
ATOMS, SPECTRA,
RADIATION

A Regression Model for a Displaced Atom Cascade under Ion Sputtering of Solids

V. A. Volpyas and P. M. Dymashevskii

St. Petersburg State University of Electrical Engineering, St. Petersburg, 197376 Russia

e-mail: thinfilm@eltech.ru

Received October 18, 2000; in final form, February 6, 2001

Abstract—A regression model of an ion–atom collision cascade (and eventually of a displaced atom cascade) resulting under ion sputtering of amorphous and polycrystalline materials is developed. The model allows the description of the elastic scattering of atomic particles in a solid using various particle interaction potentials. Based on this model, we calculate the sputtering yields for materials with atomic numbers $Z_a = 22\text{--}79$ and the sputtering rates for several multicomponent targets. The results of statistical simulation within the model developed are compared with experimental data for the sputtering of amorphous and polycrystalline materials. It is shown that our model fits the experimental data up to the statistical error and adequately characterizes the ion sputtering process. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

The sputtering of single-component amorphous and polycrystalline targets by ion bombardment is the most-studied field of the physics of material sputtering [1]. It is assumed that the direct knocking-out of atoms from their equilibrium positions in a solid prevails in this case. The available models for ion sputtering of single-component rely on the Sigmund mechanism of physical sputtering by fast ions or atoms [2]. Within this mechanism, the density of a cascade of moving displaced target atoms is described by the distribution function that meets the linearized Boltzmann equation well known from the particle transfer theory. Given the initial velocities and positions of atoms that are decelerated by means of random collisions in an infinite medium, a solution of this equation makes it possible to calculate the atom flux through an arbitrary plane. By taking the target surface as this plane and introducing the parameter E_b , the atom–target binding energy, one can determine the escape probability of the displaced atoms, as well as their energy and angular distributions. An elementary event here is a cascade of ion–atom collisions, which causes a displaced atom cascade and makes surface atoms leave the target under certain conditions. Such a sputtering mechanism is considered to be the most probable at ion energies to 1–3 eV and applies to the ion–plasma sputtering of single-components materials.

Today, due to the progress in microelectronics technology, investigations into ion and ion–plasma sputtering of multicomponent materials are becoming more and more urgent. To these materials, sputtering yield as a parameter characterizing the process cannot be applied because of a depleted layer formed on the surface. In this case, the efficiency of sputtering should be

described in terms of the target sputtering rate along the depth or a change in the target weight.

The statistical simulation of particle scattering upon ion–atom collisions [3–8] allows researchers to quantitatively estimate the sputtering rate of multicomponent targets. However, the numerical simulation of an ion–atom collision cascade and the entire cascade of displaced atoms requires the computer RAM to have an extremely large capacity so as to store huge data arrays of time-varying parameters in different phase spaces (coordinates of cascade particles, their momenta, free path lengths for cascade particles of each sort, etc.). Such a calculation is impossible even with modern computers.

As a reasonable trade-off in this situation, we suggested to mathematically simulate a cascade of ion–atom pair collisions with the subsequent description of the bombarding ion trajectory and the energy imparted to target atoms. In essence, this simulation studies individual cascade chains and extends the results to the entire atom–atom collision cascade. Such an approach combines the easy computing procedure (the entire cascade of pair ion–atom collisions and its individual chains are simulated by the Monte Carlo method) and the correct generalization of the simulation results for the displaced atom cascade by regression analysis. This allowed us to develop a regression model for an ion–atom collision cascade taking place upon ion sputtering of amorphous and polycrystalline single- and multicomponent materials.

REGRESSION MODEL OF ION–ATOM COLLISION CASCADE

Consider the form and the application of our model in detail. The trajectory of a bombarding ion in the tar-

get and momenta gained by target atoms as a result of pair ion–atom collisions are calculated by the Monte Carlo statistical simulation method adapted to particle scattering during ion–atom collisions [9]. When describing the elastic interaction of atomic particles in solids, one must bear in mind the empirical rule according to which collisions causing an appreciable scattering occur at distances on the order of half the equilibrium spacing between adjacent atoms. With so small distances, long-range attracting forces, responsible for binding forces in solids, can be neglected. By order of magnitude, these distances equal the sizes of colliding atomic particles. Therefore, pair ion–atom collisions in solids are usually simulated within the model of quasi-rigid balls [10], in which the elastic parameters involved in atomic particle scattering are calculated with real potentials of interatomic interaction. Note that this model typically provides relatively high rates of statistical simulation. For a potential of interatomic interaction, we chose the Born–Mayer potential modified [3, 4] for a wide range of colliding partners ($Z_a = 2–80$).

In a similar way, the trajectories of target atoms of various sort displaced by a bombarding ion (primarily displaced atoms) are simulated and their energy losses in each event of pair atom–atom collision with secondary displaced target atoms are calculated. The trajectories of various branches of the cascade of the primarily displaced atoms are simulated until their energy becomes smaller (because of atom–atom collisions) than the binding energy between atoms of certain sort or until the trajectories cross the target surface. The ratio of the number of displaced atoms in a cascade branch simulated that cross the surface to the number of bombarding ions defines the probability of these atoms escaping the surface, P_{mc} . Its value depends on the atomic weight, atomic number, and initial energy of a bombarding ion, as well as on the atomic weight,

atomic number, binding energy, and concentration of the target components.

As was noted, the simulation of subsequent branches of the displaced atom cascade is difficult because of the limited capabilities of modern computers. However, regularities obtained in simulating ion–atom cascade branches can be extended to all subsequent branches of the displaced atom cascade by using regression analysis. Clearly, the branches of the primarily displaced (by a bombarding ion) atoms involve a greater number of collisions and a higher initial energy than the subsequent atom–atom branches. Therefore, the escape probability for atoms crossing the surface in each of the subsequent atom–atom branches is lower than the simulated value P_{mc} for the primarily displaced atoms. The smaller the initial energy and the number of atom–atom collisions in any of the secondary branches, the smaller its contribution to the escape probability.

Quantitatively, the ion-induced escape probability P_{is} (for single-component targets, P_{is} coincides with the ion sputtering yield S ; for multicomponent targets, it characterizes the rate of ion sputtering V_{is}) can be calculated from the formula

$$P_{is} = C_n(Z_o)P_{mc}W_{ia}W_{aa}, \quad (1)$$

where $C_n(Z_a)$ is the factor characterizing the shell structure of target atoms, P_{mc} is the simulated escape probability for a primarily displaced atom (it is calculated by the Monte Carlo method), W_{ia} is the factor characterizing the degree of regression of the ion–atom collision cascade, and W_{aa} is the factor that characterizes the degree of regression of the atom–atom collision cascade.

From experimental studies [1], it follows that single-component targets made of Group I materials (Cu, Ag, and Au) feature the highest sputtering yield (Fig. 1). These atoms have the occupied d shell ($Z_d = 18$), which is assumed to define the binding energy of a target atom. To take into account the shell structure of the atoms in calculating the ion-induced escape probability P_{is} , we approximated the factor $C_n(Z_a)$ normalized to Cu atoms ($Z_a = 29$) in (1):

$$C_n(Z_a) = \begin{cases} 0.11Z_a - 2.19 & \text{for } Z_a = 22-29, \\ -0.02Z_a + 1.12 & \text{for } Z_a = 30-41, \\ 0.216Z_a - 8.56 & \text{for } Z_a = 41-47, \\ -0.01Z_a + 1 & \text{for } Z_a = 48-73, \\ 0.148Z_a - 10.49 & \text{for } Z_a = 73-79. \end{cases} \quad (2)$$

The factors W_{ia} and W_{aa} characterizing the degrees of regression of the ion–atom and atom–atom collision cascades were determined with different regression models and checked against experimental data on sputtering yields for a variety of single-component targets and on rates of sputtering multicomponent targets by noble gas ions with various energies. Eventually, we

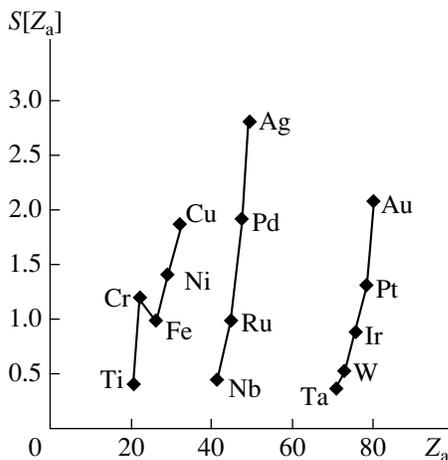


Fig. 1. Sputtering yield in view of the occupancy of the d shell vs. target atomic number Z_a at Ar^+ bombardment ($E = 0.5$ keV) [1].

obtained

$$W_{ia} = \sum_{j=1}^{n_i} \left(\frac{j}{n_i}\right)^{(K_{ri}-0.5j/n_i)}, \quad (3)$$

where n_i is the maximal number of collisions of a bombarding ion in the ion–atom collision cascade, j is the serial number of ion–atom collisions ($1 \leq j \leq n_i$), and K_{ri} is the regression coefficient for a branch of the ion–atom collision cascade, and

$$W_{aa} = \sum_{j=1}^{n_a} \left(\frac{j}{n_a}\right)^{(K_{ra}-0.5j/n_a)}, \quad (4)$$

where n_a is the maximal number of collisions of a displaced atom in the atom–atom collision cascade, j is the serial number of atom–atom collisions ($1 \leq j \leq n_a$), and K_{ra} is the regression coefficient for a branch of the atom–atom collision cascade.

The regression coefficients K_{ri} and K_{ra} were determined by comparing the energy dependence of the sputtering yield $P_{mc}(E_i)$ obtained from the statistical simulation of the trajectories of the ions and primarily displaced atoms with experimentally found energy dependences of the sputtering yield for Ti, Cu, Nb, Ag, Ta, and Au targets subjected to noble gas ion bombardment [1]. Representing K_{ri} as a function of the atomic number Z_i of a bombarding ion and its energy E_i , and K_{ra} as a function of the atomic number of the target Z_a and the energy E_i of the bombarding ion in the form

$$\begin{aligned} K_{ra}(Z_a, E_i) &= A(Z_a)E_i + B(Z_a), \\ K_{ri}(Z_i, E_i) &= C(Z_i)E_i + D(Z_i), \end{aligned} \quad (5)$$

one can generalize the results and determine the numerical values of the expansion coefficients $A(Z_a)$, $B(Z_a)$, $C(Z_i)$, and $D(Z_i)$. The dependences $A(Z_a)$ and $B(Z_a)$ for atoms with $Z_a = 22-79$ and Ar^+ ions with energies $E_i = 250-1000$ eV are depicted in Fig. 2.

With such a representation, for the sputtering of targets with $Z_a < 80$ by Ar^+ ions with energies $E_i = 100-1000$ eV, K_{ra} and K_{ri} are given by

$$\begin{aligned} K_{ra}(Z_a, E_i) &= (-2.37 \times 10^{-5}Z_a - 0.0016)E_i \\ &+ (0.0705Z_a + 0.3582), \\ K_{ri}(Z_i, E_i) &= 2.0. \end{aligned} \quad (6)$$

Thus, by calculating the regression coefficients for the cascade branches K_{ri} and K_{ra} as functions of given Z_i and E_i and Z_a and E_i , respectively, the factors W_{ia} and W_{aa} , which characterize the degrees of regression for the ion–atom and atom–atom cascades, can be found from (3) and (4). Then, having calculated the escape probability P_{mc} (the probability of the primarily displaced atom leaving the surface) by the Monte Carlo method and having taken into account the coefficient

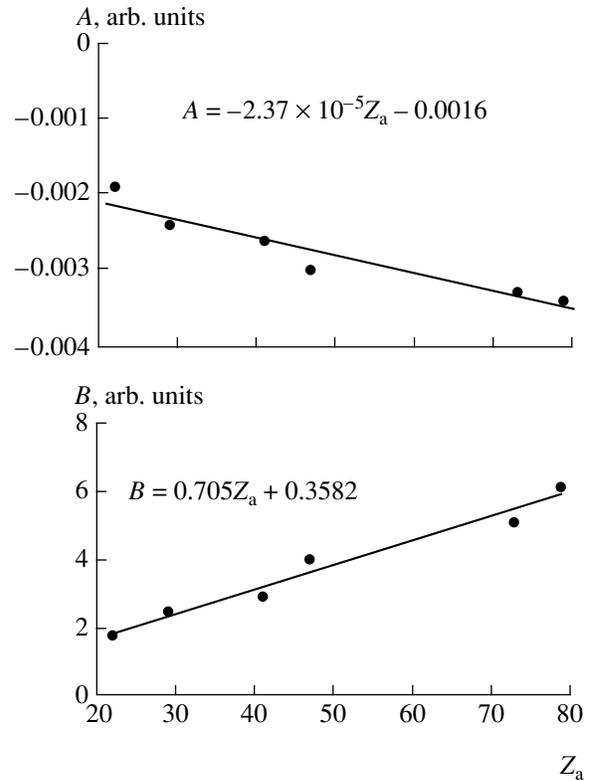


Fig. 2. Expansion coefficients A and B vs. Z_a .

$C_n(Z_a)$ [see (2)], characterizing the shell structure of target atoms, one can evaluate the escape probability P_{is} (the probability of target atoms leaving the surface by the action of a bombarding ion) given by (1). For single-component targets, P_{is} coincides with the ion sputtering yield S ; and for multicomponent targets, with the rate of ion sputtering V_{is} . Note that in the latter case, V_{is} is governed by the sputtering rate of the component having the lowest P_{is} .

DISCUSSION

Our regression model of ion–atom and atom–atom collision cascades taking place during ion sputtering allows the calculation of the sputtering yields and rates for a wide variety of targets and bombarding ions. Theoretically, the rate of sputtering single-component targets can be estimated from the sputtering yield S , which is found within the Sigmund model or with the empirically refined Matsunami formula [11]. However, as follows from the comparison with experimental data [1], the Sigmund model applies at relatively high energies of bombarding ions ($E_i \geq 1$ keV), which are of minor importance in ion–plasma sputtering problems. Comparing the results of the statistical simulation in terms of our model with the experimental data for sputtering amorphous and polycrystalline single-component materials shows that our model fairly accurately fits the experimental data (within a statistical experimental

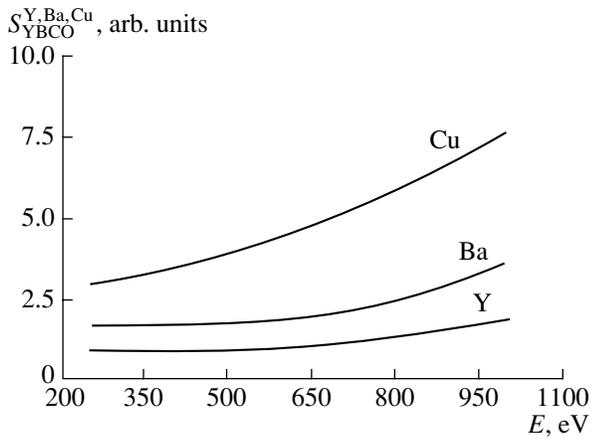


Fig. 3. Sputtering yield for the components of the YBCO ceramic vs. E within the regression model.

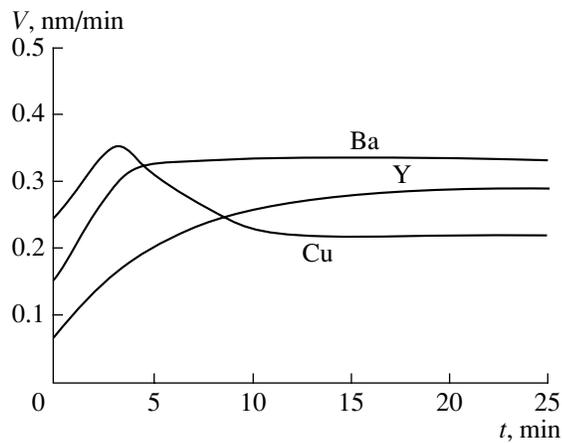


Fig. 4. Sputtering rate V of the YBCO ceramic vs. presputtering time t . Sputtering by Ar^+ ions. The discharge voltage and current density are 160 V and 4 mA/cm², respectively.

error of 10% [1]) for the sputtering yields and rates in a wide range of bombarding ion energies.

Sputtering of multicomponent targets, specifically, complex oxides of high-temperature superconductors, is a more complicated process than sputtering of single-component targets. Experiments on the sputtering of multicomponent targets [12] show that at the early stage of the process, the composition of the sputtered atomic flux differs from the stoichiometric composition of the target and varies with time. At this stage, the composition of the target surface is other than the stoichiometric composition of the target volume. Due to diffusion processes, the thickness of the nonstoichiometric layer may be about 10 nm. During the process, the composition of the sputtered atom flux becomes constant and coincident with the stoichiometric composition of the target; that is, the sputtering process becomes steady-state. It is known that the greater the bombarding ion flux, i.e., the higher the target sputtering rate, the faster the steady-state regime is set. However, in a number of processes of thin-film evaporation,

the sputtering rate is limited by the growth dynamics of a film on a substrate. Therefore, if the evaporation rate is low, the instant the steady-state conditions are established must be detected with a high accuracy.

By way of example, let us consider the sputtering of the high-temperature superconducting ceramic $\text{YBa}_2\text{Cu}_3\text{O}_7$ by Ar^+ ions and compare our calculations of the sputtering rate with experiments [12]. The calculations show (Fig. 3) that the Y atoms in the YBCO ceramic are the most difficult to sputter: under the conditions of target presputtering, the selective sputtering yield of Y in the $\text{YBa}_2\text{Cu}_3\text{O}_7$ system $S_{\text{YBCO}}^{\text{Y}} = P_{\text{is}}$ [see (1)] is roughly half as large as those of Ba and Cu. This is associated largely with the high binding energy of Y, $E_b = 20$ eV [13]. Thus, the sputtering rate of Y specifies the “sluggishness” of the YBCO target sputtering (Fig. 4), while the Ba and Cu atoms have time to adapt to the escape of the Y atoms from the surface. When ions or atoms bombard targets of complex composition, the kinetic energy of the incident particles is unequally distributed among the components, since atoms with various atomic weights and binding energies participate in the collision cascades. As the energy of incident ions grows, sputtering takes place via collisions between atoms that gained an extra energy from the incident ions rather than directly via ion–atom collisions.

REFERENCES

1. *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1981, 1983; Mir, Moscow, 1984, 1986), Vols. 1, 2.
2. P. Sigmund, *Phys. Rev.* **184**, 383 (1969).
3. A. A. Abrahamson, *Phys. Rev.* **178** (3), 76 (1969).
4. V. I. Gaydaenko and V. K. Nikulin, *Chem. Phys. Lett.* **7**, 360 (1970).
5. P. K. Petrov, V. A. Volpyas, E. K. Hollmann, *et al.*, *Vacuum* **48**, 669 (1997).
6. V. A. Volpyas, E. K. Hollmann, D. A. Plotkin, and V. I. Goldrin, *Vacuum* **51**, 227 (1998).
7. V. A. Vol'pyas, E. K. Gol'man, and M. A. Tsukerman, *Zh. Tekh. Fiz.* **66** (4), 16 (1996) [*Tech. Phys.* **41**, 304 (1996)].
8. V. A. Volpyas and A. B. Kozyrev, *Physics of Weakly-Ionized Plasma (Monograph)* (Skladen', St. Petersburg, 1997).
9. P. K. Petrov, V. A. Volpyas, and R. A. Chakalov, *Vacuum* **52**, 427 (1999).
10. V. A. Vol'pyas and E. K. Gol'man, *Zh. Tekh. Fiz.* **70** (3), 13 (2000) [*Tech. Phys.* **45**, 298 (2000)].
11. Y. Yamamura, N. Matsunami, and N. Iton, *Radiat. Eff.* **71**, 65 (1983).
12. V. A. Volpyas, E. K. Hollmann, and V. I. Goldrin, in *Abstracts of the 10th International School on Vacuum, Electron and Ion Technologies, Varna, Bulgaria, 1997*, p. 18.
13. A. L. Pivovarov, S. P. Chenakin, *et al.*, *Poverkhnost* **35** (11), 31 (1989).

