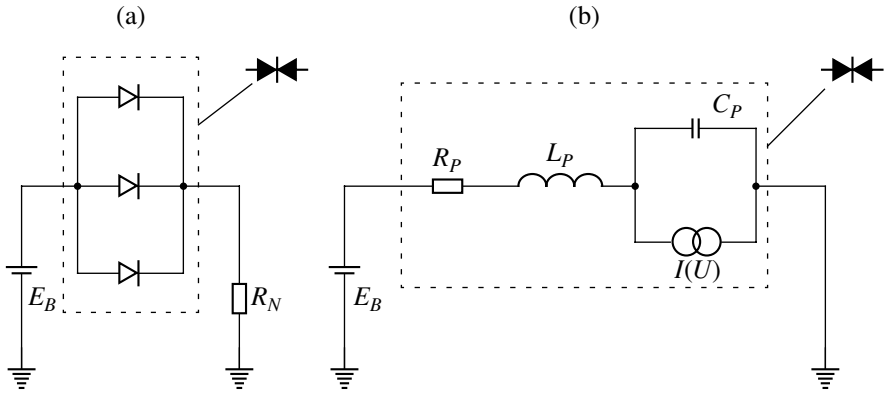


Fig. 1. RTD equivalent circuits for the (a) steady-state and (b) transient cases.

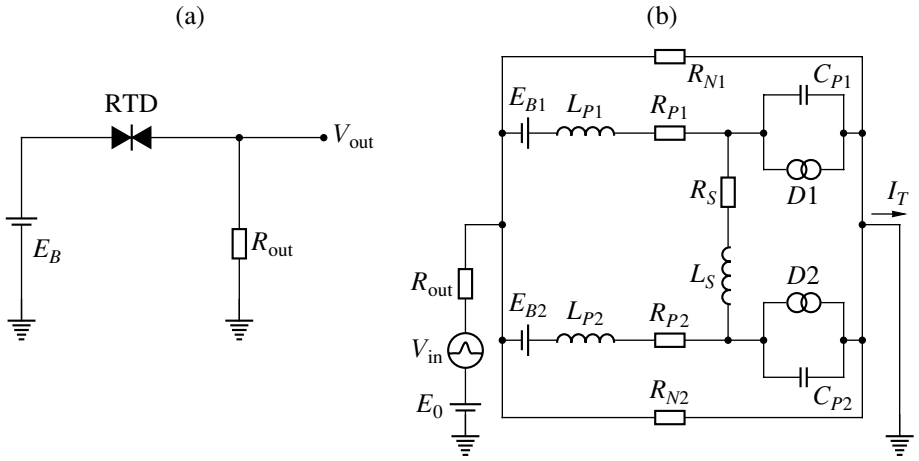


Fig. 2. Equivalent circuits for the (a) single-RTD structure and (b) integrated structure with two parallel-connected RTDs.

trical models proposed fit available experimental data for RTDs more adequately than other known models has been demonstrated in [12, 13, 15]. For example, the maximal error in simulating the *I–V* characteristic is typically 10–15%, which is a quite reasonable value in view of the strong nonlinearity of the RTD *I–V* curve. The associated technique for the identification of the model parameters from experimental data has been reported in [18, 19].

In our study, we use the ES–RTS–NANODEV [19–22] software suite for analyzing resonant tunneling devices and circuits. This suite, employing both the well-known and new electrical models, is a part of the NANODEV suite for simulating single-electron, resonant-tunneling, and quantum-interference nanodevices [23, 24].

ANALYSIS

We have reported [15, 16] that an RTD may serve as a core for generators of harmonic, relaxation, gate relaxation, etc. signals. In these applications, a dc voltage is applied to the diode. Concurrently and independently, Kawano *et al*. [25, 26] have demonstrated that similar generators can be implemented with an RTD, a capacitor, and an inductor provided that the diode is fed by a harmonic voltage with the dc component. This circuit is simple but have the inductor; therefore, the monolithic (integrated) implementation of the generator is hardly possible. That is why the discrete version of the generator was discussed in [25, 26].

Let us demonstrate the possibility of implementing signal generators of various types with a simple device structure incorporating only an RTD (i.e., without a capacitor and an inductor). Its equivalent circuit is shown in Fig. 2a. A dc bias E_B is applied to the input of the structure, and the output voltage V_{out} is taken from the resistor R_{out}. The equivalent circuit of the RTD

Values of the equivalent circuit parameters

	E_B, V	R_P, Ω	L_P, nH	C_P, pF	R_{out}, Ω
Figs. 3a, 3c	0.575	0.17	0.05	15	0.1
Figs. 3b, 3d	0.575	0.17	5	2	0.1

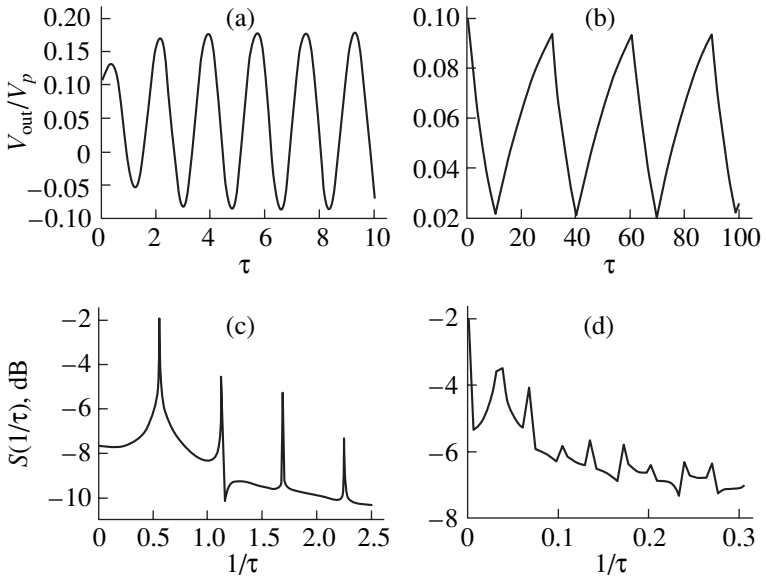


Fig. 3. Results of simulation of the simple single-RTD circuit ($\tau = t/t_0$, $t_0 = 10^{-10}$ s): (a, b) harmonic and relaxation oscillations, respectively; (c, d) their respective spectra.

alone is depicted in Fig. 1b. For the *I–V* characteristic approximated by model (1), we used experimental data for structure *A* in [27]. This was done intentionally to improve the validity of the theoretical consideration.

The results of simulation of the structure in Fig. 2a are presented in Fig. 3 for almost harmonic (Fig. 3a) and relaxation (Fig. 3b) output signals. The associated spectral characteristics of the signals are illustrated in Figs. 3c and 3d. The table lists the values of the parameters of the equivalent circuit for the signals generated. The spread of the parameter values is seen to be small.

As follows from the table, the generators of various signals can be implemented with the simple structure including an RTD and a load resistor. The parameters of the equivalent circuit, namely, R_P, L_P, C_P, and R_{out}, can be varied by varying the geometry of the passive part of the integrated structure, by additionally doping some of its regions, by varying the thickness of the structure layers, etc., since the design, process, and physical parameters of the RTD specify the values of the circuit parameters listed above. The device performance can easily be improved by slightly modifying the fabrication technology. One can use empirical formulas for R_P, L_P, C_P, and R_{out} that are derived at the design stage. It should be noted that the parameter values listed in the table are typical of actual RTDs [27–32].

Consider a simple integrated structure that includes two RTDs and is described by the equivalent circuit in Fig. 2b. We will show that this structure can be used in multilevel logic, as a frequency converter, or as a signal generator, depending on applied biases and equivalent circuit parameters.

First of all, we note that this structure is similar to that experimentally studied in [5] at $T = 100$ K. The most considerable difference is the coupling character. In [5], two RTDs are connected through a resistance only. In the circuit shown in Fig. 2b, the diodes are connected via the resistor R_S and inductor L_S. Such a connection can be viewed as interaction of two parallel-connected resonant tunneling elements in the integrated structure through their active regions. Next, our values of R_S are much lower than in [5]. The presence of R_{N1} and R_{N2} does not considerably affect the operation of the circuit.

The analytical *I–V* curve of the integrated structure (Fig. 4) establishes that the structure can be used as a multilevel logic gate. Numerical experiments indicate that the shape of this double-peak curve strongly depends on the bias E_{B2}. This is because the partial *I–V*

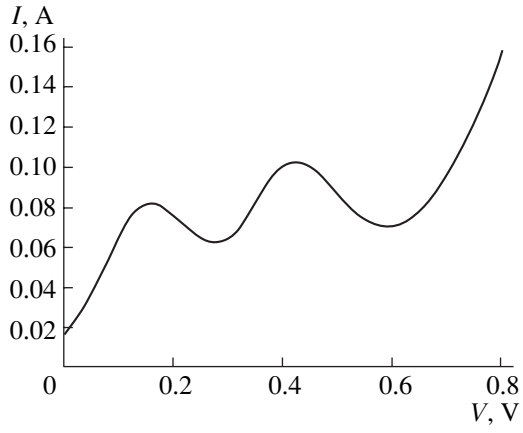


Fig. 4. *I–V* curve for the logic gate consisting of two parallel-connected RTDs.

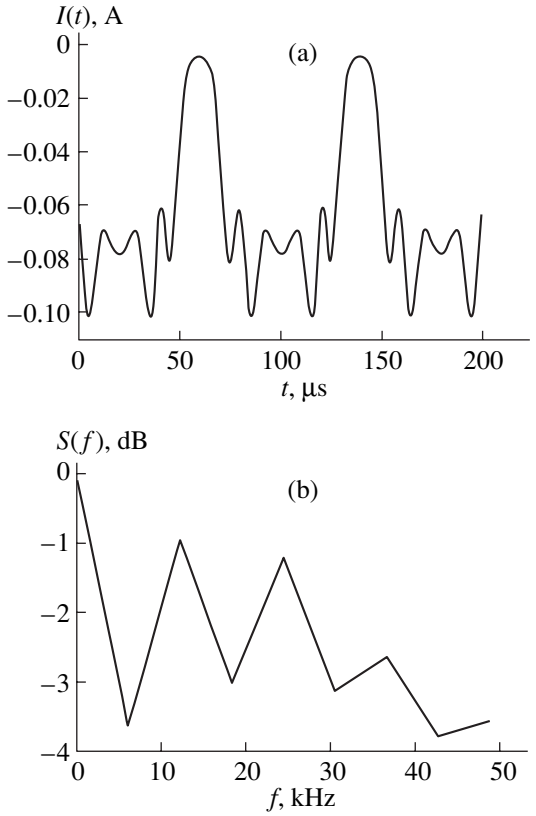


Fig. 5. Results of simulation of the double-RTD structure: (a) output current and (b) its spectrum.

characteristics of the RTDs virtually superpose when E_{B2} is below some critical value. In addition, the overall *I–V* characteristic of the structure depends on R_S and L_S, since these parameters specify feedback between the RTDs. The values of the parameters used in the numerical calculations were $E_{B1} = 0$, $E_{B2} = 0.35$ V, $R_{P1} = R_{P2} = 0.17\ \Omega$, $L_{P1} = L_{P2} = 1.01$ nH, $C_{P1} = C_{P2} = 2.96$ pF, $R_S = 1.0\ \Omega$, and $L_S = 1.01$ nH. The partial *I–V* curves are similar to that of structure *A* in [27]. The resistances R_{N1} and R_{N2} were not taken into consideration in the calculations.

The structure under study can also serve as a frequency converter. Figure 5 demonstrates the results of simulation for the parameter values listed above when a harmonic signal

$$V_{\text{in}} = A\sin(ft), \tag{2}$$

was applied to the input. Here $A = 0.35$ V and $f = 12.5$ kHz. E_0 was taken to be equal to 0.3 V.

The current shown in Fig. 5a passes through the resistor R_{out}. The corresponding spectrum of the output signal is shown in Fig. 5b. It is obvious that this structure is promising for a frequency multiplier.

Finally, it remains to demonstrate that, with a dc bias ($V_{\text{in}} = 0$), this integrated structure can be used as a signal generator if the parameters of the RTDs slightly differ. In this case, $R_{\text{out}} = 0$ and the current I_T (Fig. 2b) is taken as an output. The *I–V* curve of one RTD corresponds to the structure *A* in [27], and the *I–V* curve of the other differs by approximately 2%. Figure 6 shows the results of simulation for the circuit in Fig. 2b with $R_{P1} = 0.17\ \Omega$, $R_{P2} = 0.167\ \Omega$, $L_{P1} = 0.01$ nH, $L_{P2} = 0.0098$ nH, $C_{P1} = 0.5$ pF, $C_{P2} = 0.49$ pF, $R_S = 8.012\ \Omega$, and $L_S = 0.1$ nH.

From Fig. 6a, it follows that the output signal is nearly harmonic in this case (its spectrum is depicted in

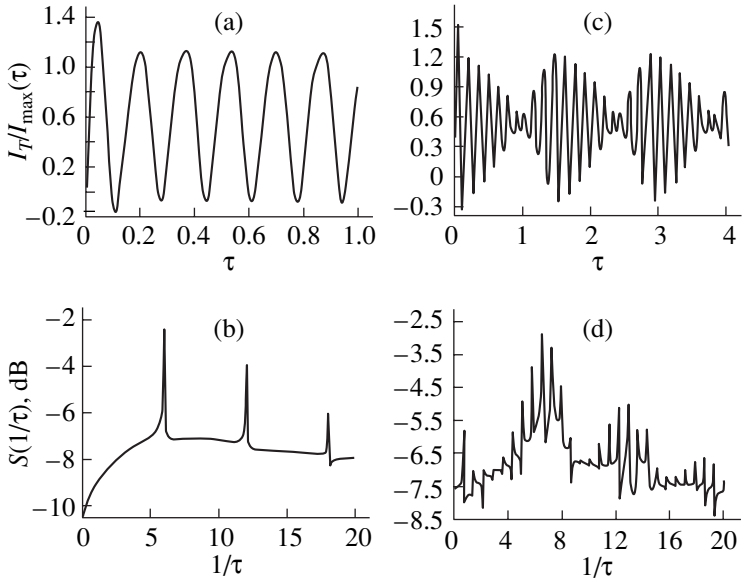


Fig. 6. Results of simulation of the double-RTD structure for $\tau = t/t_0$, $t_0 = 10^{-10}$ s, $I_{\max} \approx 0.2$ A, $E_0 = 0.55$ V, and $E_{B1} = 0$. (a, c) Oscillations and (b, d) their spectra. E_{B2} = (a, b) 0 and (c, d) 0.1 V.

Fig. 6b). Figure 6c shows the almost chaotic output signal (its associated spectral characteristic is given in Fig. 6d). Such a large discrepancy between the signals was obtained with a slight difference in only one parameter: in Fig. 6a, $E_{B2} = 0$, while in Fig. 6c, $E_{B2} = 0.1$ V.