Translated by V. Isaakyan

On Acoustic Radiation Accompanying Vibrations of a Charged Droplet

A. I. Grigor'ev and A. R. Gaibov

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

e-mail: grig@uniyar.ac.ru

Received December 4, 2000

Abstract—The dispersion equation relation for the spectrum of capillary oscillations of a charged droplet in a compressible ambient has been derived. It has been shown that such oscillations in the case of droplets, whose sizes and charges are typical of drops in clouds, fogs, or raindrops, cause the generation of sonic and ultrasonic waves in the ambient. An expression for the total intensity of the acoustic radiation generated by a single droplet has been obtained. © 2001 MAIK "Nauka/Interperiodica".

1. Investigation of the interaction between capillary oscillations of a charged drop and a field of acoustic oscillations generated in an ambient compressible medium is of interest in a wide area range of scientific problems and practical applications, from the peculiarities of sound propagation through clouds and fogs to the problems of ultrasonic coagulation of liquid-drop systems and acoustic or electrostatic levitation of large drops in advanced technologies for producing high-purity materials or high-precision measurements of physicochemical properties (see, for instance, [1–8] and references therein). Nevertheless, some issues associated with the problem in question still remain unclear. For example, in most considerations of the applications of the problem of interaction between acoustic waves and the liquid-drop systems, the drops were only assumed to be either sound-scattering objects without internal degrees of freedom or as sources or sinks of water vapor. At the same time, it is known [2] that the frequencies of capillary oscillations of water drops considered as an ideal liquid, when the droplets have sizes typical of clouds or fog (that is, from a few micrometers to several tens of micrometers), correspond to ultrasonic range oscillations, and for raindrops (with the of a radius larger than 250 μm) coincide with the frequencies correspond to of audible sound waves. If a drop is viscous, has an electrical charge, and the equalization rate of its the electrical potential over the drop is finite, then its motions relative to the ambient accompanied by deviations of its shape from spherical will cause an additional drift shift of the frequency spectrum of the capillary oscillations toward a region of lower acoustic frequencies [9–12]. The above-said means that liquid drops in natural liquid-drop systems are capable of both absorbing and generating sound or ultrasound. This is confirmed by the known distortion of the frequency spectrum and strong attenuation of sound by fog and also by the practice of reducing the high-frequency noise abatement of jet

engines in a high-frequency band of the sound spectrum by injection of a liquid-drop aerosol into the combustion chamber.

2. We shall solve a problem on of acoustic radiation of a vibrating drop with the radius R of an incompressible ideal electroconductive liquid having density ρ_1 and surface tension γ and carrying an electrical charge Q . We assume the ambient medium to be ideal, compressible, and specified by the sound speed V and density ρ_2 .

We write the equation for the media interface perturbed by the capillary wave motion in the form

$$r(\Theta, t) = R + \xi(\Theta, t),$$

where $\xi(\Theta, t)$ is the small perturbation of the drop surface and $|\xi| \ll R$.

We assume the wave motions in the drop and the ambient to be of a potential type with the velocity potentials $\psi_1(\mathbf{r}, t)$ and $\psi_2(\mathbf{r}, t)$, respectively.

The mathematical formulation of the problem of calculation of the spectrum of capillary oscillations in the system described and the intensity of the generated acoustic radiation induced by them, in an linear approximation linear with respect to ξ/R , has the form [13]

$$\Delta\psi_1 = 0, \quad (1)$$

$$\frac{1}{V^2} \frac{\partial^2 \psi_2}{\partial t^2} - \Delta\psi_2 = 0, \quad (2)$$

$$r = R: \frac{\partial \psi_1}{\partial r} = \frac{\partial \psi_2}{\partial r}, \quad (3)$$

$$\frac{\partial \xi}{\partial t} = \frac{\partial \psi_1}{\partial r}, \quad (4)$$

$$\begin{aligned} \Delta p - \rho_1 \frac{\partial \psi_1}{\partial t} + \rho_2 \frac{\partial \psi_2}{\partial t} + F_q(\Theta, t) \\ = \gamma \left[\frac{2}{R} - \frac{1}{R^2} (2 + \hat{L}) \xi \right], \end{aligned} \quad (5)$$

$$r = 0: |\psi_1| < \infty, \quad (6)$$

where Δp is the pressure differential across the drop-media interface, $F_q(\Theta, t)$ is the pressure exerted by of the electrical field at on the interface, and \hat{L} is the angular component of the Laplacian in spherical coordinates.

In addition, let us suppose the potential $\psi_2(\mathbf{r}, t)$ at infinity would meet Sommerfeld's radiation condition

$$r \rightarrow \infty: \frac{\partial \psi_2}{\partial r} + ik\psi_2 = o\left(\frac{1}{r}\right). \quad (7)$$

We assume the dependence of potentials $\psi_1(\mathbf{r}, t)$ and $\psi_2(\mathbf{r}, t)$ on time to be periodical: $\psi_1, \psi_2 \sim \exp(i\omega t)$. In this case, the wave Eq. (2) transforms into Helmholtz's equation

$$\Delta \psi_2 + k^2 \psi_2 = 0, \quad k \equiv \omega/V. \quad (8)$$

3. Excluding from consideration the radial centrosymmetrical motion of the drop (as impossible in an incompressible liquid) and the translational (which is not realized in this coordinate system) motions of the drop, we shall seek solutions of Eqs. (1) and (2) for the potentials $\psi_1(\mathbf{r}, t)$ and $\psi_2(\mathbf{r}, t)$ in the form

$$\begin{aligned} \psi_1(\mathbf{r}, t) = \sum_{n=2}^{\infty} A_n r^n P_n(\mu) \exp(i\omega t), \\ \mu \equiv \cos \Theta, \end{aligned} \quad (9)$$

$$\psi_2(\mathbf{r}, t) = \sum_{n=2}^{\infty} B_n h_n^{(2)}(kr) P_n(\mu) \exp(i\omega t), \quad (10)$$

where $P_n(\mu)$ is the Legendre polynomials, and $h_n^{(2)}(x)$ is the second Hankel's spherical function.

Substituting formulas (9) and (10) into Eq. (3) we relate the coefficients A_n with B_n by the expression

$$r = R: B_r = \frac{A_n n R^{n-1}}{k[\partial h_n^{(2)}(kr)/\partial(kr)]}. \quad (11)$$

Differentiating the dynamic boundary condition (5) with respect to time and taking into account the kinematic boundary condition (4), we obtain

$$\begin{aligned} -\rho_2 \frac{\partial^2 \psi_2}{\partial t^2} + \rho_1 \frac{\partial^2 \psi_1}{\partial t^2} \\ - \frac{\gamma}{R^2} (2 + \hat{L}) \frac{\partial \psi_1}{\partial t} - \frac{\partial}{\partial t} F_q(\Theta, t) = 0. \end{aligned} \quad (12)$$

For the time derivative of the pressure on the drop surface produced by the electrical field with respect to time we have (see Appendix)

$$\frac{\partial F_q}{\partial t} = \frac{Q^2}{4\pi\epsilon R^6} \sum_{n=2}^{\infty} A_n n(n-1) R^n P_n(\mu) \exp(i\omega t). \quad (13)$$

Substituting formulas (9)–(11) and (13) into Eq. (12), we arrive at a dispersion equation relation for the capillary oscillations in the system under consideration in the following form

$$\begin{aligned} \omega_n^2 = (n-1) \frac{\gamma}{R^3} \left[\frac{Q^2}{4\pi\epsilon\gamma R^3} - (n+2) \right] \\ \times \left[\frac{(\rho_2/\rho_1) h_n^{(2)}(kR)}{kR h_{n-1}^{(2)}(kR) - (n+1) h_n^{(2)}(kR)} - \frac{1}{n} \right]^{-1}. \end{aligned} \quad (14)$$

When differentiating Hankel's spherical function with respect to its argument, we used the formula [14]

$$\frac{d}{dx} h_n^{(2)}(x) = h_{n-1}^{(2)}(x) - \frac{(n+1)}{x} h_n^{(2)}(x).$$

In the case where $V \rightarrow \infty$, that is, if the ambient medium is assumed to be incompressible, expression (14) is reduced to the known dispersion equation [15] for capillary oscillations of an ideal incompressible electroconductive drop immersed in an ideal incompressible dielectric medium. The latter equation has the form

$$\begin{aligned} \omega^2 = \frac{(n+1)n(n-1)}{\rho_2 n + (n+1)\rho_1} \frac{\gamma}{R^3} [(n+2) - W], \\ W \equiv \frac{Q^2}{4\pi\epsilon\gamma R^3}. \end{aligned} \quad (15)$$

4. To elucidate the meaning of the roots of dispersion equation (14), let us write it for the principal mode of the capillary oscillations ($n=2$) of the drop using the relations between Hankel's spherical functions and trigonometric functions [3] as follows

$$h_n^{(2)}(x) = j_n(x) - i y_n(x),$$

$$j_2(x) = \left(\frac{3}{x^3} - \frac{1}{x} \right) \sin(x) - \frac{3}{x^2} \cos(x),$$

$$y_2(x) = \left(-\frac{3}{x^3} + \frac{1}{x} \right) \cos(x) - \frac{3}{x^2} \sin(x).$$

Here $j_2(x)$ and $y_2(x)$ are the spherical cylindrical functions of the first and second kinds, respectively. The dis-

persion equation (14) for the principal mode of the drop oscillations has the form

$$\omega_2^2 = \frac{\gamma}{R^3}(W - 4) \quad (16)$$

$$\times \left[\frac{(\rho_2/\rho_1)[3R\omega/V + i(-3 + R^2\omega^2/V^2)]}{(-9R\omega/V + R^3\omega^3/V^3) + i(9 - 4R^2\omega^2/V^2)} - \frac{1}{2} \right]^{-1}.$$

Numerical analysis of this equation shows that it has five roots. If one takes $\rho_1 = 1 \text{ g/cm}^3$, $\rho_2 = 1.3 \times 10^{-3} \text{ g/cm}^3$, $V = 3.3 \times 10^4 \text{ cm/s}$, $R = 0.01 \text{ cm}$, $\gamma = 72 \text{ dyn/cm}$, $W = 1$ (this corresponds to capillary oscillations in air at atmospheric pressure of a water drop carrying a charge equal to a quarter of the limiting one in the sense of Rayleigh's stability criterion), then the value of the five roots will be: $\omega_2^{(1)} = (2.1 \times 10^4 + i2.5 \times 10^{-12}) \text{ s}^{-1}$; $\omega_2^{(2)} = (-2.1 \times 10^4 + i2.5 \times 10^{-12}) \text{ s}^{-1}$; $\omega_2^{(3)} = (6.5 \times 10^6 + i3.7 \times 10^6) \text{ s}^{-1}$; $\omega_2^{(4)} = (-6.5 \times 10^6 + i3.7 \times 10^6) \text{ s}^{-1}$; $\omega_2^{(5)} = i6 \times 10^6 \text{ s}^{-1}$. The first two roots correspond to the drop oscillations with frequencies approximately equal to those defined by relation (15) at $n = 2$. These oscillations are attenuated very slowly on account of the loss of the energy take-off being spent for the excitation of longitudinal ultrasonic waves in the ambient. These waves have the frequencies $\omega_1^{(1)}$ and $\omega_2^{(2)}$ and are described by relation (10). The second pair of the roots corresponds to quickly attenuated oscillations of the compressible medium in a close vicinity of the drop surface, and the oscillation frequencies are of the same order as the frequencies of natural oscillations of the medium in a volume equal to the drop volume; that is, these roots correspond to a vortex motion of the medium near the drop surface. The fifth root corresponds to fast aperiodic decay of the perturbations of the external medium's density near the drop surface. From the above-said it follows that in view of a study of acoustic oscillations generated by the oscillations of an incompressible drop the first two roots of Eq. (16) are of interest.

To derive an analytical expression for the damping factor η corresponding to the first two roots of Eq. (16), we shall seek a solution of Eq. (16) in the form $\omega = \omega_0 + i\eta$, where, according to the results of the numerical calculations given above, $\eta \ll \omega_0$. Let us substitute $\omega = \omega_0 + i\eta$ into Eq. (16) and, using a linear approximation of various powers of ω with respect to η , equate the imaginary component of the obtained expression to zero. As a result, for the damping factor we find the expression

$$\eta \approx \frac{\rho_2 R^5 \omega_0^6}{81 \rho_1 V^5}. \quad (17)$$

At $\rho_2 \ll \rho_1$, for ω_0 we have $\omega_0^2 \approx 8\gamma/\rho_1 R^3$ [15], and for the damping factor of the principal mode of the capillary oscillations of the drop related to the generation of acoustic waves we obtain

$$\eta \approx \frac{6.3 \rho_2 \gamma^3}{\rho_1^4 R^4 V^5}.$$

Namely, the principal mode of oscillations is associated with inducing the acoustic waves.

5. Let us find the total intensity of the acoustic radiation produced by a vibrating drop using the known expression [13]

$$I = \rho_2 V \overline{\oint v^2} ds, \quad (18)$$

where the upper bar symbolizes the magnitude averaged over the oscillation period, that is, $\overline{v^2}$ is the mean value of the square of the velocities of medium particles in the acoustic wave; integration is performed over a closed surface encompassing the origin of coordinates.

For the integration surface we take a sphere with radius R_0 , where $R_0 \gg \lambda$, in which λ is the oscillation wavelength. Then in expression (10) the terms diminishing quickest with increasing R_0 can be dropped. To do this, let us represent Hankel's spherical function in the form of a series in negative powers of its argument as follows [14]

$$h_n^{(2)}(x) = i^{n+1} x^{-1} \exp(-ix) \times \sum_{m=0}^n \frac{(n+m)!}{m!(n-m)!} (2ix)^{-m}. \quad (19)$$

Now, at large distances r away from the drop (at $kr \gg 1$, when the sum in formula (19) may be replaced by unity), expression (10) for the velocity potential in the ambient, with allowance for formulas (11) and (19), takes the form

$$\psi_2 \approx \sum_{n=2}^{\infty} \frac{i^{n+1} A_n n R^n P_n(\mu)}{kR h_{n-1}^{(2)}(kR) - (n+1) h_n^{(2)}(kR)} \times \frac{1}{kr} \exp[-i(k_n r - \omega_n t)], \quad (20)$$

that is, the dependence of individual partial waves on the radial coordinate is determined by spherical waves. Inasmuch as $\mathbf{v} = \text{Re}[\text{grad}(\psi_2)]$, the velocity, in a linear approximation with respect to r^{-1} , can be represented in the form

$$\mathbf{v} \approx \text{Re} \left\{ - \sum_{n=2}^{\infty} \frac{i^{n+2} A_n n R^n P_n(\mu)}{kR h_{n-1}^{(2)}(kR) - (n+1) h_n^{(2)}(kR)} \times \frac{1}{r} \exp[-i(k_n r - \omega_n t)] \right\} \mathbf{n}_r, \quad (21)$$

that is, in the approximation used, the velocity has only the radial component.

Taking into consideration that $(\text{Re}[z])^2 \equiv |z|^2/2$, we see that to find the square of the velocity of particles in the acoustic wave involved in formula (18), it is necessary to perform the scalar multiplication of relation (21) by its complex-conjugate expression which can be obtained from relation (21) by changing the sign before

the imaginary unity and replacing the second Hankel's spherical functions $h_n^{(2)}$ with the first ones $h_n^{(1)}$, because $h_n^{(2)}$ and $h_n^{(1)}$ are complex conjugated functions. Ultimately, for the mean value of the square of the particle velocities in an acoustic wave we obtain the following expression

$$\overline{v^2} = - \sum_{2n=z}^{\infty} \sum_{m=z}^{\infty} \frac{A_n A_m m n P_n(\mu) P_m(\mu) R^n R^m}{[kR h_{n-1}^{(2)}(kR) - (n+1)h_n^{(2)}(kR)][kR h_{m-1}^{(1)}(kR) - (m+1)h_m^{(1)}(kR)]r^2}. \tag{22}$$

Substituting formula (22) into expression (18) and taking in the account orthogonality of the Legendre polynomials, for the total intensity of the acoustic

radiation under the constant from a vibrating drop of constant volume we obtain

$$I = \pi \rho_2 V \sum_{n=z}^{\infty} \frac{A_n^2 n^2 R^{2n}}{[kR h_{n-1}^{(2)}(kR) - (n+1)h_n^{(2)}(kR)][kR h_{n-1}^{(1)}(kR) - (n+1)h_n^{(1)}(kR)]}. \tag{23}$$

The expression obtained for the intensity of acoustic radiation of a vibrating drop is not quite convenient in practical use because it involves peak values A_n of the velocity potential that are unknown under experimental or natural conditions. To remove them, we express a the perturbation of a spherical drop in the form of a series in the Legendre polynomials

$$\xi(\mathbf{r}, t) = \sum_{n=z}^{\infty} C_n P_n(\mu) \exp(i\omega t). \tag{24}$$

Now, substituting formulas (9) and (24) into the kinematic boundary condition (4), it is not difficult to express unknown peak values A_n in terms of the amplitudes C_n of capillary oscillations of a drop, which can be easily evaluated under experimental and natural conditions, as follows

$$|A_n| \equiv |C_n(\omega_n/nR^{n-1})|. \tag{25}$$

Now, using expression (23) subject to formula (25), it is possible to obtain numerical estimates for various situations encountered in liquid-drop systems of natural origin.

6. Let us evaluate the intensity of acoustic radiation of a weakly charged raindrop ($W = 0.01$) of a minimum possible size with $R = 0.025$ cm (according to [16], the sizes of raindrops vary from 0.025 to 0.25 cm, droplets of smaller sizes pertain to a drizzle, and larger drops falling through air disintegrate because the action of aerodynamic forces overcomes the surface tension; see also [12]). We assume that the drop's oscillations vibrating on account of excitation occur in the principal

mode of its oscillations ($n = 2$) with the amplitude $C_2 = 0.1R$. A reason for the development of oscillations of the drop could be its (the drop's) collisions with smaller and slower flying droplets, as well as the hydrodynamic attraction between drops of comparable sizes (in the last case, the amplitude of the oscillations may be of an order of R). In the case considered, $\gamma = 72$ dyn/cm, $\rho_1 = 1$ g/cm³, $\omega_2 \approx 6.1 \times 10^3$ s⁻¹, $k_2 \approx 0.2$ cm⁻¹, $k_2 R \approx 4.6 \times 10^{-3}$. In addition, we take $\rho_2 = 1.3 \times 10^{-3}$ g/cm³, $V = 3.3 \times 10^4$ cm³. Eventually, from formula (23) and (25) it is not difficult to find

$$I \approx \frac{4\pi\rho_2 V C_2^2 R^2 \omega^2}{|k_2 R h_1^{(2)}(k_2 R) - 3h_2^{(2)}(k_2 R)|^2} 2.2 \times 10^{-15} \text{ erg/s}. \tag{26}$$

Let us note that, for a drop of the size and charge considered above, only the first four modes will radiate in the acoustic range, and their radiation intensity decreases with the mode number growth (by more than an order of magnitude in going from one mode to the next).

Assuming that in a cubic kilometer of space occupied by rain there are 3×10^{14} drops of the sizes chosen above (roughly one drop in 3 cm³), it is not complicated to find that the total intensity of the acoustic radiation is associated with the principal mode ($n = 2$) of capillary oscillations of the drops in a volume of 1 km³ is ≈ 0.66 erg/s at a frequency of $\omega_2 \approx 6.1 \times 10^3$ s⁻¹. Such an intensity corresponds a sound level of ≈ 22 dB at the boundary of a radiating object volume and roughly corresponds to the sound level of a quiet human voice.

The frequency of capillary oscillations of the drop decreases with increasing drop charge according to (15). Simultaneously, k_2 and $k_2 R$ will decrease, and, consequently, the radiation intensity will also decrease. For instance, at $W = 1$, the intensity I at $\omega_2 \approx 5.3 \times 10^3 \text{ s}^{-1}$ will be $\approx 7.6 \times 10^{-16} \text{ erg/s}$, and the sound level at the boundary of a radiating object one cubic kilometer in volume is $\approx 17 \text{ dB}$.