It is worth noting that the equivalent circuit in Fig. 2b can be used as a basis for electrical models for a number of new structures [33–38] that exploit the principle of coherent charge carrier transport with self-organization [39]. These structures have been theoretically treated in [33–39], and early attempts to analyze them in terms of electrical models have been made in [14, 40]. In these structures, the partial *I–V* curves of the diodes in Fig. 2b may differ much greater. Furthermore, the resistances R_{P1}, R_{P2} and the inductances L_{P1}, L_{P2} may also substantially differ, so that there is no need for the voltage sources E_{B1} and E_{B2}. The extended functionality of the new devices is due to more complex physical processes in the nanostructure rather than to the spatial separation of two resonant tunneling regions as in the structure considered before.

CONCLUSION

Using original electrical models of an RTD and the EC–RTS–NANODEV software suite, we analyzed several simple RTD-based integrated devices. In spite of the simplicity of the devices, they were shown to be capable of performing a variety of complex functions, depending on the applied (usually dc) biases and the equivalent circuit parameters. Specifically, they can serve as multilevel logic gates; frequency converters; and generators of harmonic, relaxation, and chaotic signals. With a slight difference in the values of the equivalent circuit parameters, the integrated nanostructures may provide a basis for analog and digital circuits. The feasibility of such an approach is also supported by the fact that the parameter values used in this study are typical of actual RTDs [27–32].

ACKNOWLEDGMENTS

This work was partially supported by the Belarussian Research Programs “Informatics,” “Low-Dimensional Structures,” and “Nanoelectronics.”

REFERENCES

1. Zh. I. Alferov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 3 (1998) [Semiconductors **32**, 1 (1998)].
2. *Resonant Tunneling in Semiconductors: Physics and Applications*, Ed. by L. L. Chang, E. E. Mendez, and C. Tejedor (Plenum, New York, 1991), NATO ASI Ser., Ser. B **277** (1991).
3. *Single Charge Tunneling: Coulomb Blockade Phenomena in Nanostructures*, Ed. by H. Grabert and M. H. Devoret (Plenum, New York, 1992), NATO ASI Ser., Ser. B **294** (1992).
4. *Nanostructure Physics and Fabrications*, Ed. by M. A. Reed and W. P. Kirk (Academic, Boston, 1989).
5. S. Sen, F. Capasso, A. Y. Cho, and D. Sivco, IEEE Trans. Electron Devices **34**, 2185 (1987).
6. F. Capasso, S. Sen, F. Beltram, *et al.*, IEEE Trans. Electron Devices **36**, 2065 (1989).
7. C. E. Chang, P. M. Asbeck, K.-C. Wang, and E. R. Brown, IEEE Trans. Electron Devices **40**, 685 (1993).
8. M. Kawashima, H. Hayashi, H. Fukuyama, *et al.*, Jpn. J. Appl. Phys. **39**, 2468 (2000).
9. J. I. Bergman, J. Chang, Y. Joo, *et al.*, IEEE Electron Device Lett. **20** (3), 119 (1999).
10. B. A. Biegel and J. D. Plummer, Phys. Rev. B **54**, 8070 (1996).
11. I. I. Abramov, *Modeling of Physical Processes in Silicon Integrated Circuit Elements* (Beloruss. Gos. Univ., Minsk, 1999).
12. I. I. Abramov, A. L. Danilyuk, A. V. Korolev, and E. A. Patent, in *Proceedings of the 8th International Crimean Microwave Conference, Sevastopol, 1998*, p. 599.
13. I. I. Abramov, A. L. Danilyuk, and A. V. Korolev, Izv. Beloruss. Inzh. Akad., No. 2 (6), 43 (1998).
14. I. I. Abramov, A. L. Danilyuk, and A. V. Korolev, Izv. Beloruss. Inzh. Akad., No. 1 (7), 119 (1999).
15. I. I. Abramov, A. L. Danilyuk, and A. V. Korolev, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **43** (3), 59 (2000).
16. I. I. Abramov, A. L. Danilyuk, and A. V. Korolev, Vestsi Akad. Navuk Belarusi, Ser. Fiz.-Tékh. Navuk, No. 2, 75 (2000).
17. S. Mohan, J. P. Sun, P. Mazumber, and G. I. Haddad, IEEE Trans. Comput.-Aided Des. **14**, 653 (1995).
18. I. I. Abramov and A. V. Korolev, in *Proceedings of the 7th International Scientific and Technical Conference “Topical Problems of Solid-State Electronics and Microelectronics,” Divnomorskoe, 2000*, Vol. 2, p. 16.
19. I. I. Abramov, I. A. Goncharenko, and A. V. Korolev, in *Proceedings of the 10th International Crimean Microwave Conference, Sevastopol, 2000*, p. 418.
20. I. I. Abramov, I. A. Goncharenko, A. L. Danilyuk, and A. V. Korolev, in *Proceedings of the 9th International Crimean Microwave Conference, Sevastopol, 1999*, p. 296.
21. I. I. Abramov, Yu. A. Berashevich, I. V. Sheremet, and I. A. Yakubovskiĭ, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **42** (2), 46 (1999).
22. A. V. Korolev and I. I. Abramov, in *Proceedings of the 7th International Scientific and Technical Conference “Topical Problems of Solid-State Electronics and Microelectronics,” Divnomorskoe, 2000*, Vol. 2, p. 13.
23. I. I. Abramov, I. A. Goncharenko, E. G. Novik, and I. V. Sheremet, in *Proceedings of the 6th International Crimean Microwave Conference, Sevastopol, 1996*, p. 294.
24. I. I. Abramov and E. G. Novik, *Numerical Simulation of Metallic Single-Electron Transistors* (Bestprint, Minsk, 2000).
25. Y. Kawano, Sh. Kishimoto, K. Maezawa, and T. Mizutani, Jpn. J. Appl. Phys. **38**, L1321 (1999).

26. Y. Kawano, Sh. Kishimoto, K. Maezawa, and T. Mizutani, Jpn. J. Appl. Phys. **39**, 3334 (2000).
27. J. M. Gering, D. A. Crim, D. G. Morgan, *et al.*, J. Appl. Phys. **61**, 271 (1987).
28. E. R. Brown, C. D. Parker, and T. C. L. G. Sollner, Appl. Phys. Lett. **54**, 934 (1989).
29. C. Y. Huang, J. E. Morris, and Y. K. Su, J. Appl. Phys. **82**, 2690 (1997).
30. K. J. Gan, Y. K. Su, and R. L. Wang, J. Appl. Phys. **81**, 6825 (1997).
31. K. J. Gan and Y. K. Su, J. Appl. Phys. **82**, 5822 (1997).
32. K. J. Gan and Y. K. Su, Jpn. J. Appl. Phys. **36**, 6280 (1997).
33. I. I. Abramov and A. L. Danilyuk, Appl. Phys. Lett. **71**, 665 (1997).
34. I. I. Abramov and A. L. Danilyuk, in *Proceedings of the 7th International Crimean Microwave Conference, Sevastopol, 1997*, p. 379.
35. I. I. Abramov and A. L. Danilyuk, Vestsi Akad. Navuk Belarusi, Ser. Fiz.-Tékh. Navuk, No. 3, 64 (1997).
36. I. I. Abramov and A. L. Danilyuk, Zh. Tekh. Fiz. **68** (12), 93 (1998) [Tech. Phys. **43**, 1485 (1998)].
37. I. I. Abramov and A. L. Danilyuk, Dokl. Akad. Nauk Belarusi **42** (5), 55 (1998).
38. I. I. Abramov and A. L. Danilyuk, in *Proceedings of the 8th International Crimean Microwave Conference, Sevastopol, 1998*, p. 602.
39. I. I. Abramov and A. L. Danilyuk, in *Proceedings of the 6th International Crimean Microwave Conference, Sevastopol, 1996*, p. 45.
40. I. I. Abramov, A. L. Danilyuk, and A. V. Korolev, Vestsi Akad. Navuk Belarusi, Ser. Fiz.-Tékh. Navuk, No. 1 (2001).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1196–1198. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 134–136.

BRIEF COMMUNICATIONS

Fast Generation of an Electron Beam in a Magnetron Gun with a Secondary-Emission Metallic Cathode

Yu. Ya. Volkolupov, A. N. Dovbnya, V. V. Zakutin, M. A. Krasnogolovets, N. G. Reshetnyak, and V. P. Romas'ko

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

Received October 30, 2000

Abstract—The formation of an electron layer and the generation of an electron beam in magnetron guns where secondary emission is triggered by nanosecond pulses are studied. In the guns with small cross sizes, hollow electron beams with an outer diameter of 3–6 mm are generated. The beam current is 1–2 A, and the cathode voltage is 5–7 kV. Results obtained indicate that the generation of nanosecond beam-current pulses is a possibility. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Recent years have seen extensive research on cold-cathode electron sources that employ secondary emission in crossed electric and magnetic fields [1–4]. Among their merits are long operating time, high current density, simple design, etc. Therefore, they could make emitters for long-lived high-power microwave generators [5] or fast high-voltage devices [6]. In this work, we studied the fast (within 1–10 ns) formation of the space charge and the temporal stability of beam generation in magnetron guns.

EXPERIMENTAL SETUP

Experiments to generate electron beams were conducted with the setup shown in Fig. 1. A magnetron gun is driven by modulator *1*, which produces 4- to 100-kV negative pulses U_c of width 2–10 μs and repetition rate 10–50 Hz. The pulses are applied to cathode *5*, whereas anode *6* is grounded via resistor R_3. Secondary emission takes place during the trailing edge of a voltage pulse. Two techniques to initiate secondary emission were employed. In the first case, two pulse generators *2* were used: the former with a voltage pulse amplitude U = 2–15 kV and a trailing edge duration $\tau \approx 70$ ns; the latter with U = 3.5 kV across a 50-Ω load resistor and a duration of the leading and trailing edges of ≈1 ns. In the other case, secondary emission was initiated during the 0.6-μs-wide trailing edge of a specially generated overshoot [5]. This allowed us to vary the trailing edge duration between 2 and 600 ns and the voltage pulse steepness, between 20 and 1200 kV/μs. A magnetic field of strength $H \le 3000$ Oe was generated by solenoid *4*. The beam current and size were measured 180 mm away from the anode plane with Faraday cup *7* (≈40-cm-long coax) and resistor R_4 of value equal to the wave impedance of the coax (18 Ω). The cathode voltage was measured with voltage divider R_1R_2. The beam size was measured from its image on an X-ray film and on a molybdenum foil. The cathode and the anode of the magnetron gun were made of copper and stainless steel, respectively. The anode length was 120 mm. The gun was placed in vacuum chamber *3*, which was evacuated to a pressure of $\le 10^{-6}$ torr.

When secondary emission is initiated by nanosecond pulses, one must reckon with distortions of their shapes in transmitting lines. In our setup, the initiating pulses pass through the coax, high-voltage insulator, vacuum transmission line, and the anode fixture. To determine the actual shape of an initiating nanopulse, we measured its parameters between the anode and the cathode with regard for parasitic inductances and capacitances. Figure 2 shows the time variation of the pulse shape for pulse durations of ≈2 and ≈6 ns. The waveforms were recorded by an I2-7 oscilloscope (transmission band ≈3000 MHz). It is seen that the shorter pulse almost retains its shape (the fall time is ≈2 ns), while the longer one is distorted: its fall time increased to ≈11 ns.

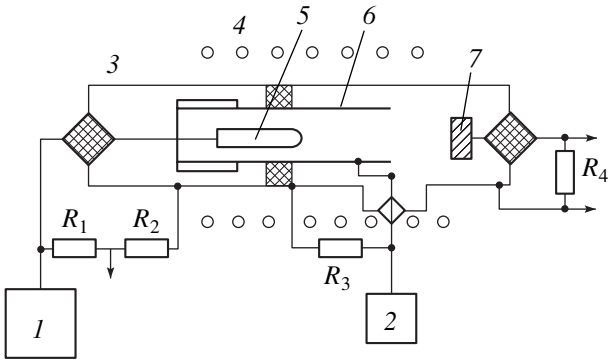


Fig. 1. Experimental setup.

1063-7842/01/4609-1196$21.00 © 2001 MAIK "Nauka/Interperiodica"

RESULTS AND DISCUSSION

Secondary emission was triggered under various conditions. With the fall time of the initiating pulses equal to ≈0.6 μs and the pulse steepness to 20–50 kV/μs, secondary emission and beam generation started 100–500 ns (depending on experimental conditions) after the beginning of the voltage pulse fall. For such a small steepness, the number of primary electrons must be significant, since only a minor part of them acquires the energy necessary to initiate secondary emission. Therefore, the accumulation of primary electrons with the necessary energy takes a long time and is of statistical character. The time spread may be as high as several tens of nanoseconds. Accordingly, the onset of current pulse generation varies in time and the electrons in the beam are distributed in energy. On the other hand, the duration of the leading edge of a beam pulse depends on the steepness of an initiating pulse; in our experiments, the duration of the leading edge of a beam pulse was several tens of nanoseconds. To improve the temporal stability, cut the beam pulse duration, and decrease the energy spread of the beam electrons, one should make the initiating pulse steeper and shorter.

With this in mind, we performed experiments where the initiating pulses had either a nanosecond fall time (Fig. 2) or a steepness of greater than 300 kV/μs. Under these conditions, the electron beam arose within 1.5–2.0 ns in the former case and ≈10 ns in the latter case after the pulse amplitude had decayed. This is also observed in Fig. 3, where typical waveforms of cathode voltage pulses (with and without the current beam) and beam current pulses from the Faraday cup are presented (cathode and anode diameters are 2 and 10 mm, respectively). The temporal instability of the onset of the beam current pulse does not exceed the initiating pulse fall time and is about several nanoseconds.

The results obtained are listed in the table. Here, d_c and D_a are the diameters of the gun cathode and anode, respectively; U_c is the cathode voltage; I is the beam current; U and τ are the amplitude and the fall time of the initiating pulse; and H is the magnetic field. During the measurements, the cathode voltage amplitude did not exceed the breakdown value for the interelectrode spacing (depending on the gun geometry, the maximal amplitude was between 15 and 25 kV). The table lists the minimal amplitude values at which beam generation still takes place. This value is of practical interest, because it sets the lower voltage limit at which secondary emission and beam generation are possible. For example, the energy of primary electrons bombarding the copper cathode should lie in the range of 0.4–0.6 keV. With such energies, the secondary emission factor reaches its maximum value and the process of secondary emission is very intense. In this case, the energy of primary electrons ranges up to ≈10% of the electron energy at the exit from the gun. This is of inter-

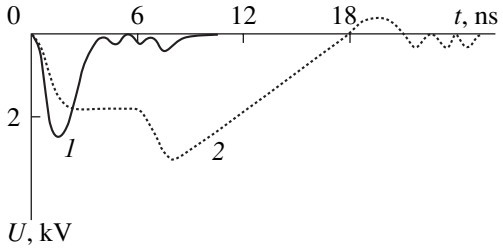


Fig. 2. Waveforms of the initiating pulse between the cathode and the anode. The pulse duration is (*1*) 2 and (*2*) 6 ns.