Substitution into formula (26) k_2 expressed in terms of ω_2 and ω_2 expressed in terms of R allows one to find that the intensity I of acoustic radiation of from a drop is inversely proportional to R^{-2} . This means that the acoustic radiation of large drops, other things being equal, will be substantially weaker. For instance, for the largest raindrop with $R = 0.25 \text{ cm}$ and at $W = 0.01$, the intensity of acoustic radiation associated with the principal mode and occurring at a frequency of 190 s^{-1} is barely $\approx 1.8 \times 10^{-17} \text{ erg/s}$. Calculation of the sound level under the same conditions as in the example above gives a very small magnitude of $\approx 1 \text{ dB}$. The human ear does not perceive such a level of sound. The sound produced by the capillary oscillations of large drops is perceived by the human ear only at large amplitudes of the drop's oscillations comparable with the radius R . In this case, the sound level is the same as that produced by the capillary oscillations of droplets with $R = 0.025 \text{ cm}$ in the example above. It should also be noted that, to the range of frequencies of acoustic oscillations perceived by the human ear, there are more than forty of the first frequencies of capillary oscillations of a drop with $R = 0.25 \text{ cm}$; however, their intensity is as low as mentioned above.

It is interesting to note that the dependence of the acoustic radiation sound intensity, induced by a vibrating drop, on the sound speed in a medium V , the medium density ρ_2 , and the frequency ω_2 of the capillary oscillations of a drop has the same form as the dependence on the same parameters of the damping factor η of capillary oscillations of a drop defined by expression (17); namely, $I \sim \eta \sim \rho_2 \omega_2^6 V^5$. This can be easily verified by substituting the expression for k_2 in terms of ω_2 and V into (26). The similarity of these dependencies mentioned above is explained by the fact that the attenuation of capillary oscillations of a drop is associated with the radiation of acoustic waves.

The obtained estimates are also valid for cumuli. As for fogs with a mean droplet radii of about $10 \mu\text{m}$, they can radiate sound waves in an audible wavelength range only if the charges of isolated individual drops are close to the limiting ones in the sense of stability with respect to their self-charge (when $W \rightarrow 4$) [9, 15].

CONCLUSION

The total acoustic radiation produced by capillary oscillations of drops in liquid-drop systems of natural

origin may exceed the hearing threshold of the human ear.

APPENDIX

Let us calculate the partial time derivative with respect to time of the pressure of an electrical field on the surface of an incompressible electroconductive liquid drop perturbed by a capillary wave motion.

The pressure of an electrostatic field with the strength $\mathbf{E}(\Theta, t)$ in a medium with the permittivity ε is defined by the relation

$$F_q(\Theta, t) = \frac{\varepsilon}{8\pi} \mathbf{E}^2.$$

The sought for derivative $\partial F/\partial t$ can be easily represented in the form

$$\frac{\partial F_q(\Theta, t)}{\partial t} = \frac{\varepsilon}{4\pi} \mathbf{E}(\Theta, t) \frac{\partial \mathbf{E}(\Theta, t)}{\partial t}.$$

However, to use this formula in the considered case, it is necessary to find an expression for the strength of the electrical field strength in the vicinity of the surface the spherical droplet perturbed by the wave motion.

We shall seek the potential Φ of the field induced by the disturbed charged droplet in the ambient space, taking into account that the potential should be a harmonic function

$$\begin{aligned} \Delta \Phi &= 0, \\ r \rightarrow \infty: \Phi &\rightarrow 0 \end{aligned} \quad (\text{A1})$$

and satisfy the condition

$$r = R + \xi: \Phi = \text{const} \quad (\text{A2})$$

at the interface.

We represent the potential Φ in the form

$$\Phi = \Phi_0 + \delta\Phi,$$

where Φ_0 is the potential of an unperturbed spherical droplet and $\delta\Phi$ is the additive an addition to the potential arising from the surface perturbation.

Then the problem formulated in (A1) and (A2) takes the form

$$\Delta(\Phi_0 + \delta\Phi) = 0, \quad (\text{A3})$$

$$r = R + \xi: \Phi_0 + \delta\Phi = \frac{Q}{\varepsilon R}. \quad (\text{A4})$$

Due to the smallness of the surface perturbation amplitude ($|\xi| \ll R$), let us expand the boundary condition (A4) in a series in the vicinity of point $\xi = 0$ and, neglecting terms of the second order of smallness, we obtain

$$r = R: \Phi_0 + \delta\Phi + \xi \frac{\partial}{\partial r} \Phi_0 = \frac{Q}{\varepsilon R}. \quad (\text{A5})$$

Now the problem splits: the problem of finding Φ

$$\begin{aligned} \Delta\Phi_0 &= 0, \\ r = R: \Phi_0 &= \frac{Q}{\epsilon R}; \end{aligned} \tag{A6}$$

and the problem of finding $\Delta\Phi$

$$\Delta(\delta\Phi) = 0, \tag{A7}$$

$$r = R: \delta\Phi + \xi \frac{\partial}{\partial r} \Phi_0 = 0. \tag{A8}$$

From Eq. (A8), taking into account that the solution to the problem (A7) is known, $\Phi(r) = Q/r$, it follows

$$r = R: \delta\Phi = -\xi \frac{\partial}{\partial r} \Phi_0 = \xi E_{or} = \xi \frac{Q}{\epsilon R^2}. \tag{A9}$$

We shall seek $\delta\Phi$ in the form of a series

$$\delta\Phi = \sum_n D_n \left(\frac{R}{r}\right)^{n+1} P_n(\mu). \tag{A10}$$

Substituting expression (A10) into formula (A9) gives

$$r = R: \delta\Phi = \sum_n D_n P_n \mu = \xi \frac{Q}{\epsilon R^2}.$$

From this expression it is easy to find the coefficients D_n

$$D_n = \frac{Q}{R^2} \int_0^\pi \xi P_n(\mu) \sin \Theta d\Theta. \tag{A11}$$

Let us note that the coefficients D_n are small values, of the same order as ϵ .

Now let us find the field strength in the drop vicinity

$$r = R + \xi: \mathbf{E} = -\nabla\Phi = -\nabla(\Phi_0 + \delta\Phi).$$

Referring this expression to the interface $r = R + \xi$ and expanding the relation derived into a series in the vicinity of the unperturbed boundary, in the linear approximation with respect to $|\xi|/R$ we find

$$\mathbf{E}|_{r=R+\xi} \approx \left[-\nabla(\Phi_0 + \delta\Phi) - \xi \frac{\partial}{\partial r} \nabla\Phi_0 \right]_{r=R}. \tag{A12}$$

Taking into account that the pressure of the electrical field at the drop surface to within the first-order perturbations is expressed in terms of only of the radial component of the field strength accurate up to infinitesimal quantities of the first order, we write an the expression for the radial component of \mathbf{E} substituting corresponding components of the vectors $\nabla\Phi_0$, $\nabla(\delta\Phi)$, and

$(\partial/\partial r)(\nabla\Phi_0)$ into (A12) as follows

$$\mathbf{E}|_{r=R+\xi} \approx \left[\frac{Q}{\epsilon R^2} + \frac{1}{R} \sum_n D_n (n+1) P_n(\mu) - \epsilon \frac{2Q}{R^3} \right]_{r=R} \mathbf{n}_r.$$

At this stage of consideration, it is possible to also find an expression for the partial time derivative of the field strength at the droplet surface with respect to time, namely

$$\frac{\partial E}{\partial t} \Big|_{r=R+\xi} \approx \left[\frac{1}{R} \sum_n \frac{\partial D_n}{\partial t} (n+1) P_n(\mu) - \frac{2Q}{\epsilon R^3} \frac{\partial \xi}{\partial t} \right]. \tag{A13}$$

The derivative $\partial D_n/\partial t$ can be found from (A11) as follows

$$\frac{\partial D_n}{\partial t} = \frac{Q}{\epsilon R^2} \int_{-1}^1 \frac{\partial \xi}{\partial t} P_n(\mu) d\mu.$$

From (A4) and (A9) it is not difficult to obtain

$$\begin{aligned} r = R: \frac{\partial D_n}{\partial t} &= \frac{Q}{\epsilon R^2} \int_{-1}^1 \frac{\partial \Psi_1}{\partial r} P_n(\mu) d\mu \\ &= \frac{Q}{\epsilon R^2} \sum_m \int_{-1}^1 A_m m R^{m-1} P_m(\mu) P_n(\mu) d\mu \exp(-i\omega t) \\ &= \frac{Q}{\epsilon R^3} \sum_m m A_m \exp(-i\omega t) R^m \int_{-1}^1 P_m(\mu) P_n(\mu) d\mu \\ &= \frac{Q}{\epsilon R^3} n C_n \exp(-i\omega t), \end{aligned} \tag{A14}$$

where $C_n \equiv R^n/A_n$.

Let us also then take into account that, according to (A4),

$$r = R: \frac{\partial \xi}{\partial t} \approx \frac{\partial \Psi_1}{\partial r} = \sum_n C_n P_n(\mu) n \frac{1}{R} \exp(-i\omega t). \tag{A15}$$

Substituting formulas (A15) and (A14) into expression (A13), we finally find

$$\begin{aligned} \frac{\partial E}{\partial t} &= \frac{Q}{\epsilon R^4} \sum_n n(n+1) C_n P_n(\mu) \exp(-i\omega t) \\ &\quad - \frac{2Q}{\epsilon R^4} \sum_n n C_n P_n(\mu) \exp(-i\omega t) \\ &= \frac{Q}{R^4} \sum_4 n(n-1) C_n P_n(\mu) \exp(-i\omega t). \end{aligned}$$

Then we obtain a relation for the sought derivative with respect to time of the field pressure at the drop's surface in the form

$$\begin{aligned} \frac{\partial F_q}{\partial t} &= \frac{\varepsilon}{8\pi} \frac{\partial E^2}{\partial t} = \frac{\varepsilon}{8\pi} 2E \frac{\partial E}{\partial t} \\ &= \frac{\varepsilon}{4\pi} \left(\frac{Q}{\varepsilon R^2} + \frac{1}{R} \sum_n D_n (n+1) P_n(\mu) - \xi \frac{2Q}{\varepsilon R^3} \right) \\ &\quad \times \left(\frac{Q}{\varepsilon R^4} \sum_n n(n-1) C_n P_n(\mu) \exp(-i\omega t) \right). \end{aligned}$$

Let us eliminate terms of the second order of smallness that are proportional to the products $D_n C_n$ and ξC_n as perturbations of the second order. Then

$$\frac{\partial F_q}{\partial t} \approx \frac{1}{4\pi} \frac{Q^2}{\varepsilon R^6} \sum_n n(n-1) C_n P_n(\mu) \exp(-i\omega t).$$

ACKNOWLEDGMENTS

The work was supported by grant 00-15-9925 of the President of the Russian Federation.

REFERENCES

1. N. A. Fuks, *Mechanics of Aerosols* (Akad. Nauk SSSR, Moscow, 1955).
2. J. W. Strutt (Lord Rayleigh), *The Theory of Sound* (Macmillan, London, 1896; Gostekhizdat, Moscow, 1955), Vol. 2.

3. H. L. Green and W. R. Lane, *Particulate Clouds: Dusts, Smokes and Mists* (Spon, London, 1964; Khimiya, Leningrad, 1969).
4. Won-Kyu Phim, Sang Kun Chung, M. T. Hyson, *et al.*, IEEE Trans. Ind. Appl. **IA-23**, 975 (1987).
5. V. Sh. Shagapov, Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana **24**, 506 (1988).
6. L. G. Kachurin, *Physical Principles of Influence on Atmospheric Processes* (Gidrometeoizdat, Leningrad, 1990).
7. P. V. R. Suryanarayana and Y. Bayazitoglu, Phys. Fluids A **3**, 967 (1991).
8. A. L. Yarin, G. Brenn, O. Kastner, *et al.*, J. Fluid Mech. **399**, 151 (1999).
9. A. I. Grigor'ev, Zh. Tekh. Fiz. **55** (7), 1272 (1985) [Sov. Phys. Tech. Phys. **30**, 736 (1985)].
10. S. O. Shiryayeva, M. I. Muniquev, and A. I. Grigor'ev, Zh. Tekh. Fiz. **66** (7), 1 (1996) [Tech. Phys. **41**, 635 (1996)].
11. S. O. Shiryayeva, Zh. Tekh. Fiz. **69** (8), 28 (1999) [Tech. Phys. **44**, 894 (1999)].
12. A. I. Grigor'ev, V. A. Koromyslov, and S. O. Shiryayeva, Zh. Tekh. Fiz. **70** (7), 26 (2000) [Tech. Phys. **45**, 840 (2000)].
13. L. D. Landau and E. M. Lifshitz, *Mechanics of Continuous Media* (Gostekhizdat, Moscow, 1953).
14. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
15. S. O. Shiryayeva, A. I. Grigor'ev, and I. D. Grigor'eva, Zh. Tekh. Fiz. **65** (2), 1 (1995) [Tech. Phys. **40**, 117 (1995)].
16. L. T. Matveev, *Course of General Meteorology. Atmosphere Physics* (Gidrometeoizdat, Leningrad, 1983).

Translated by N. Mende

On the Capillary Vibration and Stability of a Charged Bubble in a Dielectric Liquid

A. N. Zharov and A. I. Grigor'ev

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

e-mail: grig@uniyar.ac.ru

Received December 7, 2000

Abstract—The capillary vibration and stability of a charged bubble against infinitesimal volume and shape perturbations when the bubble is immersed in a viscous incompressible dielectric liquid are studied. The range of physical parameters where the noncentrosymmetric radial and axisymmetric surface motions of the bubble are unstable is found. Asymptotic analytical expressions for the damping constant of the axisymmetric capillary vibration in the low- and high-viscosity approximations are derived. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

Investigation into the capillary vibration and stability of charged bubbles in a liquid is of both scientific and applied interest [1]. Researchers come up against the problem of charged bubbles when studying acoustic and hydrodynamic cavitations, electrical discharge in liquids [2, 3], flotation and electroflotation [4], filtering [5], optical breakdown in liquids [6], bubbling [7], and heat exchange [8]. This problem also arises in the technology of separating salts of heavy metals from water solutions by electrical discharge [9] and in cavitation-assisted fusion of light nuclei [10, 11].

Microbubbles appearing in these applications are often charged. Charging can be associated with different mechanisms. In the case of electrical discharge, a plasma-filled cavity forms. If the characteristic times of diffusion for charge carriers of opposite sign in the plasma are different and smaller than the characteristic time of recombination, a part of the charge will remain on the cavity walls; that is, the cavity will be charged. For bubbles arising in the chamber of an electrohydrodynamic pump, the settling-out of electronegative impurity molecules, which is responsible for electrical-to-mechanical energy conversion [12], seems to be a more plausible mechanism of acquiring an excessive charge. Thus, two basic mechanisms of bubble charging can be distinguished: (1) the settling-out of ions coming from the interior of the bubble to its walls and (2) the settling-out of ions from the surrounding liquid or impurity ions on the bubble walls. In the former case, the amount of charge depends largely on the recombination-to-diffusion time ratio; in the latter, on the capability of the interface to adsorb ions of the liquid or impurity ions. In either case, the amount of charge on the bubble may be significant and affect its stability.

(1) Let a bubble of radius R_0 and charge Q acquired by either of the two ways form in a liquid of density ρ , viscosity ν , and permittivity ϵ . The bubble contains a

perfect gas under a pressure P_{g0} that obeys the polytropic law with a polytrope index γ and also a saturated vapor under a pressure P_v . Let the liquid pressure around the (immobile) bubble be P_∞ and the surface tension at the liquid–gas interface be σ .

In the general case, such a bubble is nonequilibrium and moves under the action of the resulting pressure [13]

$$P(R) = P_v + P_{g0} \left(\frac{R_0}{R} \right)^{3\gamma} + \frac{Q^2}{8\pi\epsilon R^4} - \frac{2\sigma}{R} - P_\infty, \quad (1)$$

where R is the current radius of the bubble executing centrosymmetric vibration. If $P(R) > 0$, the bubble expands; if $P(R) < 0$, it shrinks; and if $P(R) = 0$, the bubble is in the equilibrium state.

We will solve the boundary-value problem of the arbitrary motion of the charged bubble wall in the liquid. It is natural to assume that the motion of the bubble boundary will cause motions in both the liquid and the gas–vapor mixture. However, in view of the actual densities of liquids and gases, the gas motions can be neglected in the first order of smallness. In fact, using the Cauchy–Lagrange integral, we find that the motion of the medium changes the pressure by a value $\delta P \sim \rho \partial \phi / \partial t$, which is proportional to the density of the medium. Since the density of a liquid is three orders of magnitude larger than that of a gas in most cases, the contribution from the change in the gas density to the pressure balance at the interface can be neglected. Therefore, we will consider only the motion of the surrounding liquid. The gas inside the bubble will be considered immobile.

Mathematically, the problem stated is represented by the continuity equation for the liquid [14]

$$\operatorname{div} \mathbf{u} = 0 \quad (2)$$

and Navier–Stokes equation

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla P + \nu \Delta \mathbf{u}, \quad (3)$$

where the parenthesized vectors mean the scalar product.

At the interface, described by the equation

$$F(\mathbf{r}, t) = r - R(t) - \xi(\vartheta, \varphi, t), \quad (4)$$

the boundary conditions

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \mathbf{u} \nabla F = 0, \quad (5)$$

$$\boldsymbol{\tau}(\mathbf{n} \nabla) \mathbf{u} + \mathbf{n}(\boldsymbol{\tau} \nabla) \mathbf{u} = 0, \quad (6)$$

$$P + P_\sigma + 2\rho \nu \mathbf{n}(\mathbf{n} \nabla) \mathbf{u} - P_V - P_g - P_q = 0, \quad (7)$$

should be satisfied. Here, $\boldsymbol{\tau}$ and \mathbf{n} are the vectors tangential and normal to the bubble surface, respectively; P_σ is the pressure due to the surface tension; P_g is the gas pressure in the bubble; and P_q is the electrical pressure.

Set (2)–(7) should be complemented by the equation of state for the gas–vapor mixture in the bubble

$$P(V) = P_V + P_g, \quad (8)$$

condition of the bubble volume constancy during shape variations due to vibration

$$\int_V dV = \frac{4\pi}{3} R^3(t) \quad (9)$$

and condition for the center-of-mass immobility

$$\int_V \mathbf{r} dV = 0, \quad (10)$$

where integration is over the bubble volume.

Recall that we will consider both the radial vibration of the bubble, which is attended by a change in its volume, and the vibration due to the deviation of the bubble shape from spherical with the bubble volume remaining unchanged [see Eq. (9)].

(2) Set (2)–(10) completely describes the motion of a liquid near an immobile gas–vapor bubble; hence, its solution yields the complete spectrum of possible vibrations of the immobile bubble in the liquid. The problem stated is essentially nonlinear; therefore, we will linearize it with respect to the deviation of its shape from spherical, $\xi(\vartheta, \varphi, t)$ and the field velocity in the liquid, $\mathbf{u}(\mathbf{r}, t)$ (both parameters are of the same order of smallness).

Continuity equation (2) is initially linear. The Navier–Stokes equation after linearization takes the form

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{1}{\rho} \nabla P + \nu \Delta \mathbf{u}. \quad (11)$$

The linearization of the kinematic boundary condition yields

$$r = R(t): \frac{dR}{dt} + \frac{\partial \xi}{\partial t} = u_r. \quad (12)$$

Upon linearization, it will suffice to take dynamic boundary condition (6) on the spherical surface of the bubble $r = R(t)$. For this surface, the radial unit vector \mathbf{e}_r serves as the normal vector \mathbf{n} . As the tangential vector $\boldsymbol{\tau}$, one should take successively the polar unit vector \mathbf{e}_ϑ and the azimuth unit vector \mathbf{e}_φ .

For $\boldsymbol{\tau} = \mathbf{e}_\vartheta$, we obtain

$$r = R(t): \frac{\partial u_\vartheta}{\partial r} + \frac{1}{r} \frac{\partial u_r}{\partial \vartheta} - \frac{u_\vartheta}{r} = 0. \quad (13)$$

For $\boldsymbol{\tau} = \mathbf{e}_\varphi$,

$$r = R(t): \frac{\partial u_\varphi}{\partial r} + \frac{1}{r \sin \vartheta} \frac{\partial u_r}{\partial \varphi} - \frac{u_\varphi}{r} = 0. \quad (14)$$

The linearization of the dynamic boundary condition yields

$$r = R(t) + \xi: P + P_\sigma - 2\rho \nu \frac{\partial u_r}{\partial r} - P_V - P_g - P_q = 0. \quad (15)$$

The condition of the constancy of the bubble volume when its shape is other than spherical due to vibration is given by

$$\int_\Omega \xi d\Omega = 0, \quad (16)$$

and the condition of the center-of-mass immobility takes the form

$$\int_\Omega \xi \mathbf{e}_r d\Omega = 0, \quad (17)$$

where Ω is solid angle.