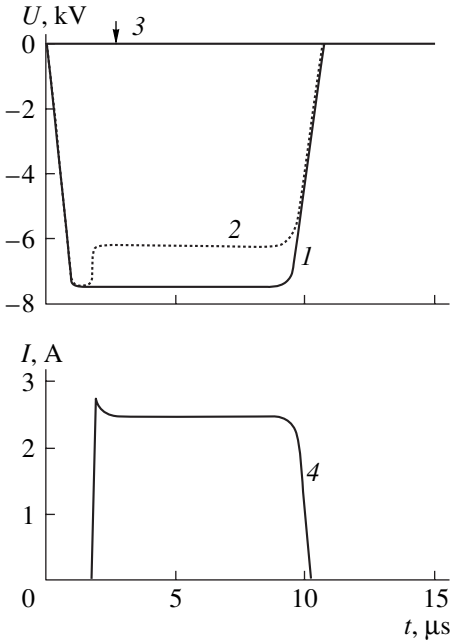


Fig. 3. Waveforms of (*4*) the beam current pulse from the Faraday cup and the cathode voltage (*2*) with and (*1*) without the beam current. Arrow *3* indicates the instant the initiating pulse is applied.

Table

Gun no.	d_c, mm	D_a, mm	U_c, kV	I, A	H, Oe	U, kV	τ, ns
1	2	7	7	1.9	3000	2.4	2
2	2	10	5	0.8	1900	4	13
3	2	10	7	1.6	2100	3	11
4	3	14	8	2.3	1400	3	14

est for the formation of both a stable electron layer and an electron beam with a considerable energy spread.

The study of beam generation vs. pulse steepness showed that this dependence has a threshold. For beam generation to be stable at nanosecond fall times, the steepness must be much greater than at larger times. For example, beam generation in gun no. 1 occurred at a steepness of more than 1000 kV/μs and in gun no. 2, at a steepness above 300 kV/μs. These values are more than one order of magnitude higher than those necessary for generating the beam at a fall time of 0.6 μs. The same is true for gun no. 3 and gun no. 4.

When the gun is triggered by a voltage pulse with a nanosecond fall time, the energy spectrum of the beam is improved because of a decrease in the amount of fast electrons that are generated during the trailing edge of a small-amplitude pulse.

Our experimental data are in satisfactory agreement with results obtained by numerically simulating the formation of the electron layer in the crossed fields. Our computations, as well as those in [7, 8], indicate that, during the fall time of the initiating voltage pulse (1 or 2 ns), the electrons may gain energy large enough for secondary emission to be initiated. With such short fall times, the amount of primary electrons is small; however, because of the very high steepness, they have a chance of gaining the necessary energy for a small number of gyroperiods, so that intense secondary emission takes place.

In both cases, the cross section of the beams was a ring with the uniform azimuth distribution of the intensity. The inner diameter was found to be roughly equal to the cathode diameter, and the "wall" thickness was 1–1.5 mm. For example, for gun no. 2, the beam has outer and inner diameters of 4 and 2 mm, respectively, 180 mm away from the anode plane.

CONCLUSION

Thus, our experiments have demonstrated the possibility of forming the space charge and generating an electron beam in a magnetron gun with a secondary emission cathode within ≈2 ns. This enables the synchronization of the beam current pulses with a nanosecond accuracy. With a cathode voltage of 5–7 kV and a magnetic field strength of 1900–3000 Oe, hollow electron beams with a current of 1–2 A and an outer diameter of 3–6 mm have been obtained.

REFERENCES

1. V. M. Lomakin and L. V. Panchenko, Élektron. Tekh., Ser. 1, No. 2, 33 (1970).
2. J. F. Skowron, Proc. IEEE **61** (3), 69 (1973).
3. S. A. Cherenshchikov, Élektron. Tekh., Ser. 1, No. 6, 20 (1973).
4. A. N. Dovbnya, V. V. Zakutin, N. G. Reshetnyak, *et al.*, in *Proceedings of the 5th European Particle Accelerator Conference*, Ed. by Myers A. Pacheco, R. Rascual, *et al.* (Inst. of Physics Publ., Bristol, 1996), Vol. 2, p. 1508.
5. V. V. Zakutin, A. N. Dovbnya, N. G. Reshetnyak, *et al.*, in *Proceedings of the 1997 Particle Accelerator Conference, Vancouver, 1997*, Ed. by M. Comyn, M. K. Craddok, M. Reiser, and J. Thomson, Vol. 3, p. 2820.
6. A. I. Vishnevskiĭ, A. I. Soldatenko, and A. I. Shendakov, Izv. Vyssh. Uchebn. Zaved., Radioélektron. **11**, 555 (1968).
7. A. V. Agafonov, V. P. Tarakanov, and V. M. Fedorov, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled. **1** (2–3), 134 (1997).
8. A. V. Agafonov, V. P. Tarakanov, and V. M. Fedorov, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled., No. 4, 11 (1999).

Translated by V. Isaakyan

Technical Physics, Vol. 46, No. 9, 2001, pp. 1199–1201. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 137–139.

BRIEF COMMUNICATIONS

Twinning in Bismuth Single Crystals Induced by Thermal Cycling

O. M. Ostrikov

Sukhoĭ State Technical University, Gomel, 246746 Belarus

Received November 1, 1999; in final form, May 5, 2000

Abstract—The role of twinning in bismuth single crystal plastic deformation under thermal cycling (cooling to 77 K and heating to 373 K) is studied. It is found that, under these conditions, the twinning of the crystals proceeds in several stages. © *2001 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Under operating conditions, engineering materials are often subjected to thermal cycling. For instance, the rubbing metallic parts of an internal combustion engine heat up at its start and cool down when it stops. Thermal stresses that occur in the material under these conditions change its dislocation structure and may cause the breakdown of the material. That is why the study of the effect of thermal cycling and, in particular, the effect of the resulting thermal flows on the mechanical properties of solid crystals is of great physical interest.

In crystals, internal stress relaxation caused by temperature variation does not always happen through dislocation glide. In a number of important engineering materials (α-Fe, Fe + 3.5% Si, Ti, AlTi, etc.), the relaxation may be due to twinning. Twinning is often associated with crack nucleation in a crystal [1–3]; therefore, studying twin nucleation under thermal cycling would help to extend the service life of the materials.

The object of our work is to elaborate a physical model of thermal stress relaxation in bismuth single crystals and to discover the role of twinning in this process.

EXPERIMENTAL

Bismuth is a convenient model material [4–6] that is prone to twinning and does not require a large temperature drop to initiate twin nucleation. Also, the twinning laws found in this material can be extended to other metals, since twinning dislocations are similar in all crystals: they are Shockley partial dislocations and differ only by magnitude and sense of the Burgers vector.

The bismuth single crystals were grown by the Bridgman method from a 5N-purity raw material. Specimens measuring $4 \times 5 \times 10$ mm were obtained by cleaving the ingot along the (111) cleavage plane. Because of the layer structure and of the crystals and pronounced cleavage properties in them, the as-cleaved surface can be examined metallographically without additional treatment.