(3) To scalarize the linearized set of equations (2) and (11)–(17), we decompose the vector velocity field into three mutually orthogonal components: potential, poloidal, and toroidal [15]:

$$\mathbf{u} = \mathbf{N}_1 \psi_1(\mathbf{r}, t) + \mathbf{N}_2 \psi_2(\mathbf{r}, t) + \mathbf{N}_3 \psi_3(\mathbf{r}, t), \quad (18)$$

where the projectors \mathbf{N}_i and their Hermitean conjugates \mathbf{N}_i^+ are selected as follows:

$$\begin{aligned} \mathbf{N}_1 &= \nabla, & \mathbf{N}_2 &= \nabla \times \mathbf{r}, & \mathbf{N}_3 &= \nabla \times (\nabla \times \mathbf{r}), \\ \mathbf{N}_1^+ &= -\nabla, & \mathbf{N}_2^+ &= \mathbf{r} \times \nabla, & \mathbf{N}_3^+ &= (\mathbf{r} \times \nabla) \times \nabla. \end{aligned} \quad (19)$$

Then, we can write the orthogonality condition for these operators in the form

$$\mathbf{N}_j^+ \mathbf{N}_i = 0 \quad \text{at } j \neq i. \quad (20)$$

In view of (19), continuity equation (2) takes the

form

$$\mathbf{N}_1^+ \mathbf{u} = 0. \quad (21)$$

Substituting (18) into (21), taking into account orthogonality condition (20), and bearing in mind that $\mathbf{N}_1^+ \mathbf{N}_1 = -\Delta$, we come to the equation for the potential function:

$$\Delta \psi_1 = 0. \quad (22)$$

The solution of (22) for the liquid (at $r > R$) is easily obtained as the series in spherical functions:

$$\psi_1(\mathbf{r}, t) = \sum_{n=0}^{\infty} \sum_{m=-n}^{m=n} \frac{C_{nm}^1(t)}{r^{n+1}} Y_{nm}(\vartheta, \varphi). \quad (23)$$

In expression (23), we separate the zero-mode ($n=0$) terms that correspond to the radial vibration of the bubble boundary and to the changes in the bubble shape that do not alter its volume:

$$\begin{aligned} \psi_1(\mathbf{r}, t) &= \frac{C_{00}^1(t)}{r} \\ &+ \sum_{n=1}^{\infty} \sum_{m=-n}^{m=n} \frac{C_{nm}^1(t)}{r^{n+1}} Y_{nm}(\vartheta, \varphi) = \psi_1^{(r)} + \psi_1^{(s)}, \end{aligned} \quad (24)$$

where $\psi_1^{(r)}$ describes the radial vibration and $\psi_1^{(s)}$ describes the changes in the shape at constant volume.

With regard for (24), expression (18) can be recast as

$$\begin{aligned} \mathbf{u} &= \mathbf{N}_1 \psi_1^{(r)}(\mathbf{r}, t) + \mathbf{N}_1 \psi_1^{(s)}(\mathbf{r}, t) \\ &+ \mathbf{N}_2 \psi_2(\mathbf{r}, t) + \mathbf{N}_3 \psi_3(\mathbf{r}, t). \end{aligned} \quad (25)$$

This form will be used in the subsequent analysis. Before proceeding to the scalarization of the Navier–Stokes equation, we represent the pressure in the liquid as the sum of two terms responsible for the radial vibration [superscript (r)] and for the changes in the shape [superscript (s)] (by analogy with the potential component ψ_1 of the velocity):

$$P = P^{(r)} + P^{(s)}. \quad (26)$$

Let us introduce the function H :

$$\nabla H = \frac{1}{\rho} \nabla P. \quad (27)$$

Assuming that a change in the pressure does not affect the density of the liquid in (27), we find in view of (26)

$$\begin{aligned} H &= \frac{P - P_{\infty}}{\rho} = \frac{P^{(r)} + P^{(s)} - P_{\infty}}{\rho} \\ &= \frac{P^{(r)} - P_{\infty}}{\rho} + \frac{P^{(s)}}{\rho} \equiv H^{(r)} + H^{(s)}. \end{aligned} \quad (28)$$

Substituting (25) into Navier–Stokes equation (11) and using (20) and (27), we arrive at three equations for the scalar functions:

$$H = -\frac{\partial \psi_1^{(r)}}{\partial t} - \frac{\partial \psi_1^{(s)}}{\partial t}, \quad (29)$$

$$\frac{\partial \psi_j}{\partial t} = \nu \Delta \psi_j; \quad j = 2, 3. \quad (30)$$

Substituting (28) into (29) and taking into account the orthogonality of the spherical functions (i.e., equating the terms proportional to various spherical functions) yields

$$P^{(r)} = P_{\infty} - \rho \frac{\partial \psi_1^{(r)}}{\partial t}, \quad (31)$$

$$P^{(s)} = -\rho \frac{\partial \psi_1^{(s)}}{\partial t}. \quad (32)$$

The former equation corresponds to the radial vibration, and the other describes the changes in the shape at constant volume.

The solution of Eq. (30) limited at $r \rightarrow \infty$ has the form [16]

$$\begin{aligned} \psi_j &= \sum_{n=0}^{\infty} \sum_{m=-n}^{m=n} C_{nm}^j k_n \left(\sqrt{\frac{S}{\nu}} r \right) Y_{nm} e^{St}; \\ &j = 2, 3, \end{aligned} \quad (33)$$

where $k_n(z)$ is the modified spherical Bessel function of the 3rd kind and S is the eigenvalue having the dimension of frequency.

To scalarize the boundary conditions, we will take advantage of the expressions for the velocity field components in the spherical coordinate system [15]:

$$\begin{aligned} u_r &= \frac{\partial \psi_1^{(r)}}{\partial t} + \frac{\partial \psi_1^{(s)}}{\partial t} - \frac{1}{r} \Delta_{\Omega} \psi_3, \\ u_{\vartheta} &= \frac{1}{r} \frac{\partial \psi_1^{(s)}}{\partial \vartheta} + \frac{1}{\sin \vartheta} \frac{\partial \psi_2}{\partial \varphi} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \psi_3}{\partial \vartheta} \right), \\ u_{\varphi} &= \frac{1}{r \sin \vartheta} \frac{\partial \psi_1^{(s)}}{\partial \varphi} - \frac{\partial \psi_2}{\partial \vartheta} + \frac{1}{r \sin \vartheta} \frac{\partial}{\partial r} \left(r \frac{\partial \psi_3}{\partial \varphi} \right), \end{aligned} \quad (34)$$

where Δ_{Ω} is the Laplacian angular part.

The substitution of (34) into scalarized kinematic boundary condition (12) yields

$$r = R(t): \frac{dR}{dt} + \frac{\partial \xi}{\partial t} = \frac{\partial \psi_1^{(r)}}{\partial t} + \frac{\partial \psi_1^{(s)}}{\partial t} - \frac{1}{r} \Delta_{\Omega} \psi_3. \quad (35)$$

From Eq. (35), which relates the scalar functions ψ_j and the surface perturbation ξ , and relationships (23)

and (33), it readily follows that ξ can be represented as the infinite series in spherical functions:

$$\xi = \sum_{n=0}^{\infty} \sum_{m=-n}^{m=n} Z_{nm}(t) Y_{nm}. \quad (36)$$

Note that condition (16) (the constancy of the volume during capillary vibration) yields $Z_{00}(t) = 0$ and condition (17) (the center-of-mass immobility), $Z_{1m}(t) = 0$. Therefore, summation in (36) must start with $n = 2$. Since (35) is related to (23) and (33), one can conclude that summation in the series for the functions $\psi_1^{(s)}$ and ψ_3 also must start with $n = 2$. Thus,

$$\xi = \sum_{n=2}^{\infty} \sum_{m=-n}^{m=n} Z_{nm}(t) Y_{nm}, \quad (37)$$

$$\psi_1^{(s)} = \sum_{n=2}^{\infty} \sum_{m=-n}^{m=n} \frac{C_{nm}^1(t)}{r^{n+1}} Y_{nm}(\vartheta, \varphi), \quad (38)$$

$$\psi_3 = \sum_{n=2}^{\infty} \sum_{m=-n}^{m=n} C_{nm}^3 k_n \left(\sqrt{\frac{S}{V}} r \right) Y_{nm} e^{St}. \quad (39)$$

In view of (37)–(39) and the orthogonality of the spherical functions, expression (35) can be written in the form of two equalities that describe the centrosymmetric capillary vibration of the bubble (taking place at the varying volume) and the axisymmetric capillary vibration of its shape at constant volume:

$$r = R(t): \frac{dR}{dt} = \frac{\partial \psi_1^{(r)}}{\partial r}, \quad (40)$$

$$r = R(t): \frac{\partial \xi}{\partial t} = \frac{\partial \psi_1^{(s)}}{\partial r} - \frac{1}{r} \Delta_{\Omega} \psi_3. \quad (41)$$

The substitution of (34) into (13) and (14) yields

$$\frac{\partial}{\partial \vartheta} \left\{ 2 \frac{\partial}{\partial r} \left(\frac{\psi_1^{(s)}}{r} \right) + \frac{\partial^2 \psi_3}{\partial r^2} - \frac{1}{r^2} (2 + \Delta_{\Omega}) \psi_3 \right\} \quad (42)$$

$$+ \frac{r}{\sin \vartheta} \frac{\partial}{\partial \varphi} \left\{ \frac{\partial}{\partial r} \left(\frac{\psi_2}{r} \right) \right\} = 0,$$

$$\frac{1}{\sin \vartheta} \frac{\partial}{\partial \varphi} \left\{ 2 \frac{\partial}{\partial r} \left(\frac{\psi_1^{(s)}}{r} \right) + \frac{\partial^2 \psi_3}{\partial r^2} - \frac{1}{r^2} (2 + \Delta_{\Omega}) \psi_3 \right\} \quad (43)$$

$$- r \frac{\partial}{\partial \vartheta} \left\{ \frac{\partial}{\partial r} \left(\frac{\psi_2}{r} \right) \right\} = 0.$$

These equations must be solved on the spherical surface $r = R(t)$ of the bubble.

Let us apply the operator $(1/\sin \vartheta)(\partial/\partial \vartheta)$ to Eq. (42) and the operator $(1/\sin \vartheta)(\partial/\partial \varphi)$ to Eq. (43) and then combine the equations to obtain

$$r = R(t): 2 \frac{\partial}{\partial t} \left(\frac{\psi_1^{(s)}}{r} \right) + \frac{\partial^2 \psi_3}{\partial r^2} - \frac{1}{r^2} (2 + \Delta_{\Omega}) \psi_3 = 0. \quad (44)$$

Applying the operator $(1/\sin \vartheta)(\partial/\partial \varphi)$ to Eq. (42) and the operator $(1/\sin \vartheta)(\partial/\partial \vartheta)$ to Eq. (43) and then subtracting the equations, we obtain

$$r = R(t): \frac{\partial}{\partial r} \left(\frac{\psi_2}{r} \right) = 0. \quad (45)$$

The Laplacian and electrical pressures, like the pressure in the liquid, can be represented as the sum of two terms [17]. The first one is proportional to the zero-order spherical function, and the other is the series in spherical functions starting with the second-order function:

$$P_{\sigma} = \frac{2\sigma}{R(t)} + \frac{\sigma}{R^2(t)} \sum_{n=2}^{\infty} \sum_{m=-n}^{m=n} (n-1)(n+2) \quad (46)$$

$$\times Z_{nm}(t) Y_{nm} \equiv P_{\sigma}^{(r)} + P_{\sigma}^{(s)},$$

$$P_q = \frac{Q^2}{8\pi\epsilon R^4(t)} + \frac{Q^2}{4\pi\epsilon R^5(t)} \sum_{n=2}^{\infty} \sum_{m=-n}^{m=n} (n-1) \quad (47)$$

$$\times Z_{nm}(t) Y_{nm} \equiv P_q^{(r)} + P_q^{(s)}.$$

In view of (34), (46), and (47) and the orthogonality of the spherical functions, Eq. (15) is split into two equations:

$$r = R(t): P^{(r)} + P_{\sigma}^{(r)} - 2\rho v \frac{\partial^2 \psi_1^{(r)}}{\partial r^2} \quad (48)$$

$$- P_V - P_g - P_q^{(r)} = 0,$$

$$r = R(t): P^{(s)} + P_{\sigma}^{(s)}$$

$$- 2\rho v \frac{\partial}{\partial r} \left(\frac{\partial \psi_1^{(s)}}{\partial r} - \frac{1}{r} \Delta_{\Omega} \psi_3 \right) - P_q^{(s)} = 0. \quad (49)$$

Eventually, the problem being solved is subdivided into two problems: determination of the centrosymmetric radial vibration of the bubble and finding the spectrum of capillary waves on the surface of the constant-volume bubble.

(4) (i) Consider the former problem. Substituting the expression $\psi_1^{(r)} = C_{00}^1(t)/r$ into kinematic boundary condition (40) and determining the constant $C_{00}^1(t)$, we obtain

$$\psi_1^{(r)} = -\frac{R^2(t) dR(t)}{t dt}. \quad (50)$$

Substituting (50) into dynamic boundary condition (48) and taking into account (1), (31), (46), and (47), we arrive at the equation for the radial motion of the bubbles:

$$R(t) \frac{d^2 R(t)}{dt^2} + 2 \left(\frac{dR(t)}{dt} \right)^2 + \frac{4\nu}{R(t)} \frac{dR(t)}{dt} - \frac{1}{\rho} P(R(t)) = 0. \tag{51}$$

Equation (51) is an ordinary nonlinear second-order differential equation. It follows from the equations of linear hydrodynamics and coincides with the well-known Raleigh equation (which is derived from the nonlinearized Navier–Stokes equation) up to the coefficient before the first-order derivative squared (this coefficient does not influence stability analysis in the linear approximation). The study of the radial motions of a bubble in a liquid (see, e.g., [13]) involves the linearization of Eq. (51) in the vicinity of the equilibrium state (point of rest), which meets the condition $P(R) = 0$. In this case, the bubble radius is given by $R(t) = R + \delta R(t)$, where $\delta R(t)$ is of the first order of smallness. Then, Eq. (51) is recast as

$$R \frac{d^2 \delta R(t)}{dt^2} + \frac{4\nu}{R} \frac{d \delta R(t)}{dt} - \frac{1}{\rho} \frac{dP(R)}{dR} \delta R(t) = 0. \tag{52}$$

The characteristic equation for ordinary linear differential equation with constant coefficients (52) has the form

$$R\lambda^2 + \frac{4\nu}{R}\lambda - \frac{1}{\rho} \frac{dP(R)}{dR} = 0, \tag{53}$$

hence,

$$\lambda = -\frac{2\nu}{R^2} \pm \sqrt{\frac{4\nu^2}{R^4} + \frac{1}{\rho R} \frac{dP(R)}{dR}}. \tag{54}$$

It is seen from (54) that the radial motion of the bubble ($\sim e^{\lambda t}$) can be both stable and unstable depending on the magnitude and sign of the derivative $dP(R)/dR$ (the sign of this derivative depends on a specific equilibrium state of the bubble). In-depth stability analysis for the radial motion of the bubble has been described in [13], and we will not dwell on it.

(ii) Now let us find the spectrum of the capillary vibration of the bubble shape at constant volume. In our stability analysis, we will assume that the surface perturbation ξ and the potential function $\psi_1^{(s)}$ exponentially vary with time (e^{St}). Because of this, we put $Z_{nm}(t) = Z_{nm} e^{St}$ and $C_{nm}^1(t) = C_{nm}^1 e^{St}$ in (37) and (38). Bearing in mind that $\delta R(t)$ in the expression $R(t) = R + \delta R(t)$ is a time-varying quantity of first order of smallness, we refer boundary conditions (41), (44), and (49)

to the equilibrium state $r = R$. Then, substituting (37)–(39) into (41) yields

$$SZ_{nm} + \frac{(n+1)}{R^{n+2}} C_{nm}^1 - \frac{n(n+1)}{R} k_n \left(\sqrt{\frac{S}{\nu}} R \right) C_{nm}^3 = 0. \tag{55}$$

With regard for the recurrence relations for modified spherical Bessel functions of the 3rd order,

$$\frac{dk_n(z)}{dz} = \frac{n}{z} k_n(z) - k_{n+1}(z), \tag{56}$$

$$\frac{d^2 k_n(z)}{dz^2} = \left(1 + \frac{n(n-1)}{z^2} \right) k_n(z) + \frac{2}{z} k_{n+1}(z),$$

the substitution of (38) and (39) into (44) gives

$$-\frac{2(n+2)}{R^{n+3}} C_{nm}^1 + \left\{ \frac{S}{\nu} + \frac{2(n-1)(n+1)}{R^2} \right\} \tag{57}$$

$$\times k_n \left(\sqrt{\frac{S}{\nu}} R \right) + \frac{2}{R} \sqrt{\frac{S}{\nu}} k_{n+1} \left(\sqrt{\frac{S}{\nu}} R \right) \Big] C_{nm}^3 = 0.$$

Finally, the substitution of (38) and (39) into (49) in view of (32), (46), (47), and (56) leads us to

$$\frac{(n-1)}{R^2} \left[\sigma(n+2) - \frac{Q^2}{4\pi\epsilon R^3} \right] Z_{nm} - \frac{\rho}{R^{n+1}} \left[S + \frac{2\nu(n+1)(n+2)}{R^2} \right] C_{nm}^1 + \frac{2\rho\nu n(n+1)}{R} \left(\sqrt{\frac{S}{\nu}} k_{n+1} \left(\sqrt{\frac{S}{\nu}} R \right) \right. \tag{58}$$

$$\left. - \frac{n-1}{R} k_n \left(\sqrt{\frac{S}{\nu}} R \right) \right) C_{nm}^3 = 0.$$

Equations (55), (57), and (58) make up a set of uniform linear equations for the unknown coefficients Z_{nm} , C_{nm}^1 , and C_{nm}^3 . This set will have a nontrivial solution if and only if its determinant equals zero, which leads to the dispersion relation

$$S^3 + \frac{2\nu S(n+1)(2n+1)}{R^2} \left(S + \frac{2\nu(n+2)(n-1)}{R^2} \right) + \frac{(n+1)(n-1)(n+2)\alpha_n}{\rho R^3} \left(S - \frac{2\nu(2n+1)}{R^2} \right) + \frac{2\sqrt{S\nu}}{R} \left(S^2 - \frac{2\nu S(n+1)(n+2)(n-1)}{R^2} \right) + \frac{(n+1)(n-1)(n+2)\alpha_n}{\rho R^3} \Big) \frac{k_{n+1}(\sqrt{S\nu^{-1}}R)}{k_n(\sqrt{S\nu^{-1}}R)} = 0, \tag{59}$$

where

$$\alpha_n = \sigma - \frac{Q^2}{4\pi\epsilon(n+2)R^3}.$$

Equation (59) describes the spectrum of the capillary vibration of the bubble surface at constant volume. This equation must be analyzed numerically; however, analytical expressions can be derived in the asymptotic approximations of low- and high-viscous liquids.