The specimens were stressed by thermal cycling: cooling to 77 K in liquid nitrogen and subsequent heating to 373 K in boiling water. A cooling/heating cycle lasted 40 s.

Using a PMT-3 optical microscope, we measured the twin dimensions and detected the instant of their nucleation.

RESULTS AND DISCUSSION

The {110}⟨001⟩-type twin density on the (111) bismuth single crystal surface vs. the number n of thermal cycles is presented in Fig. 1. We have

$$\rho = \frac{N}{S}, \tag{1}$$

where N is the number of twins on the area S of the (111) bismuth single crystal surface.

As is seen from Fig. 1, the curve $\rho = \rho(n)$ has three stages: (1) the incubation stage ($0 < n < 25$), (2) the stage where the number of the twins grow ($25 < n < 200$), and (3) the saturation stage ($n > 200$).

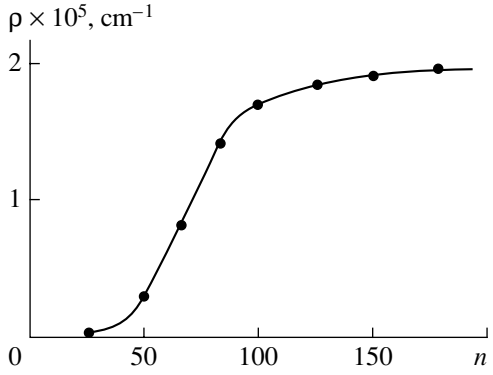


Fig. 1. Twin density ρ on the (111) single-crystal bismuth surface vs. the number n of thermal cycles.

1063-7842/01/4609-1199$21.00

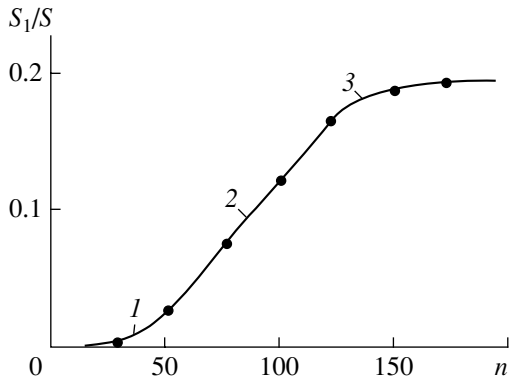


Fig. 2. Relative area occupied by twins on the (111) surface vs. the number of thermal cycles. (*1*–*3*) Twinning steps.

At the incubation stage, dislocations collect on the $\{11\bar{1}\}$ planes becomes difficult because the defects already formed cause stresses that block the sources of dislocations. This makes the twinning process, the Schmidt factor of which is close to that of dislocation glide on the $\{11\bar{1}\}$ planes, energetically more favorable.

Note that twinning takes place for the most part when dislocation glide is difficult or impossible. In general, the dislocations cannot glide because of the symmetry selection rule, the low-temperature limitations, and limitations due to impurities. In our case, twinning is associated with energetically favorable dislocation glide on the $\{11\bar{1}\}$ planes.

The ratio S_1/S, where S_1 is the surface area of twins and S is the area on the (111) surface, vs. the number n of thermal cycles is presented in Fig. 2. The value of S_1 was defined as the sum of the products of the lengths and widths of individual twins. The dimensions were measured on the (111) single-crystal bismuth surface. The curves $S_1/S = f(n)$ and $\rho = \rho(n)$ are of a similar shape, but the interpretation of the results shown in Fig. 2 is different. During the incubation period, the twins are not detected. The growth of S_1 after 25 cycles is accounted for largely by twins of length to 10 μm and width to 2–3 μm. Then, the ratio S_1/S (Fig. 2) rises because the number of the twins grows and they increase in size. The change in their size is due to alternating-sign stresses generated by thermal cycling. The effect of the alternating-sign stresses on the twinning kinetics in bismuth single crystals has been described in [7–11]. In those papers, the Bauschinger effect and twin-boundary pinning have been studied. According to [10, 11], the twin boundaries stop moving in the direction normal to them after a number of loading/unloading cycles. A similar situation occurs upon thermal cycling. The twins with at a length of 20 μm (sometimes 40 μm) and a width of 5 μm cease to grow. This happens after 200 thermal cycles.

Thus, at the stage of twin growth in bismuth single crystals subjected to thermal cycling, we can separate three characteristic steps (Fig. 2).

1. Nucleation of Twins

At this step, the $\{11\bar{1}\}$ perfect dislocations are split into partial twinning ones with the formation of the twin–mother crystal interfaces. As a rule, this process takes place owing to $\{11\bar{1}\}$ dislocation clusters formed during the incubation period. Because of the thermal stresses and difficult dislocation glide along the $\{11\bar{1}\}$ system, the stresses in the clusters relax either through the basal slip along (111) or through the generation of $\{110\}\langle 001\rangle$ twinning dislocations. The latter process is energetically more favorable due to the greater Schmidt factor and the superposition of the thermal stresses and those caused by the $\{11\bar{1}\}$ dislocations.

2. Increasing the Number of Twing and Their Growth

At this step, the number of the twins increases not only because of the thermal and dislocation stresses but also because of the stresses produced by incoherent twin boundaries in the crystal. The twins expand mostly by the translation of the twinning dislocations along the existent twin boundaries. This results in the displacement of the twin edges and causes the twin growth in the direction normal to the boundary.

3. Twinning Saturation

At this step, the twinning pattern on the (111) surface stabilizes and the curve $S_1/S = f(n)$ saturates because of the pinning of the twin boundaries by Frank dislocations generated in the vicinity of the boundaries. According to [12], Frank dislocations appear by the reaction

$$[\bar{1}01]_{0.92} + [0\bar{1}\bar{1}]_{3.08} \longrightarrow [\bar{1}\bar{1}0]_{3.08}. \qquad (2)$$

Here, the subscripts are the ratios of the Burgers vector squared to the lattice parameter squared (that is, the relative dislocation energy). In this case, the interacting dislocations lie in the (111) and $(11\bar{1})$ crystallographic planes and the resultant dislocation, in (001).

Thus, like dislocation glide along $\{11\bar{1}\}$, the twinning process saturates (Fig. 1). Basal slip along (111), which is observed at the third step of twinning as well, succeeds twinning. The thermal stresses here have only an indirect effect because of the symmetry selection rule, suppressing the plastic deformation process considered. In this case, as at the last step of twinning, basal slip is initiated by the stresses from the $\{11\bar{1}\}$ dislocation clusters and by the twin boundaries. The

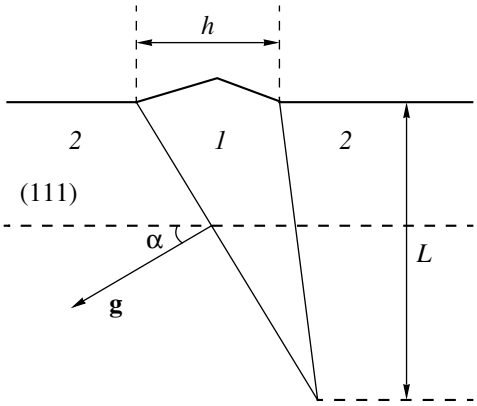


Fig. 3. Direction of the elastic forces **g** of a twin boundary in relation to the twin boundary and to the (111) cleavage plane: (*1*) twin of length *L* and width *h* and (*2*) mother crystal.

direction of the elastic forces **g** of a twin boundary in reference to the boundary and to the (111) cleavage plane is shown in Fig. 3. The angle α depends on the twin length *L* and width *h* and is about 10°; that is, for the basal slip initiated by the twin boundary, the Schmidt factor roughly equals 0.2.