(5) Consider the case of a low-viscous liquid. Modified spherical integral-order Bessel functions of the 3rd kind are given by [16]

$$k_n(z) = \frac{\pi e^{-z}}{2z} \sum_{k=0}^n \frac{(n+k)!}{k!(n-k)!} (2z)^{-k}. \quad (60)$$

If the viscosity of the liquid is so low that $\nu \ll R^2S$, the argument $z = \sqrt{S\nu}R$ of the spherical cylindrical function is much greater than unity ($z \gg 1$). Then, from (60), one readily obtains the asymptotic relation

$$\frac{k_{n+1}(z)}{k_n(z)} = 1 + \frac{n+1}{z} + o\left(\frac{1}{z}\right). \quad (61)$$

Substituting (61) into (59) and leaving the terms of first order of smallness with respect to viscosity, we come to the dispersion relation

$$S^3 + \left(\frac{2\sqrt{S\nu}}{R} + \frac{4\nu(n+1)^2}{R^2}\right)S^2 + \left(S + \frac{2\sqrt{S\nu}}{R} - \frac{2\nu n}{R^2}\right) \times \frac{(n+1)(n-1)(n+2)\alpha_n}{\rho R^3} = 0. \quad (62)$$

Directing the viscosity to zero in (62), we deduce the expression for the frequency of the capillary vibration of the bubble in perfect liquid:

$$S = \pm i \sqrt{\frac{(n+1)(n-1)(n+2)\alpha_n}{\rho R^3}} \equiv \pm i\omega_0. \quad (63)$$

It follows from (63) that the bubble is unstable against deformation at constant volume if $\alpha_n \leq 0$. Within our model, this condition should be met jointly with the condition $P(R) = 0$; thus, we arrive at the set of equations

$$\alpha_n = \sigma - \frac{Q^2}{4\pi\epsilon(n+2)R^3} \leq 0, \quad (64)$$

$$P(R) = P_V + P_{g0} \left(\frac{P_0}{R}\right)^{3\gamma} + \frac{Q^2}{8\pi\epsilon R^4} - \frac{2\sigma}{R} - P_\infty = 0.$$

Equation (64) implies that the bubble becomes unstable when the electrical pressure equals the Laplacian pressure and the pressure of the gas-vapor mixture equals the pressure in the liquid. Physically, such a conclusion is quite obvious: if the pressure of the mixture

in the bubble exceeds the pressure in the liquid, any thermal perturbation of the surface shape is eliminated because of the pressure drop between the bubble and the liquid; if the pressure drop has the opposite sign, any thermal perturbation of the surface shape will grow.

In view of the circumstance that the second mode in (64) may lose stability first, one easily obtains the condition of bubble instability against shape variation:

$$\frac{Q^2}{16\pi\epsilon\sigma R_0^3} \geq \sqrt{\frac{P_{g0}}{P_\infty - P_V}}.$$

In a viscous liquid, surface shape variations will decay because of viscous energy dissipation. Thus, the frequency of the surface vibration of the bubble in a viscous liquid can be represented in the form $S = \pm i\omega_0 + \delta S$, where δS is of first order of smallness with respect to viscosity. Substituting this expression into (62) yield the correction to the frequency

$$\delta S = -\frac{\nu(n+2)(2n+1)}{R^2}. \quad (65)$$

This correction defines the damping constant for the surface vibration of the bubble in a low-viscous liquid, which was calculated long ago by Lamb [18].

(6) For the case of a high-viscous liquid $\nu \gg R^2S$, the argument of the cylindrical function tends to zero. Then, only the higher-order terms of the singular part in series (60) must be left:

$$k_n(z) = \frac{\pi e^{-z}}{2z} \left(\frac{(2n)!}{n!(2z)^n} + \frac{(2n-1)!}{(n-1)!(2z)^{n-1}} + \frac{(2n-2)!}{(n-2)!2!(2z)^{n-2}} + o\left(\frac{1}{z^{n-2}}\right) \right), \quad (66)$$

$$k_{n+1}(z) = \frac{\pi e^{-z}}{2z} \left(\frac{(2n+2)!}{(n+1)!(2z)^{n+1}} + \frac{(2n+1)!}{n!(2z)^n} + \frac{(2n)!}{(n-1)!2!(2z)^{n-1}} + o\left(\frac{1}{z^{n-1}}\right) \right). \quad (67)$$

From (66) and (67),

$$\frac{k_{n+1}(z)}{k_n(z)} = \frac{2n+1}{z} + \frac{z}{2n-1} + o(z). \quad (68)$$

Substituting (68) into (59) yields the dispersion relation for the capillary vibration of the bubble at constant volume for a high-viscous liquid:

$$S^2 + \frac{2\nu S(n+2)(2n^2+1)}{R^2(2n+1)} + \omega_0^2 = 0. \quad (69)$$

From (69) at $\nu R^{-2} \gg \omega_0$, one readily obtains

$$S_1 = -\frac{\omega_0^2 R^2}{2\nu} \frac{2n+1}{(n+2)(2n^2+1)} + O\left(\frac{1}{\nu^3}\right),$$

$$S_2 = -\frac{2\nu(n+2)(2n^2+1)}{R^2} + O\left(\frac{1}{\nu}\right). \tag{70}$$

It follows from (70) that the bubble becomes unstable (i.e., $S \geq 0$) when $\alpha_n \leq 0$, or $\omega_0^2 \leq 0$ (as for the case of low viscosity), but the instability increment here is small, since $\omega_0 \leq \nu R^{-2}$. It is worth noting that the instability increment for a charged bubble is inversely proportional to the viscosity, as for a charged viscous drop [19].

At $\omega_0^2 > 0$, both roots of (70) are negative and define the damping constants of the capillary vibration of the bubble surface. The first root is much larger and is of greater interest from the physical point of view (as for the case of waves on the planar surface of a viscous liquid [20]), since vibrations related to the second root rapidly decay.

(7) The conclusions following from the asymptotic considerations are supported by numerical calculations. They were performed by jointly using Eqs. (59) and

$P(R) = 0$ written in dimensionless variables where $R_0 = \rho = \sigma = 1$ and all the physical parameters are expressed in terms of their characteristic scales:

$$P^* = \frac{\sigma}{R_0}, \quad R^* = R_0, \quad Q^* = \sqrt{R_0^3 \sigma},$$

$$\nu^* = \sqrt{\frac{R_0 \sigma}{\rho}}, \quad S^* = \sqrt{\frac{\sigma}{\rho R_0^3}}.$$

In this basis, the problem is characterized by three dimensionless parameters: the Rayleigh parameter $W = Q^2/(16\pi\epsilon)$, the parameter that describes the pressure in the bubble $\beta = P_{g0}/2$, and the parameter that defines the pressure in the liquid $\beta_* = (P_\infty - P_\nu)/2$. Depending on the values of these parameters, the equation $P(R) = 0$ has one or two solution or does not solutions at all, as follows from (1) and Descartes' rule of signs [13]. If $\beta_* > 0$ and the other two parameters are arbitrary (region A), we have one solution. In region B ($\beta_{*cr} < \beta_* < 0$ and the Rayleigh and gas parameters are smaller than their critical values), two equilibrium states arise. Finally, in region C (where $\beta_* < \beta_{*cr}$, $W > W_{cr}$, and $\beta > \beta_{cr}$), no solution is found.

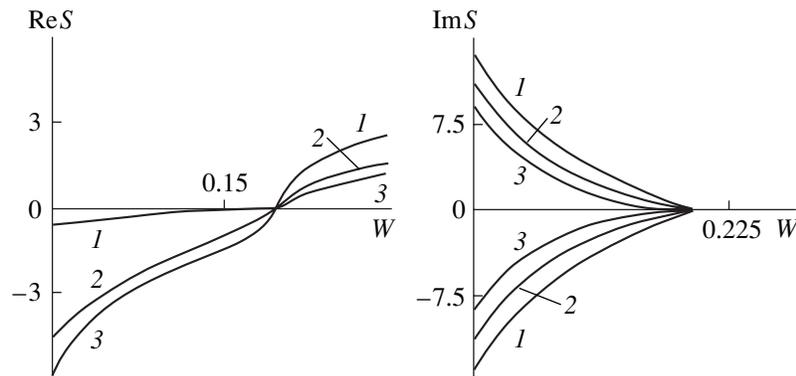


Fig. 1. Complex frequency vs. Rayleigh parameter for $n = 2$, $\gamma = 1$, $\beta = 0.2$, and $\beta_* = 1$. $\nu = (1) 0.005$, $(2) 0.06$, and $(3) 0.1$.

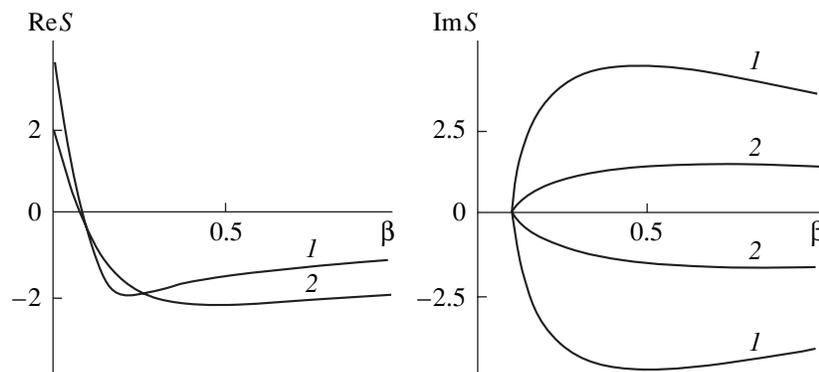


Fig. 2. Complex frequency vs. parameter β that characterizes the gas pressure in the bubble for $n = 2$, $\gamma = 1$, $W = 0.1$, and $\beta_* = 1$. $\nu = (1) 0.05$ and $(2) 0.3$.

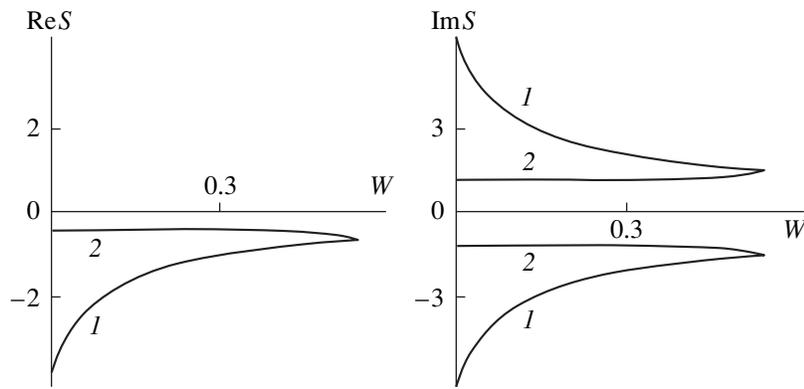


Fig. 3. Same as in Fig. 1 for $\beta_* = -0.5$ and $\nu = 0.1$. Equilibrium states with a (1) smaller and a (2) larger radius.

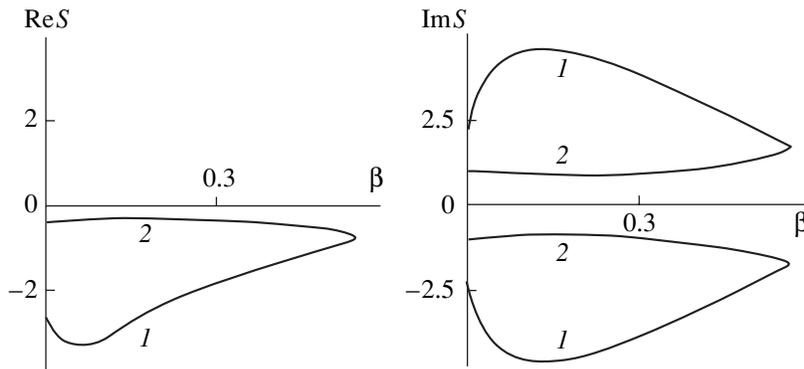


Fig. 4. Same as in Fig. 2 for $W = 0.05$, $\beta_* = -0.5$, and $\nu = 0.1$. Equilibrium states with a (1) smaller and a (2) larger radius.

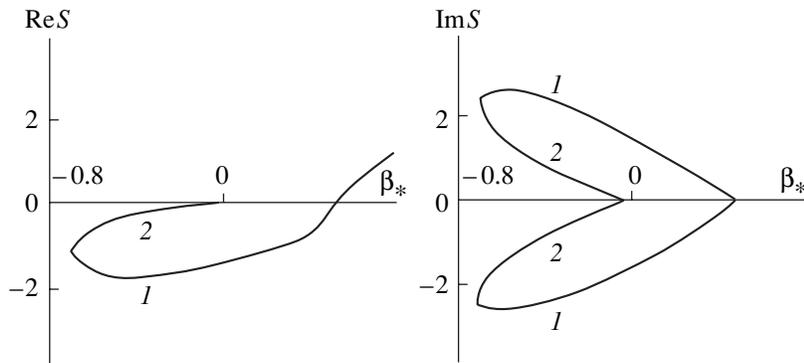


Fig. 5. Complex frequency vs. parameter β_* that characterizes the pressure in the liquid for $n = 2$, $\gamma = 1$, $W = 0.2$, $\beta = 1$, and $\nu = 0.1$. Equilibrium states with a (1) smaller and a (2) larger radius.

In region A, the frequencies of the surface vibration and the damping constant decrease with increasing W .

At a certain critical value of $W = \sqrt[3]{\beta/\beta_*}$, the damping constant vanishes and then becomes negative, turning to the instability increment (Fig. 1). The instability increment decreases as the viscosity grows. In this range, the stability of the bubble considerably depends on the initial pressure of the gas. When the gas pressure decreases, the real part of the complex frequency also (as in the case of increasing Rayleigh parameter) passes

through zero, causing instability, as follows from the curves $S = S(R(\beta))$ (Fig. 2).

In region B, the bubble is absolutely stable against surface perturbations. For example, when the Rayleigh parameter grows, the capillary vibration frequency and the damping constant decrease for the equilibrium bubble with a lower radius and, conversely, increase for the bubble with a larger radius (Fig. 3). The reason is that when the Rayleigh parameter grows, two solutions of the equation $P(R) = 0$ converge at a boundary point between regions B and C and then disappear. An

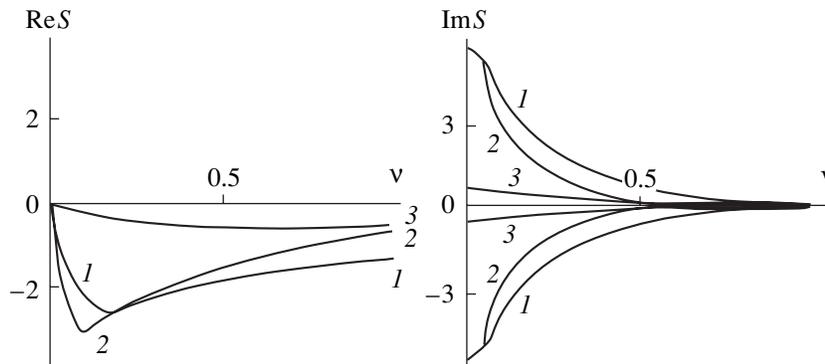


Fig. 6. Complex frequency vs. viscosity for $n = 2$, $\gamma = 1$, and $W = 0.05$. $\beta_* = (1, 2) 0.6$ and $(3) -0.6$. $\beta = (1) 0.5$ and $(2, 3) 0.1$. 3, equilibrium state with a larger radius.

increase in the gas pressure also does not affect the shape stability of the bubble, as is seen in Fig. 4.

In region *B*, an increase in both the pressure in the liquid and the Rayleigh parameter does not affect the shape stability. As follows from Fig. 5, an increase in the pressure in the liquid decreases the frequencies and the damping constant for the equilibrium bubble with both a small and a large radius. However, for $\beta_* = 0$, the larger-radius equilibrium state disappears, while the smaller-radius one passes to region *A* and soon becomes unstable.

The complex frequency vs. viscosity behavior is similar for any of the regions and equilibrium states. Specifically (Fig. 6), the capillary vibration frequencies drop with increasing viscosity. The damping constant first linearly increases in the low-viscosity range and then starts to decrease, which is in good qualitative and quantitative agreement with the above asymptotic calculations.

CONCLUSION

The bubble becomes unstable against shape variation only if the pressure in the liquid exceeds the pressure of the saturated vapor–gas mixture in the bubble. Conversely, the bubble is unstable against volume variation only if the pressure in the liquid is lower than in the bubble. The charge on the bubble significantly diminishes the critical pressure in the liquid at which shape instability occurs. The more viscous the liquid, the smaller the instability increment. The effect of viscosity on the damping constant is different in low- and high-viscous liquids. In the former case, the damping constants for the surface and radial components linearly increase. At a high viscosity, the damping constant for the surface component starts to decrease, while that for the radial component continues growing.

REFERENCES

1. A. N. Zharov and S. O. Shiryayeva, *Elektrokhim. Obrab. Met.*, No. 6, 9 (1999).
2. V. V. Skorykh, *Zh. Tekh. Fiz.* **56** (8), 1569 (1986) [*Sov. Phys. Tech. Phys.* **31**, 930 (1986)].
3. V. F. Klimkin, *Pis'ma Zh. Tekh. Fiz.* **16** (4), 54 (1990) [*Sov. Tech. Phys. Lett.* **16**, 146 (1990)].
4. A. L. Droyaronov, *Elektrokhim. Obrab. Met.*, No. 4, 39 (1993).
5. M. M. Khasanov, *Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza*, No. 2, 66 (1994).
6. N. F. Bunkin and A. F. Lobevev, *Pis'ma Zh. Tekh. Fiz.* **19** (21), 38 (1993) [*Tech. Phys. Lett.* **19**, 682 (1993)].
7. V. L. Zelenko and V. P. Myasnikov, *Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza*, No. 3, 59 (1992).
8. M. K. Bologa, S. M. Klimov, and S. I. Chuchkalov, *Elektrokhim. Obrab. Met.*, No. 2, 52 (1992).
9. A. N. Malinin, V. E. Sabinin, and A. N. Sidorov, *Pis'ma Zh. Tekh. Fiz.* **20** (1), 57 (1994) [*Tech. Phys. Lett.* **20**, 28 (1994)].
10. A. G. Lipson, V. A. Klyuev, B. V. Deryagin, *et al.*, *Pis'ma Zh. Tekh. Fiz.* **16** (19), 89 (1990) [*Sov. Tech. Phys. Lett.* **16**, 763 (1990)].
11. A. G. Lipson, B. F. Lyakhov, B. V. Deryagin, *et al.*, *Pis'ma Zh. Tekh. Fiz.* **17** (21), 33 (1991) [*Sov. Tech. Phys. Lett.* **17**, 763 (1991)].
12. A. I. Zhakin, *Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza*, No. 2, 14 (1988).
13. A. I. Grigor'ev and A. N. Zharov, *Zh. Tekh. Fiz.* **70** (4), 8 (2000) [*Tech. Phys.* **45**, 389 (2000)].
14. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
15. S. O. Shiryayeva, A. É. Lazaryants, A. I. Grigor'ev, *et al.*, Preprint No. 27, IMRAN (Inst. of Microelectronics, Russian Academy of Sciences, Yaroslavl, 1994).
16. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
17. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
18. H. Lamb, *Hydrodynamics* (Cambridge Univ. Press, Cambridge, 1932; Ogiz-Gostekhizdat, Moscow, 1947).
19. S. O. Shiryayeva, *Zh. Tekh. Fiz.* **70** (9), 30 (2000) [*Tech. Phys.* **45**, 1128 (2000)].
20. V. G. Levich, *Physicochemical Hydrodynamics* (Akad. Nauk SSSR, Moscow, 1952).

Translated by V. Isaakyan

The Exact Solution of the Equation for Diffusion of a Three-Component Mixture through a Capillary

O. E. Aleksandrov

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

e-mail: aleks@dpt.ustu.ru

Received November 20, 2000; in final form, February 5, 2001

Abstract—The exact solution of the equation for steady-state diffusion of a three-component mixture is presented for a closed system of two capillary-connected bulbs. The solution, obtained in the whole range of concentrations, makes it possible to compute the concentration profiles of the components and the distribution of the mixture density over the capillary. It is shown that an extremum of the mixture density may arise inside of the capillary when the components greatly differ in diffusion coefficients. Specifically, when the mixture in the lower bulb is denser than in the upper one and the stratification of the system is stable, an inverse density gradient region may occur in the capillary, this region being unstable against gravitational convection. © 2001 MAIK “Nauka/Interperiodica”.

Free convection in an inhomogeneous three-component gas mixture in the gravity field when the mixture at the bottom is denser than at the top was discovered in 1966. As has been demonstrated in pioneering works [1–3] and subsequent experimental studies [4, 5], the convection shows a number of intriguing features.

(1) In passing from stable to convective diffusion, the component fluxes grow several tens or several thousands times.

(2) The convection rate reaches a maximum at a certain pressure (≈ 3 MPa) and then is reduced nearly to zero as the pressure further increases.

(3) The anomalous separation of the mixture when a heavy component penetrates through the capillary faster than a light one has been observed.

(4) Three convection conditions—monotone, oscillatory, and random—have been observed.

(5) The recurring kinetic phase transition has been discovered: during a single event of three-component mixing, the conditions in the two-bulb system alternate in the order stable diffusion \rightarrow convection \rightarrow stable diffusion \rightarrow convection \rightarrow stable diffusion as the difference in component concentrations decreases.