It is obvious that the basal slip is also saturated after a number of cycles. This process is favored by the twin boundaries, which serve as stoppers for the basal slip. This sets conditions for microcracking.

Notice that, in materials that are not prone to twinning, cracking at thermal cycling may appear earlier than in those prone to twinning, because the energy is not spent on twin nucleation and growth. Hence, the tendency to twinning can be considered as a factor that improves the material plasticity and extends the time to cracking.

CONCLUSION

We showed that both twinning and plastic deformation in bismuth single crystals subjected to thermal cycling develop in stages.

REFERENCES

1. V. M. Finkel', A. P. Korolev, A. M. Savel'ev, and V. A. Fedorov, Fiz. Met. Metalloved. **48**, 415 (1979).
2. V. A. Fedorov, V. M. Finkel', and V. P. Plotnikov, in *Interaction of Lattice Defects and Properties of Metals* (Tul'sk. Politekh. Inst., Tula, 1979), pp. 143–148.
3. V. M. Finkel', *Physics of Destruction* (Metallurgiya, Moscow, 1970).
4. O. M. Ostrikov, Fiz. Met. Metalloved. **86** (6), 106 (1998).
5. O. M. Ostrikov, Fiz. Met. Metalloved. **87** (1), 94 (1999).
6. O. M. Ostrikov, Inzh.-Fiz. Zh. **72**, 592 (1999).
7. V. I. Bashmakov and V. P. Soldatov, Fiz. Met. Metalloved. **16**, 768 (1963).
8. V. I. Bashmakov and N. G. Yakovenko, Fiz. Met. Metalloved. **26**, 606 (1968).
9. V. I. Bashmakov and N. G. Yakovenko, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 1, 48 (1969).
10. V. I. Bashmakov, M. E. Bosin, F. F. Lavrent'ev, and I. I. Papirov, Probl. Prochn., No. 1, 80 (1974).
11. V. I. Bashmakov, M. E. Bosin, and P. L. Pachomov, Phys. Status Solidi A **9**, 69 (1972).
12. F. F. Lavrent'ev, Fiz. Met. Metalloved. **18**, 428 (1964).

Translated by B. Malyukov

Technical Physics, Vol. 46, No. 9, 2001, pp. 1202–1204. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 71, No. 9, 2001, pp. 140–142.

BRIEF COMMUNICATIONS

Experiments with Pulsed High-Voltage Discharge in the Atmosphere

Yu. Ya. Volkolupov, M. A. Krasnogolovets, M. A. Ostrizhnoĭ, and V. I. Chumakov

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

Received October 10, 2000

Abstract—The characteristics of a spark atmosphere discharge initiated at an interelectrode voltage U = 500 kV and a discharge initiated with a thermally bursting wire are studied. © *2001 MAIK "Nauka/Interperiodica".*

A high-voltage atmosphere discharge excited with a high-voltage (several hundred kilovolts) pulse generator produces high-power radiation whose spectrum lies in the rf and optical bands. The visualization of the discharge helps to determine the characteristics of the radiation, to indirectly find the features of the discharge, and to evaluate the optimal parameters of experimental facilities. This paper reports the characteristics of a high-voltage atmosphere discharge initiated in the spark gap of the generator at U = 500 kV and a discharge initiated by a generator-terminating thermally bursting wire (BW). BW pulses longer than 1 μs were obtained, and the discharge current waveform as a function of load, as well as the radiation energy density as a function of BW material, were studied.

In the experiment, we used a column-type Arkadyev–Marx pulse voltage generator (PVG) with a 10-stage voltage multiplier [1]. The maximum output voltage was 600 kV at an initially accumulated energy of 3.8 kJ. The atmosphere discharge was initiated between the point electrodes of the PVG (the breakdown conditions) or by means of the BW with the electrodes removed. The discharge current was measured with a Rogowski loop. The optical radiation of the discharge was detected in a wide range of 0.3–10.2 μm by an IMO-2N device and in individual spectral bands with a set of optical filters.

Figure 1 shows the waveforms of the PVG current. When the spark is initiated between the point electrodes, damped oscillations with a frequency of ≈800 Hz are observed. The loss Δ in the equivalent oscillatory circuit of the discharge is small, hence, the slow damping of the current amplitude oscillations. Such a behavior is typical of PVG-initiated discharges [1, 2].

The time behavior of the current in the case of the BW is significantly different (Figs. 1b, 1c). The initial pulse of the current is higher, and the oscillations decay noticeably faster. This behavior is inconsistent with the linear model of oscillatory circuit, can be described only in terms of an equivalent circuit including nonlinear resistor and capacitor (see Fig. 1). The effect of the BW parameters on the current is seen in Figs. 1b and 1c. Using a 0.04-mm-diameter tungsten BW instead of a copper one results in a still shorter current pulse and increases the pulse amplitude. In addition, the discharge becomes aperiodic.

The PVG discharge is illustrated in Fig. 2. The spark develops as a typical single- or multichannel process whose shape and sizes depend on electrode configuration, gap width, and applied voltage [3]. The use of the BW stabilizes the shape of the discharge area and makes it possible to significantly extend the discharge

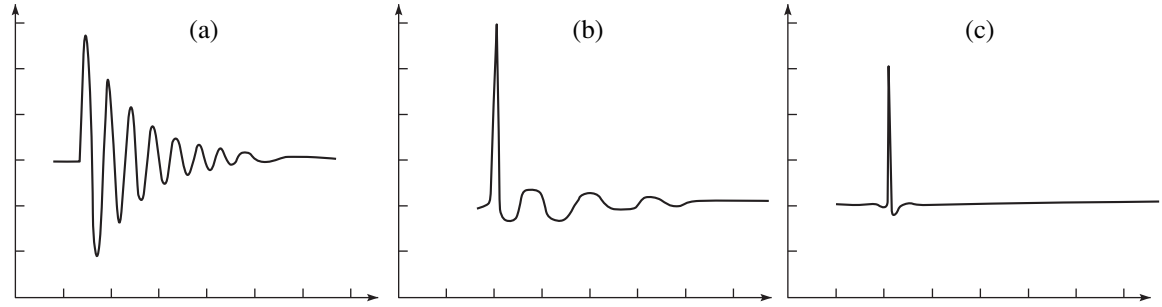


Fig. 1. Waveforms of the current generated by the PVG discharge: (a) point electrodes, (b) copper BW, and (c) tungsten BW. Vertical and horizontal scale divisions are (a) 2 kA and 10 μs, (b) 5 kA and 5 μs, and (c) 10 kA and 5 μs, respectively.

1063-7842/01/4609-1202$21.00

gap. The interelectrode spacing in Figs. 2a and 2b is 25 cm. In Figs. 2c and 2d, the wire is 124 cm long.

Photography using optical filters allows one to observe the spatial temperature distribution in the discharge channel and to estimate the sizes of individual temperature regions [4]. As a rule, four temperature regions of the discharge are distinguished. The highest-temperature region of a diameter no more than 1 mm is at the center (Figs. 2a, 2c). The plasma coat adjacent to the high-temperature region has a transverse size on the order of the electrode diameter. The outer tubular shell coaxial with the high-temperature channel is about 1 mm thick. Some regions of the discharge channel are surrounded by asymmetric low-temperature shells. When viewed through a low-pass filter, the discharge region appears uniform and the regions with different temperatures are not distinguished (Fig. 2d).