To find reasons for the anomalous convection, a system of equations for three-component steady-state diffusion through a capillary was solved in the case when one of the components has a low concentration [5]. However, for finite concentrations of the components, the solution turns out to be physically unreasonable: the concentration profile of one of the components shows an extremum in the capillary. In this paper, the exact solution for a system of equations for steady-state three-component diffusion throughout the concentration range is presented for a closed system of two capillary-connected bulbs.

We consider steady-state diffusion of a three-component mixture of ideal gases through a long capillary: $L \gg R$, where L and R are the capillary length and radius, respectively (Fig. 1). For definiteness, the components will be numbered by subscripts $i = 1, 2$, and 3 so that $m_1 < m_2 < m_3$, where m_i is the molecular weight of the i th component.

Let the concentrations of the components and the pressures in the bulbs be kept constant. Then, three-component steady-state diffusion through the capillary is described by the system of equations

$$T = \text{const}; \quad \sum_i c_i = 1; \quad n \sum_i c_i \mathbf{u}_i = 0;$$

$$\text{div}(n c_i \mathbf{u}_i) = 0; \quad i = 1, 2; \quad (1)$$

$$\sum_{j=1}^3 \frac{c_i c_j}{D_{ij}} (\mathbf{u}_i - \mathbf{u}_j) = -\text{grad}(c_i); \quad i = 1, 2,$$

where T is the gas temperature; n is the gas density, which varies along the capillary due to the barotropic

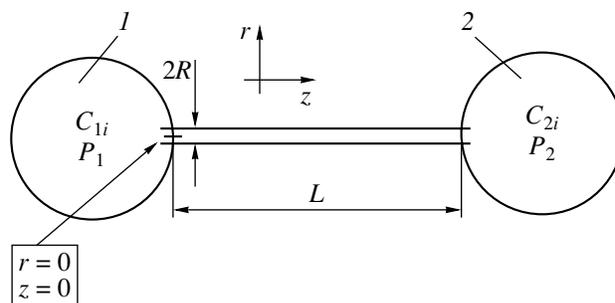


Fig. 1. Diffusion through the capillary: (1, 2) bulbs.

effect and gravitation; c_i is the molar concentration of the i th component; \mathbf{u}_i is the vector of the averaged velocity of the i th component; and D_{ij} is the coefficient of interdiffusion for the i th and j th components. The last equation in (1) is known as the Stefan–Maxwell equation. System (1) is complete, because it involves twelve independent equations and twelve unknowns (c_i, \mathbf{u}_i).

Since the capillary is long and the gas as a whole does not move in it, the components of the mixture be uniformly distributed in the transverse direction, so that one should solve the one-dimensional diffusion problem. The averaging of system (1) over the capillary cross section yields the following system of equations (the temperature equation is omitted):

$$\sum_i c_i = 1; \quad n \sum_i c_i u_i = 0;$$

$$nc_i u_i = \text{const}_i = j_i = \frac{J_i}{S}; \quad i = 1, 2; \quad (2)$$

$$\sum_{j=1}^3 \frac{c_i c_j}{D_{ij}} (u_i - u_j) = -\frac{dc_i}{dz}; \quad i = 1, 2,$$

where J_i is the total flux of the i th component through the capillary, S is the channel cross-sectional area, and $j_i = nc_i u_i$ is the i th component flux density.

The boundary conditions for system (2) are given by

$$c_1(0) = C_{11}, \quad c_2(0) = C_{21}, \quad c_1(L) = C_{12},$$

$$c_2(L) = C_{22}, \quad (3)$$

where C_{ik} is the component concentrations ($i = 1, 2$ is the component number and $k = 1, 2$ is that of the bulbs).

In view of the notation introduced into (2), one can transform the Stefan–Maxwell equation to

$$\sum_{j=1}^3 \frac{1}{nD_{ij}} (c_j j_i - c_i j_j) = -\frac{dc_i}{dz}; \quad i = 1, 2. \quad (4)$$

Taking into account that $j_i = \text{const}$ and $j_3 = -(j_1 + j_2)$, we get the system of linear inhomogeneous equations with constant coefficients for two independent concentrations:

$$\frac{dc_1}{dz} - c_1 j_2 \left(\frac{1}{D_{12}} - \frac{1}{D_{13}} \right) + c_2 j_1 \left(\frac{1}{D_{12}} - \frac{1}{D_{13}} \right) = \frac{-j_1}{D_{13}},$$

$$c_1 j_2 \left(\frac{1}{D_{12}} - \frac{1}{D_{23}} \right) + \frac{dc_2}{dz} - c_2 j_1 \left(\frac{1}{D_{12}} - \frac{1}{D_{23}} \right) = \frac{-j_2}{D_{23}}, \quad (5)$$

where j_i are arbitrary constants of integration that should be determined from the boundary conditions.

The solution of a system of linear inhomogeneous equations is the sum of the particular solution of the

inhomogeneous system, c_{Fi} , and the general solution of the homogeneous system, c_{0i} :

$$c_i = c_{Fi} + c_{0i}.$$

The general solution of the homogeneous system can be written as

$$c_{0i} = \sum_{l=1}^2 A_{li} e^{S_l z}. \quad (6)$$

Using the notation

$$\lambda = -\left(\frac{j_1}{D_{23}} + \frac{j_2}{D_{13}} + \frac{j_3}{D_{12}} \right),$$

$$B = \frac{j_1}{D_{12}D_{13}} + \frac{j_2}{D_{12}D_{23}} + \frac{j_3}{D_{13}D_{23}},$$

we will obtain

$$S_1 = 0, \quad S_2 = \lambda. \quad (7)$$

Substituting (6) and (7) into the initial system of equations yields the expressions for the coefficients A_{li} ; eventually, the general solution of the homogeneous system can be written as

$$c_{01}(z) = j_1 K_1 + \left(\frac{1}{D_{12}} - \frac{1}{D_{13}} \right) K_2 e^{\lambda z},$$

$$c_{02}(z) = j_2 K_1 + \left(\frac{1}{D_{12}} - \frac{1}{D_{23}} \right) K_2 e^{\lambda z}, \quad (8)$$

where K_i are constants of integration.

The partial solution of the inhomogeneous system can be found as a linear function of z :

$$c_{F1}(z) = -j_1 \frac{B}{\lambda} z,$$

$$c_{F2}(z) = -j_2 \frac{B}{\lambda} z - \frac{D_{12}(1-B)}{\lambda(D_{13}-D_{12})}. \quad (9)$$

The general solution of inhomogeneous system (5) is the sum of the partial solution of the inhomogeneous system, c_{Fi} , and the general solution of the homogeneous system, c_{0i} . The constants of integration K_i and j_i should be determined from boundary conditions (3). The fluxes j_i cannot be expressed in terms of the concentrations at the ends of the channels, and the values of K_i are obtainable from the first and second boundary conditions (3) in the form of

$$K_1 = \frac{1}{\lambda} \left(C_{11} \left(\frac{1}{D_{12}} - \frac{1}{D_{23}} \right) + C_{12} \left(\frac{1}{D_{12}} - \frac{1}{D_{13}} \right) \right),$$

$$K_2 = \frac{1}{\lambda} (C_{11} j_2 + C_{12} j_1).$$

The flux densities j_i can be found by numerically solving the set of the penultimate and last equations

from boundary conditions (3). The resulting solution can be used to calculate the mixture density profile in the capillary:

$$\rho(z) = \sum_i n m_i c_i(z), \quad (11)$$

here, we neglect the diffusion barotropic effect because of its smallness in a continuous medium. Substituting concentrations (8) and (9) into (11) and differentiating the resulting expression, we obtain the position z_{ex} of the density extremum in the capillary:

$$z_{\text{ex}} \lambda = \ln \left(\frac{\rho V B}{K_2 \lambda^2 \left[\left(\frac{1}{D_{12}} - \frac{1}{D_{13}} \right) \Delta m_{13} - \left(\frac{1}{D_{12}} - \frac{1}{D_{23}} \right) \Delta m_{23} \right]} \right), \quad (12)$$

where $\rho V = \sum_i m_i j_i$ is the density of the mass flux from the bulb containing the denser mixture to that with the mixture of lower density.

The extremum is due to the exponential term in (8), which becomes a constant when

$$\frac{j_1}{D_{23}} + \frac{j_2}{D_{13}} + \frac{j_3}{D_{12}} = 0,$$

in this case, the density extremum is absent.

To clarify the type of the extremum, one must calculate the second-order derivative of the density with respect to the z coordinate:

$$\left. \frac{d^2}{dz^2} \left(\frac{\rho}{n} \right) \right|_{z=z_{\text{ex}}} = \rho V \left(\frac{j_1}{D_{12} D_{13}} + \frac{j_2}{D_{12} D_{23}} + \frac{j_3}{D_{13} D_{23}} \right). \quad (13)$$

The value of the second-order derivative at the extreme point is obviously nonzero; hence, this is an extremum, rather than an inflection. However, we cannot predict the sign of the second-order derivative; in other words, both the maximum and the minimum of the mixture density may occur in the capillary.

Notice that only the density of the mixture may show an extremum: the concentration of the components is a monotonic function of z , decreasing from the bulb with a higher concentration to that with a lower one. Thus, the concentration profiles in [5] are incorrect.

For diffusion in a vertical capillary (in the gravitational field), the density extremum in the capillary may result in the convection instability of the diffusion process even when the lower density mixture is at the top. The experimental investigations into the convective instability of diffusion of various three-component gas mixtures through a vertical capillary have been

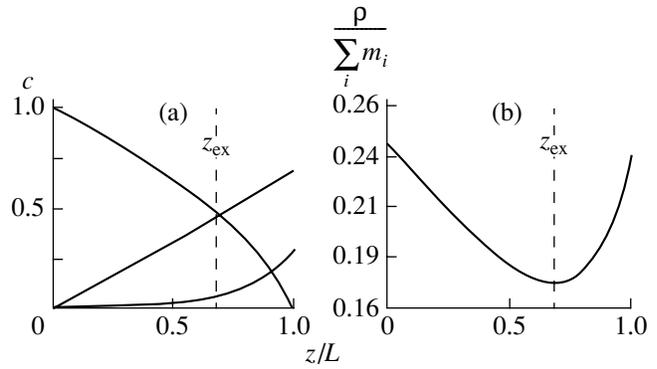


Fig. 2. Ar–(0.7 He + 0.3 R12) mixture (R12 is Freon used in rf plasma etching). (a) Concentration profile along the capillary and (b) density profile of the three-component mixture at $n = 1$. z_{ex} was calculated with (12). $j_{\text{He}} = -4.238 \times 10^{-5}$ m/s and $j_{\text{Ar}} = 4.247 \times 10^{-5}$ m/s.

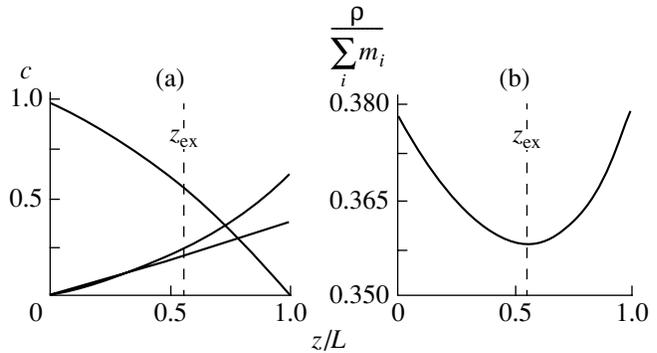


Fig. 3. Same as in Fig. 2 for N_2 –(0.38 H_2 + 0.62 CO_2) mixture, $j_{\text{H}_2} = -2.512 \times 10^{-5}$ m/s, and $j_{\text{N}_2} = 2.896 \times 10^{-5}$ m/s.

reported in [5]. The diffusion instability showed up as a significant rise in the flux of the components through the channel as compared with molecular diffusion. The compositions of the mixtures and the concentrations of the components for which the instability appeared were listed in [5].

The profiles of the density in the capillary were calculated with (8) and (9) for several mixtures used in the experiments [5] (see Figs. 2–5). Convection has been experimentally observed in these mixtures.

The capillary length L has no effect on the shape of the concentration and density profiles; therefore, all the computations were performed for the unit length. For the same reason, we can obtain the concentration and density profiles for any segment of the z axis (Figs. 2–5) if the proper concentrations are kept in the bulbs.

The stability against gravitational convection is always disturbed [5] when the density extremum in the capillary occurs (Figs. 2–5). Thus, in the process of diffusion in a three-component mixture, gravitational convection is due to the minimum of the mixture density inside the capillary. Here, two points are noteworthy.

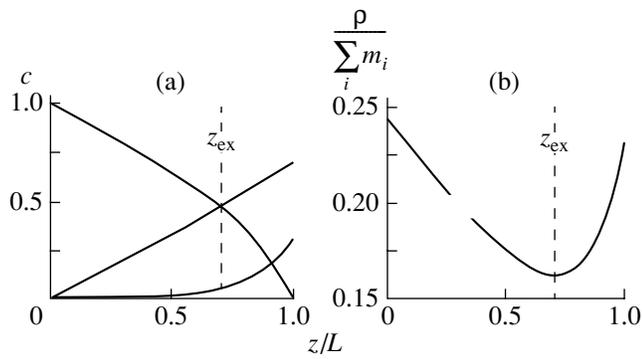


Fig. 4. Same as in Fig. 2 for Ar-(0.7 He + 0.3 Freon 12) mixture. $j_{\text{H}_2} = -4.829 \times 10^{-5}$ m/s and $j_{\text{N}_2} = 4.833 \times 10^{-5}$ m/s.

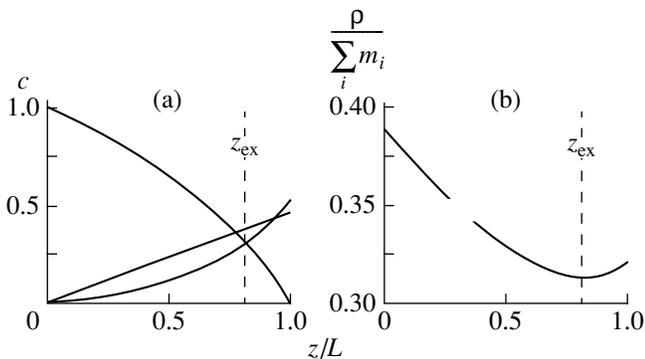


Fig. 5. Same as in Fig. 2 for Ar-(0.7N₂ + 0.3CO₂) mixture. $j_{\text{He}} = -2.944 \times 10^{-5}$ m/s and $j_{\text{N}_2} = 3.041 \times 10^{-5}$ m/s.

First, the density extremum inside the capillary is the original cause of the convection; however, once the convection has arisen, the above formulas become invalid. The density profile with an extremum will reappear if the convection stops. Second, the relative change in the density ($\rho_{\text{min}}/\rho_{\text{max}}$) for the gas mixtures shown in Figs. 2–5 ranges between 0.72 and 0.94. For a 1-m-long vertical capillary, the relative change in the density due

to gravitation is specified by the Boltzmann factor

$$\frac{\rho_{\text{min}}}{\rho_{\text{max}}} = \exp\left(-\frac{mg \, 1 \, \text{m}}{kT}\right),$$

which is equal to 0.9995 for the heaviest gas (Freon). Thus, we can assert that the change in the density due to three-component diffusion predominates in the case of a vertical capillary as well. The shorter the capillary, the less the relative change in the density due to gravitation; at the same time, the change due to three-component diffusion remains constant. Moreover, since the explicit expression for the density (pressure) profile along the capillary has not been used in the solution, it can easily be extended for a vertical capillary by multiplying by the barometric Boltzmann factor. In the general case, the exponent in the Boltzmann factor depends on the vertical coordinate; however, since the relative change in the density is small, one can roughly use the average molecular weight of the mixture in the capillary.

The occurrence of the density minimum is due to the presence of the third component in the mixture. For this minimum to arise, the diffusion coefficients of the components must significantly differ.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 98-01-00879).

REFERENCES

1. L. Miller and E. A. Mason, *Phys. Fluids* **8**, 711 (1966).
2. L. Miller, T. H. Spurling, and E. A. Mason, *Phys. Fluids* **10**, 1809 (1967).
3. B. A. Ivakin, P. E. Suetin, and G. S. Kharin, *Tr. Ural. Politekh. Inst. im. S.M. Kirova*, No. 172, 158 (1969).
4. Yu. I. Zhavrin, *Zh. Tekh. Fiz.* **54**, 943 (1984) [*Sov. Phys. Tech. Phys.* **29**, 561 (1984)].
5. V. N. Kosov and V. D. Seleznev, *Anomalous Diffusion Instability of Multicomponent Isothermal Gaseous Mixtures* (Print, Almaty, 1998).

Translated by B. Maluykov

**GAS DISCHARGES,
PLASMA**

Acceleration of a Multicomponent Plasma in the Cathode Region of a Vacuum Arc

I. A. Krinberg

Irkutsk State University, bul'v. Gagarina 20, Irkutsk, 664003 Russia

e-mail: krinberg@physdep.isu.runnet.ru

Received January 16, 2001

Abstract—A general solution to the problem of the steady-state spherical expansion of a current-carrying multicomponent plasma into a vacuum is derived. It is shown that, in vacuum arc discharges, the main force accelerating the cathode material, which becomes a plasma at distances of 1 to 300 μm from the cathode surface, is the electron pressure gradient force maintained by Joule heating. It is established that ions of different charges move with the same hydrodynamic velocity, which is uniquely determined by the mass and mean charge of the ions and the maximum electron temperature in the cathode region. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

A distinctive feature of electric discharges in a vacuum is rapid phase transitions of the cathode material from a solid state to liquid, gaseous, and plasma states (the plasma state is divided into the three substates: a dense nonideal plasma, a moderately rarefied plasma, and a collisionless plasma). In such discharges, the originally immobile cathode material is progressively accelerated to velocities of about 10^6 cm/s and, over the acceleration distance, the particle density decreases from $N \cong 10^{23}$ to $\leq 10^{10}$ cm^{-3} [1].

Currently, in the literature, there is considerable discussion of the mechanism for ion acceleration. A unified viewpoint regarding the region where the acceleration rate is the highest is lacking: the three candidates are (i) the region of a nonideal plasma with a very high density or even of a gas with a solid state (metallic) density [2, 3], (ii) the region of the hydrodynamic flow of a collisional ideal plasma [4–7], and (iii) the region of the collisionless motion of plasma particles [8–10]. Of course, the mechanism for ion acceleration should differ between the three regions of different particle densities. In the first region, the acceleration is thought to be primarily due to the energy release at the hydrodynamic and electromagnetic discontinuities that arise from phase transitions of the cathode material during the discharge [2, 3]. In the second region, a quasineutral plasma jet is thought to be accelerated by the pressure gradient force maintained by the Joule heating when the current flows through the plasma [4–7]. In the third region, the ions of different charges are thought to be accelerated independently [8, 10] (in the stage of collisionless plasma motion) by an anomalous electric field that may be generated in a certain part of the interelectrode gap as a result of the possible formation of a potential hump between the anode and the cathode and

is directed toward the anode [8]. Because of this, the velocity of the ions of the same mass m but of different charge numbers Z_k is approximately equal to $V_k \cong (2Z_k e \Delta\Phi / m)^{1/2}$, where $\Delta\Phi$ is the potential difference across the acceleration region.

Over a distance of several microns from the cathode, the plasma parameters are impossible to measure; consequently, the relative importance of different acceleration mechanisms can only be evaluated from the results of measurements at larger distances. It is well known [1, 11] that, in low-current ($I_{\text{arc}} < 300$ A) vacuum arcs, the ion composition and velocity are independent of the length l and geometry of the interelectrode gap, provided that the gap is long enough ($l \geq 0.3$ – 1 cm). This fact gives an upper estimate of the distances over which the cathode material is ionized and accelerated. That is why the results of measurements of the ion velocities at distances $r \geq 1$ cm from the cathode surface can serve as a check on the validity of theoretical ideas about ion acceleration in the cathode region. It should be noted that, for high currents ($I_{\text{arc}} > 300$ A) and short current rise times, the ions can be additionally accelerated outside the cathode region [12, 13] because of the compression of the plasma by its own magnetic field.