Table 1 lists the linear energy densities of the discharge radiation in various frequency bands. The difference in the energy of the discharge radiation in the case of BWs having different diameters and made of different materials shows that parameters of the plasma in the discharge channel depend predominantly on the wire rather than on the atmosphere.

The BW energy introduced into the discharge channel is estimated from the Joule law

$$\frac{\partial Q}{\partial t} = \frac{1}{\sigma} j^2, \tag{1}$$

where Q is the specific energy being released in the wire, σ is the wire conductivity, and j is the current density in the wire.

The temperature dependence of the metal conductivity at temperatures up to the temperature of vaporization can be approximated by the expression [5]

$$\sigma = \frac{\sigma_0}{1 + \beta Q}, \tag{2}$$

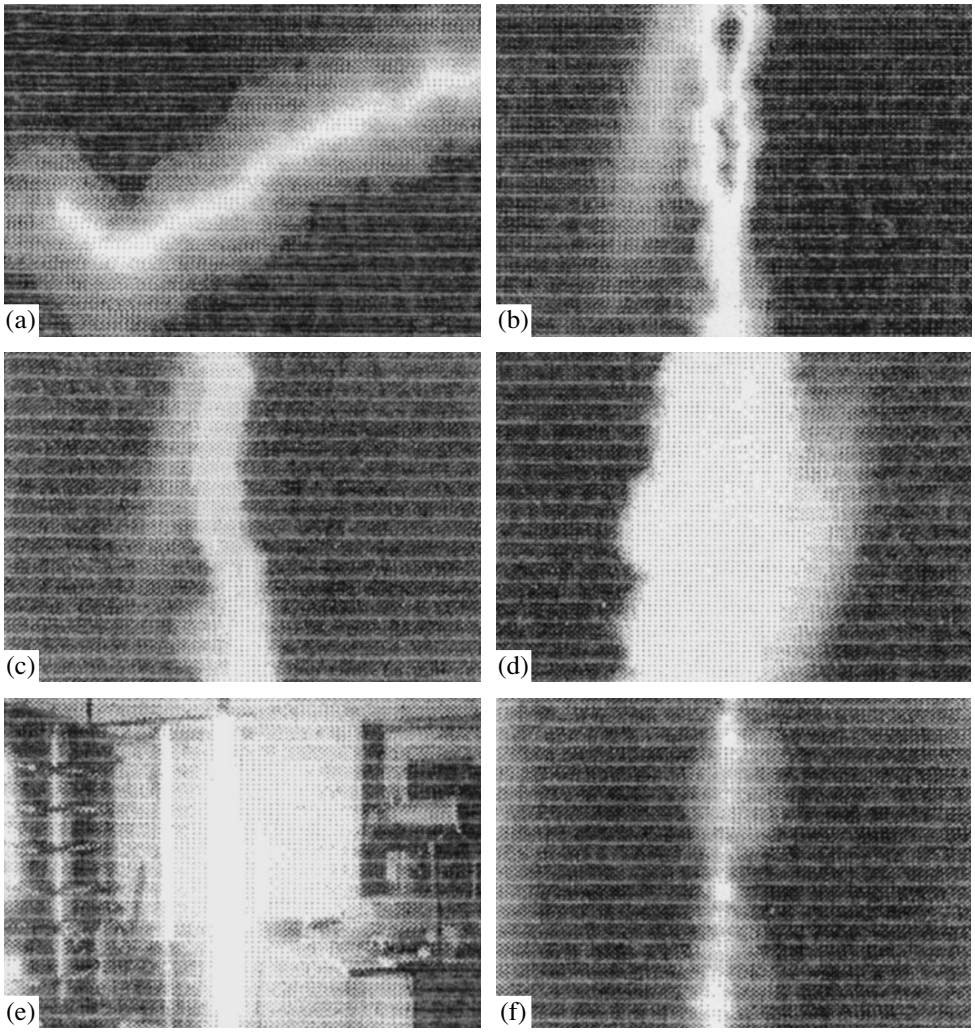


Fig. 2. Discharge as observed through the optical filters: (a) spark discharge, ZhZS-17 filter; (b) spark discharge, UFS-8 filter; (c) spark discharge, TS-10 filter; (d) bursting copper wire, KS-17 filter; (e) bursting tungsten wire, NS-10 filter; and (f) bursting tungsten wire, UFS-8 filter.

Table 1

PVG	Gap width δ, cm	Linear energy density of radiation, mJ/cm			
		without filter	UFS-6 filter	FS-6 filter	KS-17 filter
Interelectrode spacing[1]	25	3.5	1.5	0	0
BW/Cu[2]	124	4.0	0	0	0
BW/W[2]	124	7.5	0	1.0	1.5

[1] Data taken at a distance of 0.75 m.
[2] Data taken at a distance of 1.35 m.

Table 2

Material	Wire length, m	Wire radius, μm	Conductivity, 10^6 $(\Omega\ \mathrm{m})^{-1}$	Heat of sublimation, 10^5 J/kg	Temperature coefficient $\times 10^9$	Integral of current, 10^{17} $A^2\ s/m^4$	Energy released in the wire, J
Al	1.25	50	39.2	105	2.15	1.09	1801
Cu	1.25	50	63.3	58.2	1.31	1.95	4164
W	1.25	20	18.2	–	–	–	–

where σ_0 is the metal conductivity at the zero temperature and β is the temperature coefficient of conductance.

Substituting (2) into (1), we obtain

$$1 + \beta Q = \exp\left(\frac{\beta}{\sigma_0}\int_0^t j^2 dt\right). \qquad (3)$$

With (3), the energy being released in the wire becomes

$$W_1 = Q\pi r^2 l_1 = \frac{\pi r^2 l_1}{\beta}\left[\exp\left(\frac{\beta}{\sigma_0} I\right) - 1\right], \qquad (4)$$

where r is the wire radius, l_1 is the wire length, and $I = \int_0^t j^2\, dt$ is the integral of the current (see, e.g., [5]).

Table 2 lists the energies introduced into the discharge according to (4). It is seen that the conditions for the wire thermal burst are met: $W_1 \cong (2–3)W_2$, where W_2 is the heat of sublimation of the material [6].

Thus, the technique for visualizing high-current optical discharges can be applied to high-voltage discharges in the atmosphere. A benchmark for investigating an atmospheric plasma generated by a spark in wide gaps or by a burst of a metal wire is developed. Our results can be used for shortening high-voltage pulses that generate short-pulse radiation and in charged particle accelerators that excite high-power ultrashort pulses.

REFERENCES

1. S. M. Smirnov and P. V. Terent'ev, *High-Voltage Pulse Generators* (Énergiya, Moscow, 1964).
2. A. M. Asner, *Stosspannungs-Mestechnik* (Springer-Verlag, Berlin, 1974; Énergiya, Moscow, 1979).
3. J. M. Meek and J. D. Craggs, *Electrical Breakdown of Gases* (Clarendon, Oxford, 1953; Inostrannaya Literatura, Moscow, 1960).
4. V. I. Chumakov, M. A. Ostrizhnoĭ, Yu. A. Volkolupov, *et al.*, Radiotekhnika **115** (2000).
5. H. Knoepfel, *Pulsed High Magnetic Fields* (North-Holland, Amsterdam, 1970; Mir, Moscow, 1972).
6. N. N. Stolovich, V. G. Maksimov, and N. S. Minitskaya, Zh. Tekh. Fiz. **44**, 2132 (1974) [Sov. Phys. Tech. Phys. **19**, 1321 (1975)].

Translated by A. Khzmalyan