The experimental data available on the ion velocities are contradictory: first, the absolute values of the ion velocities measured in experiments with the same cathode material are different and, second, in some experiments [9, 14], the ion velocities are observed to increase with ion charge, while this is not so in other experiments [15–17]. The model calculations for a Cu cathode [6] also do not exhaustively clarify the situation. According to these calculations, the velocity of Cu^{2+} ions is higher than that of Cu^+ ions by about 10–15%. Conversely, in [9, 14], the velocity of Cu^{2+} ions was observed to be higher by 25–30%, and the velocity

of Cu^{3+} ions, by 45–50%. On the other hand, the experiments of [15] revealed no dependence of the velocity of these ions on their charge.

This paper is devoted to constructing a theory that describes the acceleration of a multicomponent gas-discharge plasma and makes it possible to calculate the velocities of ions of different charges for any cathode material. The question of what might be the dominant acceleration mechanism is answered by comparing theoretical predictions with the experimentally observed interrelations among the parameters of the plasma jet, without making any reference to the parameters that are derived indirectly from experimental measurements.

EQUATIONS OF ION MOTION

It is well known [1] that, in the cathode region of a vacuum arc, the plasma exists in the form of a set of individual microjets emitted from cathode microspots (i.e., emission centers or current cells) with a diameter of about $a \approx 1\text{--}3 \mu\text{m}$ and a current of about $I \approx 1\text{--}10 \text{ A}$. The acceleration of a microjet can be studied in a one-dimensional approximation. As was shown in [18], the results obtained in this way agree well with the results computed for the jet axis from a two-dimensional model. We consider the spherical expansion of a plasma from a spherical surface of radius $r_0 \equiv a$ into a cone having a solid angle Ω .

Let a multicomponent plasma be composed of electrons and several ion species with the same mass m but with different charge numbers Z_k . The equation of motion of the ions of species k has the form

$$mN_k \frac{d_k V_k}{dt} = -\frac{d(N_k T_+)}{dr} - Z_k e N_k \frac{d\Phi}{dr} + \sum_j R_{kj} + R_{ke}, \quad (1)$$

where N_k and V_k are the density and velocity of the ions of species k , the ion temperature T_+ is assumed to be the same for all ion species, Φ is the electric potential, r is the distance from the center of the sphere, and $d_k/dt = V_k d/dr$ in a steady state.

The ion–electron and ion–ion frictional forces are described by the relationships

$$R_{ke} = \alpha_0 Z_k^2 f_k m_e N_e (V_e - V_k) / \tau_e, \quad (2)$$

$$R_{kj} = Z_k^2 Z_j^2 f_k f_j m N_+ (V_j - V_k) / \tau_+, \quad (3)$$

where the quantities

$$\tau_e = \frac{3\sqrt{m_e} T_e^{3/2}}{4\sqrt{2\pi} e^4 N_+ \ln \Lambda}, \quad \tau_+ = \frac{3\sqrt{m} T_+^{3/2}}{4\sqrt{\pi} e^4 N_+ \ln \Lambda} \quad (4)$$

are the characteristic electron–ion and ion–ion collision times in the case of singly charged ions [19]. The rest of the notation in formulas (2)–(4) is as follows: m_e , N_e , T_e , and V_e are the mass, density, temperature, and veloc-

ity of the electrons, respectively; $N_+ = \sum_k N_k$ is the total ion density; $f_k = N_k/N_+$ is the relative fraction of the ions of species k ($\sum_k f_k = 1$); $\ln \Lambda$ is the Coulomb logarithm; and $\alpha_0 = 0.51, 0.44$, and 0.40 for $Z_k = 1, 2$, and 3 [19] (in what follows, we neglect the dependence $\alpha_0(Z)$ and set $\alpha_0 \approx 0.45$; we also assume that $\ln \Lambda = 4$). Because of the plasma quasineutrality, we have $N_e = \sum_k Z_k N_k = \langle Z \rangle N_+$, where $\langle Z \rangle = \sum_j Z_j f_j$ is the mean ion charge number.

When collisions between the particles are sufficiently frequent, the plasma can be described in the hydrodynamic model. In this case, it is convenient to use the equation of ion motion in terms of the hydrodynamic velocity

$$V = \sum_k N_k V_k / N_+ = \sum_k f_k V_k \quad (5)$$

and the diffusion-like equations for the relative (diffusive) ion velocities $U_k = V_k - V$ [19, 20]. This diffusion approximation implies that $d_k V_k/dt \approx dV/dt$; in other words, the terms proportional to dU_k/dt can be neglected in comparison with the terms proportional to U_k/τ , which are present in the expression for the frictional forces [19]. The equation for the hydrodynamic velocity V can be obtained by summing the equations of motion (1) for all ion species and taking into account

the relationship $\sum_k \sum_j R_{kj} = 0$:

$$mN_+ \frac{dV}{dt} = -\frac{d(N_+ T_+)}{dr} - \langle Z \rangle e N_+ \frac{d\Phi}{dr} + R_e. \quad (6)$$

Here,

$$R_e = \sum_k R_{ke} \approx \frac{e N_e j}{\sigma}, \quad \sigma = \frac{e^2 N_e \tau_e}{\alpha_0 \langle Z^2 \rangle m_e}$$

is the plasma electric conductivity, $\langle Z^2 \rangle = \sum_k Z_k^2 f_k$, $j = I/S$ is the absolute value of the current density, and $S = \Omega r^2$ is the cross-sectional area of the jet. Since, in a steady state, the mass and charge fluxes along a plasma jet, $G = mN_+ VS$ and $I = eN_e(V_e - V)S$, remain unchanged, the electron-to-ion velocity ratio is $V_e/V = \text{const} \approx 1/\eta \gg 1$ and the current strength can be represented as

$$I = eN_e VS/\eta, \quad (7)$$

where $\eta = \langle Z \rangle eG/mI$ is the dimensionless erosion coefficient, which is only weakly dependent on the cathode material, $\eta = 0.05\text{--}0.1$ [1]. We sum up the equation of inertialess electron motion,

$$-\frac{d(N_e T_e)}{dr} + ZeN_e \frac{d\Phi}{dr} - R_e = 0, \quad (8)$$

and Eq. (6) in order to arrive at the equation of plasma motion

$$mN_+ \frac{dV}{dt} = -\frac{d(p_e + p_+)}{dr}, \quad (9)$$

where $p_e = N_e T_e$ and $p_+ = N_+ T_+$ are the electron and ion pressures, respectively.

From Eq. (9), we can see that the main mechanism for plasma acceleration is associated with the pressure gradient, which is maintained at a sufficiently high level by the current that flows through the plasma and heats it.

The equations for the diffusive velocities can be derived from Eqs. (1) by substituting into it the accelerations $d_k V_k/dt \approx dV/dt$ which follow from Eq. (6). Taking into account the inequality $m_e \tau_e \ll m/\tau_+$, we finally arrive at the following set of linear algebraic equations for the relative ion velocities U_k :

$$\sum_j Z_k^2 Z_j^2 f_j (U_j - U_k) = \left(T_+ \frac{d \ln f_k}{dr} + (Z_k - \langle Z \rangle) e \frac{d\Phi}{dr} + \left(1 - \frac{Z_k^2}{\langle Z^2 \rangle} \right) \frac{e \langle Z \rangle j}{\sigma} \right) \tau_+ \quad (10)$$

In order to solve these equations, we must take into account the fact that they are linearly dependent, so that one of them should be replaced with the equation

$$\sum_k f_k U_k = 0, \quad (11)$$

which is a direct consequence of relationship (5).

HEAT BALANCE EQUATIONS

As the current flows through the plasma, the Joule heat is almost completely (to within terms on the order of m_e/m) deposited in the electron gas because of the stochastization of the electron current velocity $V_e - V$ in electron-ion collisions [19]. As a result, the electrons are heated and the plasma becomes thermally nonequilibrium ($T_e > T_+$) and only afterward a fraction of the Joule heat is transferred from the electrons to the ions due to heat-exchange processes. Consequently, without allowance for heat conduction, the heat balance equations for the electrons and ions can be written in the form [19]

$$\frac{3}{2} p_e \frac{d_e}{dt} \ln \left(\frac{p_e}{N_e^\gamma} \right) = \frac{j^2}{\sigma} - Q_T, \quad (12)$$

$$\frac{3}{2} p_+ \frac{d}{dt} \ln \left(\frac{p_+}{N_+^\gamma} \right) = Q_T, \quad (13)$$

where $\gamma = 5/3$ is the adiabatic index and the heat

exchange rate is described by the expression

$$Q_T = \frac{3 \langle Z^2 \rangle m_e N_e}{m \tau_e} (T_e - T_+). \quad (14)$$

We take into account relationships $d_e/dt = V_e d/dr$, $d/dt = V d/dr$, and $V_e \gg V$ and sum up Eqs. (12) and (13). As a result, we arrive at the following heat balance equation in terms of only the electron temperature:

$$\frac{d}{dr} \ln \left(\frac{p_e}{N_e^\gamma} \right) = \frac{2ej}{3\sigma T_e}. \quad (15)$$

HYDRODYNAMIC ION VELOCITY

In order to solve for the hydrodynamic velocity of all ion species, we turn to the cold ion approximation ($T_+ = 0$); as will be shown below, this approximation is valid by virtue of the inequality $\langle Z \rangle T_e \gg T_+$. Taking into account the relationships $d/dt = V d/dr$ and $N_e = \langle Z \rangle N_+$ and using expression (7) and Eqs. (9) and (15), we obtain the equations

$$\frac{dV}{dr} = \frac{2V}{r} (1-f) \left/ \left(\frac{V^2}{V_s^2} - 1 \right) \right., \quad (16)$$

$$\frac{dT_e}{dr} = \frac{2T_e}{r}$$

$$\times \left(\left(\gamma \frac{V^2}{V_s^2} - 1 \right) f - (\gamma - 1) \frac{V^2}{V_s^2} \right) \left/ \left(\frac{V^2}{V_s^2} - 1 \right) \right., \quad (17)$$

where $V_s = (\gamma \langle Z \rangle T_e / m)^{1/2}$ is the local speed of sound and $f(T_e, r) = eI / (5\sigma T_e \Omega r)$.

It is easy to see that, at the spherical surface at which the ion velocity is equal to the local speed of sound, the denominator of the right-hand sides of Eqs. (16) and (17) vanishes. Consequently, for the solution to be continuous, the numerators of the right-hand sides of both these equations should also vanish at this surface. We thus arrive at the requirement $f(T_e, r) = 1$. Denoting the radius of this critical surface by $r = r_*$ and the electron temperature at this surface by T_* and taking into account relationships (4), we can rewrite the requirement as

$$\frac{eI}{5T_* \sigma (T_*) \Omega r_*} = \frac{\langle Z^2 \rangle I}{\langle Z \rangle \Omega r_*} \left(\frac{C_*}{T_*} \right)^{5/2} = 1, \quad (18)$$

where $C_* \approx 0.1 \text{ eV}(\text{cm/A})^{2/5}$.

Under the approximate equality $\langle Z^2 \rangle \approx \langle Z \rangle^2$, expression (18) passes over to the expression obtained previously in [7, 22].

We also introduce the dimensionless coordinate $x = r/r_*$ and the dimensionless temperature $\Theta = T_e/T_*$ and

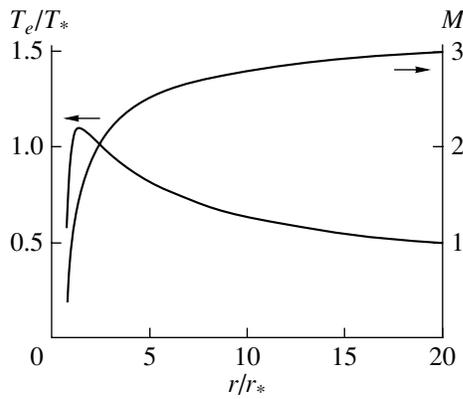


Fig. 1. Profiles of the Mach number and relative electron temperature along the plasma jet.

define the Mach number $M = V/V_*$ as the ratio of the local ion velocity to the speed of sound $V_* = (\gamma\langle Z \rangle T_*/m)^{1/2}$ at the critical surface. Taking into account relationships (18) and $\sigma \propto T_e^{3/2}$, we have $f(T_e, r) = \Theta^{-5/2}x^{-1}$ at an arbitrary surface, in which case Eqs. (16) and (17) become

$$\frac{dM}{dx} = \frac{2M}{x}(1 - \Theta^{-5/2}x^{-1})/(M^2\Theta^{-1} - 1), \quad (19)$$

$$\frac{d\Theta}{dx} = \frac{2\Theta}{x} \left(\left(\frac{5}{3}M^2\Theta^{-1} - 1 \right) \Theta^{-5/2}x^{-1} - \frac{2}{3}M^2\Theta^{-1} \right) / (M^2\Theta^{-1} - 1). \quad (20)$$

The boundary conditions for these equations have the form $M = 1$ and $\Theta = 1$ at $x = 1$. The derivatives at the surface $x = 1$ that are required to solve Eqs. (19) and (20) can be found by applying l'Hôpital's rule to their right-hand sides: $dM/dx = 1.88$ and $d\Theta/dx = 0.75$. One can see that Eqs. (19) and (20) contain no parameters of the cathode jet, so that the solutions $M(x)$ and $\Theta(x)$ are universal functions of the dimensionless distance r/r_* [7]. The profiles of the solutions $M(x)$ and $\Theta(x)$ over distances $x < 20$ from the cathode surface are shown in

Table

r/r_*	M	T_e/T_*
1	1	1
10	2.78	0.63
10^2	3.25	0.26
10^3	3.42	0.10
10^4	3.52	0.03
10^5	3.57	0.01

Fig. 1, and some of the values of the solutions at longer distances x are given in the table.

Taking into account the fact that the critical surface is located at a small distance from the spherical surface from which the plasma is emitted ($\Delta r = r_* - r_0 \approx 0.2r_0$ [7, 22]), we can set $r_* \approx r_0 \approx a$, where $a \approx 1-3 \mu\text{m}$ is the diameter of a cathode microspot [1]. From the table, we can see that, in this case, the Mach number attains its limiting value $M_{\text{lim}} \approx 3.5$ at distances larger than 1 mm from the emitting surface. Since, in all of the published experiments, the ion velocities were definitely measured at larger distances, the ion velocity in a vacuum arc over the main part of the interelectrode gap can be set equal to

$$V_{\text{lim}} = M_{\text{lim}}(\gamma\langle Z \rangle T_*/m)^{1/2} \approx (20\langle Z \rangle T_*/m)^{1/2}. \quad (21)$$

The electron temperature T_* at the critical surface can be calculated from relationship (18). However, a unified viewpoint regarding the diameter of a microspot, $a \approx r_*$, and the current I emitted by the microspot is still lacking. That is why theoretical ideas about the ion acceleration mechanism should be verified by comparing theoretical predictions with the experimentally observed interrelations among the parameters of the cathode plasma jet. This can be done by using relationship (21), which contains exclusively the parameters that can be measured experimentally.

Wieckert [6] showed that, in the case of a copper cathode, the ion composition becomes essentially fixed near the maximum of the electron temperature T_e^m , which corresponds approximately to the temperature at the critical surface, $T_e^m \approx T_*$ (see Fig. 1). Consequently, in the region $r/r_* > 1$, the ion composition can be assumed to be unchanged. For the same reason, the electron temperature values T_e^{exp} obtained by Anders [23] for the fixed ion composition are close to the temperature T_* . Figure 2 compares the ion velocity calcu-

lated from relationship (21) for the values of $T_* \approx T_*^{\text{exp}}$ and $\langle Z \rangle$ taken from [23, 24] with the velocities measured in four series of experiments with cathodes made of different materials. The closed squares, open triangles, and open circles correspond to the ion velocities measured in the experiments of E.M. Oks's group by the methods of short-time current jump [15, 16], current break [16], and current pulse modulation [17], respectively. The crosses show the results of earlier measurements carried out by Davis and Miller [9, 10] with the help of an energy analyzer. From Fig. 2, we can see that the ion velocities measured by Davis and Miller are systematically lower (by 30–50%) than those calculated theoretically or recorded by Oks's group. As was mentioned in [16], a possible cause of this discrepancy is that, in the experiments of Davis and Miller [9], the discharge chamber was kept at an insufficiently high

vacuum so that the ions could have been decelerated by the residual gas.

The results calculated from relationship (21) agree well with the data from measurements by the current jump method in experiments with Mg, Al, Ti, Cu, Pb, Bi, and Zr cathodes [15], the discrepancy being at most 10% (Fig. 2). For the same cathode materials, the disagreement between the experimental results obtained by the current break method [16] and relationship (21) amounts to 20–30%. The ion velocities measured by the current pulse modulation method [17] for 41 cathode materials differ, on average, from the calculated ion velocity by 10%, in which case the velocities measured in experiments with the six lightest cathode materials—C, Mg, Al, Si, Ca, and Ti—systematically exceed (by 50–80%) the calculated velocity and the velocities measured by the other methods. The reason for such a large discrepancy is still unclear. Presumably, light ions, whose gyrofrequency $\omega = eB/mc$ is high, were significantly affected by the generated magnetic field B , which additionally accelerated them outside the cathode region at high currents $I_{\text{arc}} = 400\text{--}600$ A (at which some of the experiments of [17] were carried out).

Figure 3 displays the ion velocities calculated from Eq. (21) for 75 conducting cathode materials. For 50 of these materials, the velocities were calculated using the data available from the measurements of the ion composition and the electron temperatures derived from it [23]. For the remaining 25 materials, the values of $\langle Z \rangle$ and T_e were predicted in [23] on the basis of the periodicity in properties of the chemical elements. That is why the velocities V_{lim} obtained for these 25 materials are merely the ion velocities that might be expected in experiments with vacuum arc cathodes made from the corresponding materials. Figure 3 clearly shows that the ion velocity depends not only on the ion mass m , which increases monotonically (by a factor of 30) with atomic number, but also on the electron temperature and mean ion charge, which both vary periodically by a factor of three, $T_e^{\text{exp}} = 1.5\text{--}4.5$ eV [23] and $\langle Z \rangle = 1\text{--}3$ [23, 24]. One can see that, except for the six lightest cathode materials mentioned above, the calculated ion velocity agrees fairly well with the experimentally measured velocities. We can thus conclude that the primary mechanism by which the ions reach the observed velocities is the gasdynamic mechanism. Now, let us estimate the electron density range in which the ions are accelerated in a vacuum arc. Taking into account the fact that the critical surface is close to the emitting surface and turning to expression (7) for the current strength, we can approximately set $I \approx eN_e V_* a^2 / \eta$. For typical parameter values of a plasma microjet ($I \approx 1\text{--}3$ A, $a \approx 1\text{--}3$ μm , and $\eta \approx 0.05\text{--}0.1$ [1]), this approximate expression and the relationship $V_* = V_{\text{lim}} / M_{\text{lim}} \approx (2\text{--}6) \times 10^5$ cm/s yield $N_e \approx 10^{19}\text{--}10^{20}$ cm $^{-3}$, which agrees with

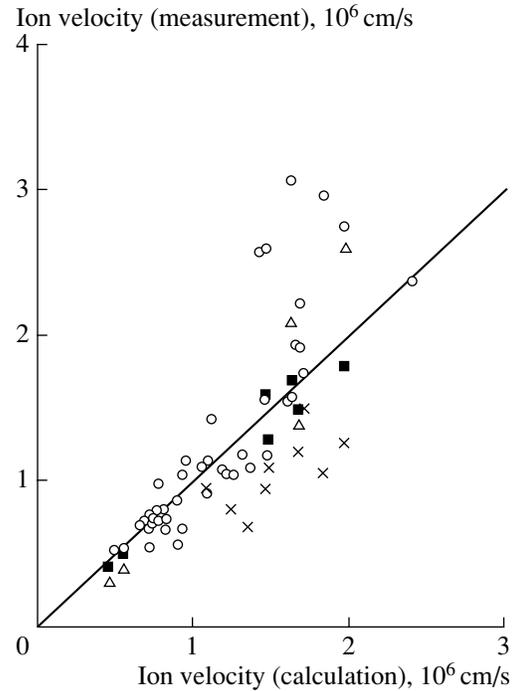


Fig. 2. Comparison of the calculated ion velocity (solid line) with the ion velocities measured in the experiments of [15] (squares), [16] (triangles), [17] (circles), and [9, 10] (crosses).

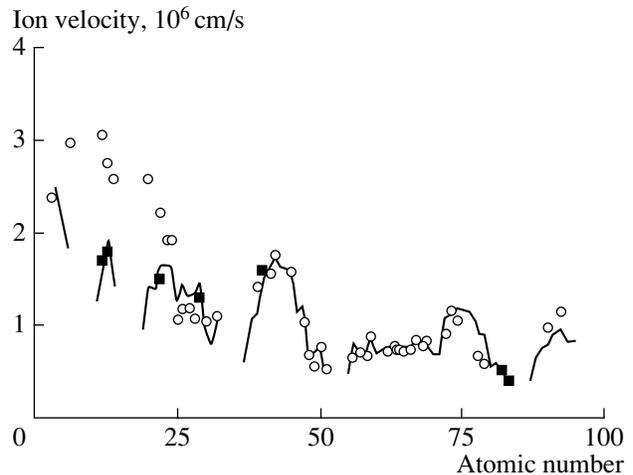


Fig. 3. Ion velocity calculated from relationship (21) as a function of the atomic number of the cathode material (solid curve) and ion velocities measured by the current jump method [15, 16] (squares) and by the current pulse modulation method [17] (circles).

the estimates $N_e \approx 10^{18}\text{--}10^{20}$ cm $^{-3}$ obtained in [22, 23] for a region where the ion composition is frozen. From the table, we can see that the ions are accelerated primarily over distances of about $r \approx (1\text{--}100)a$ from the emitting surface and that the electron density over these distances decreases to $N_e \approx 10^{14}\text{--}10^{16}$ cm $^{-3}$.

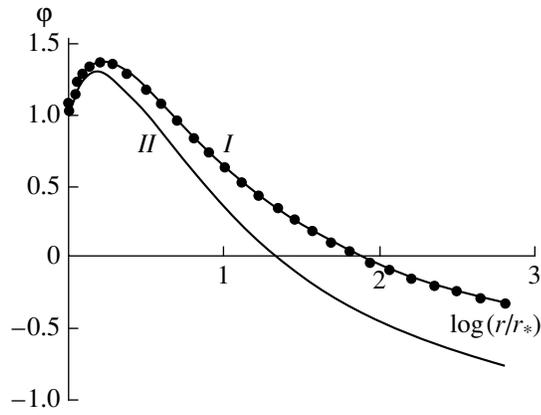


Fig. 4. Potential profile along the axis of the plasma jet for $\eta = (I) 0.1$ and $(II) 0.05$.

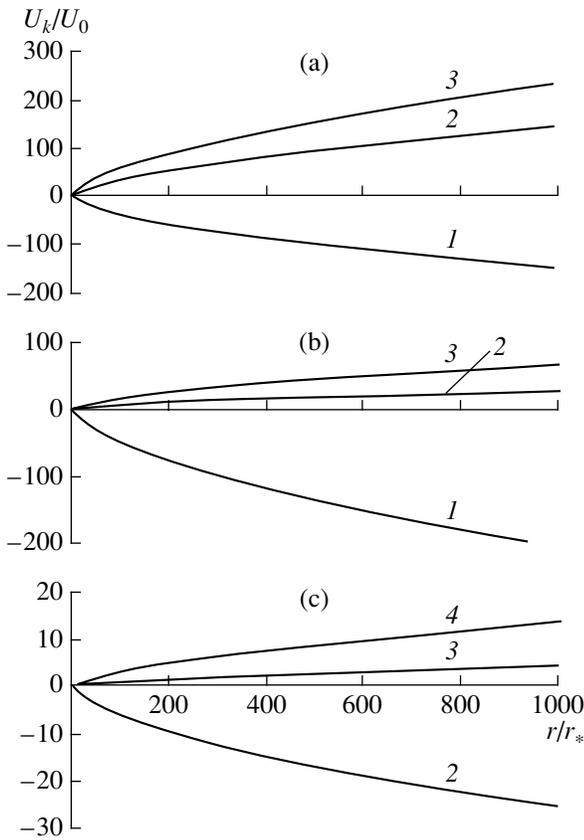


Fig. 5. Profiles of the relative diffusive velocities of different ion species along the plasma jet for $\langle Z \rangle = (a) 1.5$, $(b) 2$, and $(c) 3$. The numerals above the curves show the ion charge number.

DIFFUSIVE ION VELOCITY

Equation (10) implies that the main reasons for the deviation of the velocities $V_k = V + U_k$ of different ion species from the hydrodynamic velocity V are (i) the effect of the density gradient of a particular ion species, (ii) the dependence of the acceleration rate in an electric

field on the difference $Z_k - \langle Z \rangle$, and (iii) the dependence of the ion–electron frictional force on $Z_k^2 - \langle Z^2 \rangle$. As was shown above, $f_k \approx \text{const}$ in the region $r > r_*$. Consequently, the diffusive velocities U_k are different mainly because of the difference of the charge of a particular ion species from the mean ion charge. In order to solve Eqs. (10), we switch to the dimensionless velocities $u_j = U_j/U_0$ and represent the equations in standard form:

$$\sum_j b_{kj} u_j = F_k, \tag{22}$$

$$F_k(x) = x^2 M(Z_k - \langle Z \rangle) \frac{d\phi}{dx} + 5 \langle Z \rangle M t^{-3/2} \left(1 - \frac{Z_k^2}{\langle Z^2 \rangle} \right), \tag{23}$$

$$U_0 = \frac{\alpha_0 \langle Z^2 \rangle}{5\eta} \left(\frac{2m_e}{m} \right)^{1/2} \left(\frac{T_+}{T_*} \right)^{3/2} V_*, \tag{24}$$

where $b_{kj} = Z_k^2 Z_j^2 f_j$ ($k \neq j$) and $b_{kk} = Z_k^4 f_k - Z_k^2 \langle Z^2 \rangle$.

When deriving Eqs. (23) and (24), we used relationships (4), (7), and (18) and introduced the dimensionless potential $\phi = e\Phi/T_*$. In [7, 22], it was shown that this potential can also be written as $\phi = 1 + 5/6\eta(M^2 - 1) + 5/2(\Theta - 1)$ and that, by virtue of Eqs. (19) and (20), its gradient is an explicit algebraic function of M , Θ , and η . Figure 4 presents the dependence $\phi(x)$ calculated for the values $\eta = 0.05$ and 0.1 , which bound the main range of η values under consideration.

Equations (22) and (23) imply that the relative diffusive ion velocities u_k are uniquely determined by the universal functions $M(x)$ and $\Theta(x)$ and the relative ion content (i.e., the relative ion fractions f_k and the corresponding values of $\langle Z \rangle$ and $\langle Z^2 \rangle$) and that they are independent of the ion mass, current, and the characteristic dimension of a plasma microjet. The calculations showed that acceleration by the potential gradient $d\phi/dx$ [the first term on the right-hand side of Eq. (23)] contributes only about 20–25% of the velocity u_k (this contribution coincides with the corresponding estimate made by Hantzsche [18] for a Cu cathode). Thus, we can conclude that the difference in diffusive ion velocities originates primarily from different frictional forces acting between the electrons and different ion species. Figure 5 shows the relative diffusive velocities $u_k(x)$ of different ion species that were obtained by solving Eqs. (11) and (22) for $\eta = 0.1$ and the three typical ion charge distributions: (a) $f_1 = 0.5, f_2 = 0.5, f_3 = 0.001, \langle Z \rangle = 1.5$, and $\langle Z^2 \rangle = 2.5$ (for cathodes made from Mg, Mn, and Sn); (b) $f_1 = 0.15, f_2 = 0.7, f_3 = 0.15, \langle Z \rangle = 2$, and $\langle Z^2 \rangle = 4.3$ (for cathodes made from Ti, Cu, Cr, Pt, and Au); and (c) $f_1 = 0, f_2 = 0.2, f_3 = 0.6, f_4 = 0.2, \langle Z \rangle = 3$, and $\langle Z^2 \rangle = 9.4$ (for cathodes made from Nb, Ta, W, and Hf).

The profiles calculated for $\eta = 0.05$ differ from those shown in Fig. 5 by less than 5%. We can see that the larger the mean ion charge, the smaller the difference $u_k - u_j$ in the relative diffusive ion velocities. The reason for this is that the ion-ion frictional force (3) increases in proportion to $(Z_k Z_j)^2$. Consequently, the difference in the relative velocities is the largest, $u_k - u_j = 10^2 - 10^3$, at $\langle Z \rangle = 1.5$ (Fig. 5a).

It should be noted that the differences in the absolute ion velocities, $V_k - V_j = U_k - U_j = (u_k - u_j)U_0$, depend largely on the quantity $U_0 \propto T_+^{3/2}$. The ion temperature $T_+(x)$ can be found from Eq. (13) and expression (14). Using the above dimensionless variables and defining the dimensionless ion temperature as $\Theta_+ = T_+/T_*$, we arrive at the equation $d\Theta_+/dx + p\Theta_+ = q$, where $p = 2[2M^2(\Theta^{3/2}M^2x - 1) + 9(M^2 - \Theta)\eta/\alpha_0][3\Theta^{3/2}M^2x^2(M^2 - \Theta)]$ and $q = 6\eta(\alpha_0\Theta^{1/2}M^2x^2)^{-1}$. Formally, the solution to this equation can be written in terms of definite integrals of the functions $p(x)$ and $q(x)$. The solution simplifies substantially in the region $x > 2$, where $M^2 \gg \Theta$, in which case the product $\Theta^{1/2}M^2 \approx 4-6$ can be assumed to vary gradually over the entire region $x = 2-1000$ (see table and Fig. 1). For the most typical value of the erosion coefficient, $\eta = 0.1$, we have $9\eta/\alpha_0 = 2$ and $p(x) = 4/(3x)$, and the solution takes the form $\Theta_+(x) \approx 4(\Theta^{1/2}M^2x)^{-1}$. If we approximately set $\Theta^{1/2}M^2 \approx 4$, the solution $\Theta_+(x) \approx 1/x$ can be extended to the region $2 \geq x \geq 1$, provided that the electron and ion temperatures at the critical surface $x = 1$ are the same. Then, we have $T_+(x) \approx T_*(r_*/r)$. A comparison with Fig. 1 shows that the ion temperature decreases more sharply than the electron temperature. Substituting the ion temperature into Eq. (24) yields

$$U_0 \approx 3 \times 10^{-3} \frac{\langle Z^2 \rangle}{\eta \mu^{1/2} x^{3/2}} V_*, \quad (25)$$

where $\mu = m/m_p$ and m_p is the proton mass.

From the profiles shown in Fig. 5 and relationship (25), we can see that, even in the limiting case in which the ion charge is large, the ion mass is small ($\langle Z^2 \rangle / \sqrt{\mu} \approx 1$), and the erosion coefficient is low ($\eta = 0.05$), the diffusive velocities do not exceed the value $U_k = u_k U_0 \approx 10^{-2} V_*$, while the hydrodynamic velocity is about $V \approx (1-3)V_*$. Consequently, in all possible cases, the velocities of different ion species in low-current vacuum arcs should differ by less than 1%. This conclusion agrees completely with the results that were obtained experimentally in [15-17] and showed no dependence of the ion velocities on the ion charge. In early experimental works [9, 14], it was observed that ions with larger charges were accelerated to higher velocities. Bugaev *et al.* [16] explained this fact as being due not to the acceleration but rather to the deceleration of ions by the residual gas: the results of their measurements revealed

that the ions with a smaller charge are decelerated more strongly than the ions with a larger charge.

Note that the presence of the potential hump (Fig. 4) is often regarded as an argument in favor of significant ion acceleration under the action of the electric field, which gives rise to differences in the ion velocities. In actuality, a humped potential profile serves merely to maintain plasma quasineutrality and a constant electron-to-ion current ratio and plays a relatively minor role in ion acceleration.

DISCUSSION OF THE RESULTS

The above theoretical investigation of the characteristic features of the acceleration of a multicomponent current-carrying plasma (with a fixed ion composition) during its expansion into vacuum and a comparison of the theoretical results with a large amount of experimental data make it possible to draw the following conclusions.

(i) The cathode material is accelerated primarily after the phase transition to a plasma state. The plasma produced is accelerated during the gasdynamic expansion into vacuum over distances $r \leq 100a$ (where $a = 1-3 \mu\text{m}$ is the diameter of an emission center) from the cathode surface. Over these distances, the electron density decreases from about $N_e \approx 10^{18}-10^{20}$ to $\approx 10^{14}-10^{16} \text{ cm}^{-3}$. In the main acceleration stage, the ions acquire more than 70% of their velocity and, accordingly, more than 90% of their energy. Of course, in the initial stage of motion (at $V < V_*$), the cathode material can also be accelerated during phase transitions to a liquid metal and then to a superdense gas [2, 3]. However, in this initial acceleration stage, the ions acquire no more than 20% of their final velocity.

(ii) In the acceleration region, the ions move with the hydrodynamic velocity, which is determined by the ion mass, mean ion charge, and the maximum electron temperature in the cathode region. The calculated velocities to which the ions are accelerated agree well with the extensive data from experiments with cathodes made of different materials.

(iii) The velocities of different ion species differ only slightly (by less than 1%) from the hydrodynamic velocity of the plasma jet.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project nos. 99-02-18163 and 00-02-16792.

REFERENCES

1. G. A. Mesyats, *Ectons in Vacuum Arc Discharge: Breakdown, Spark, Arc* (Nauka, Moscow, 2000).
2. N. B. Volkov and A. Z. Nemirovskii, *J. Phys. D* **24**, 693 (1991).

3. A. Z. Nemirovskii and E. A. Litvinov, in *Proceedings of the 12th Symposium on High Current Electronics, Tomsk, Russia, 2000*, p. 60.
4. G. A. Lyubimov, Dokl. Akad. Nauk SSSR **225**, 1045 (1975) [Sov. Phys. Dokl. **20**, 830 (1975)].
5. B. Ya. Moïzhes and V. A. Nemchinskii, Zh. Tekh. Fiz. **50**, 78 (1980) [Sov. Phys. Tech. Phys. **25**, 43 (1980)].
6. C. Wieckert, Contrib. Plasma Phys. **27**, 309 (1987).
7. I. A. Krinberg, M. P. Lukovnikova, and V. L. Papernyi, Zh. Éksp. Teor. Fiz. **97**, 806 (1990) [Sov. Phys. JETP **70**, 451 (1990)].
8. A. A. Plyutto, V. N. Ryzhkov, and A. T. Kapin, Zh. Éksp. Teor. Fiz. **47**, 494 (1964) [Sov. Phys. JETP **20**, 328 (1964)].
9. W. D. Davis and H. C. Miller, J. Appl. Phys. **40**, 2212 (1969).
10. J. Kutzner and H. C. Miller, J. Phys. D **25**, 686 (1992).
11. I. G. Brown and J. E. Galvin, IEEE Trans. Plasma Sci. **17**, 679 (1989).
12. S. P. Gorbunov, V. I. Krasov, and V. L. Papernyi, J. Phys. D **30**, 922 (1997).
13. I. A. Krinberg, J. Phys. D **29**, 2049 (1996).
14. V. M. Lunev, V. G. Padalka, and V. M. Khoroshikh, Zh. Tekh. Fiz. **47**, 1491 (1977) [Sov. Phys. Tech. Phys. **22**, 858 (1977)].
15. A. S. Bugaev, V. I. Gushenets, A. G. Nikolaev, *et al.*, IEEE Trans. Plasma Sci. **27**, 882 (1999).
16. A. S. Bugaev, V. I. Gushenets, A. G. Nikolaev, *et al.*, Zh. Tekh. Fiz. **70** (9), 37 (2000) [Tech. Phys. **45**, 1135 (2000)].
17. G. Yu. Yushkov, A. Anders, E. M. Oks, *et al.*, J. Appl. Phys. **88**, 5618 (2000).
18. E. Hantzsche, IEEE Trans. Plasma Sci. **23**, 893 (1995).
19. S. I. Braginskii, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1963), Vol. 1.
20. J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954; Inostrannaya Literatura, Moscow, 1961).
21. I. A. Krinberg and M. P. Lukovnikova, J. Phys. D **29**, 2901 (1996).
22. I. A. Krinberg, Phys. Plasmas **1**, 2822 (1994).
23. A. Anders, Phys. Rev. E **55**, 969 (1997).
24. I. G. Brown, Rev. Sci. Instrum. **65**, 3061 (1994).

Translated by O. Khadin

GAS DISCHARGES, PLASMA

Physical Processes in the Final Phase of the Plasma Focus

Yu. N. Zhukov, A. I. Markoliya, A. F. Popov, and A. F. Chachakov

Sukhumi Physicotechnical Institute, Sukhumi, 384914 Abkhazia, Georgia

Received March 2, 2000; in final form, January 31, 2001

Abstract—A description of the final phase of the plasma focus is proposed based on the classical plasma physics concept. It is shown that, taking into consideration the cumulative plasma heating and the occurrence of a shock wave reflected from the axis, it is possible to qualitatively explain the experimentally observed phenomena, such as current disruption, the generation of a strong electric field, and other accompanying effects. If a weak poloidal magnetic field is initially present in the plasma column, then this field can increase in magnitude to a value comparable to the azimuthal magnetic field. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

A detailed review of the current state of the experimental and theoretical research on high-power pulsed discharges is presented in [1–3] (see also the literature cited therein). The physics of a number of processes is now well understood. However, some effects observed after the plasma implosion at the axis of the system, such as a sharp decrease in the plasma conductivity and the generation of different types of radiation, are still poorly understood. There are different points of view on the physics of the processes occurring in this phase. The best developed are the model of thermonuclear neutron generation in a deuterium plasma with allowance for the axial outflow of particles from the high-temperature region [2] and the model describing the generation of hard radiation due to the acceleration of particles in the induced electric field [3]. The deuteron spectra and the neutron yield calculated in [3] are in good agreement with the experimental data. In that paper, however, the cumulative plasma heating and the possible occurrence of a shock wave reflected from the axis were not taken into consideration. In the present paper, it is shown that, taking these factors into consideration, we can explain, at least qualitatively, the experimentally observed phenomena, such as current disruption (without invoking microscopic instabilities), the generation of an electric field, and other accompanying effects. Furthermore, the proposed model describes the increase in the magnitude of an initially weak longitudinal magnetic field to a value comparable to the azimuthal magnetic field of the pinch.

PHYSICS OF THE PROCESSES OCCURRING IN A PINCH AFTER THE ARRIVAL OF A SHOCK WAVE AT THE SYSTEM AXIS

Experiments show [1, 4] that a pinch discharge begins with a gas breakdown near an insulator, where a plasma current-carrying shell is formed. Under the action of the electrodynamic force, the shell is acceler-

ated toward the system axis and drives the neutral gas. As the converging shock wave approaches the system axis, the current-carrying shell collapses [5]. The temperature and pressure at the front of the wave grow, whereas the density remains constant and equal to the density at the front of a strong shock wave. The collapse is followed by the appearance of a reflected shock wave, which propagates in the direction opposite to the nonsteady plasma flow.

For the process to develop as described above, the plasma column should be axially uniform. However, the highest plasma parameters have been obtained in the plasma column constrictions, where the plasma ejection in the longitudinal direction plays an important role. This is typical of the plasma focus, in which the current-carrying shell takes the shape of a funnel with the throat facing the system exit. However, in this case too, the implosion process turns out to be similar to that considered above if the angle between the axis and the tangent to the current-carrying shell is less than the critical angle $\alpha_c = \arcsin 1/\gamma$, where γ is the adiabatic index [6].

The plasma temperature in the compression phase of the current-carrying shell is not high because of great ionization and radiation losses and is on the order of ten electronvolts [4]. For this reason, the electron magnetization factor is low ($\omega_{ce}\tau_e \leq 1$, where $\omega_{ce} = eH/m_e c$ is the electron cyclotron frequency and τ_e is the electron scattering time) and the current and the electric field are simply related by Ohm's law. In the course of implosion, the plasma temperature increases substantially and, in high-power discharges, reaches several kiloelectronvolts. As a result, the magnetization factor becomes much higher than unity and the conduction conditions change significantly. As was shown in [7], Ohm's law in its simple form is applicable to a high-temperature plasma only if the equilibrium condition is satisfied:

$$\nabla P = (1/c)\mathbf{J} \times \mathbf{H}, \quad (1)$$

where P is the plasma pressure.

It follows from Eq. (1) that the density of the current flowing across the magnetic field is equal to

$$\mathbf{J} = c[\mathbf{H} \times \nabla P]/H^2 \quad (2)$$

and is fully determined by the plasma pressure gradient. A specific feature of the reflected shock wave arising in the final phase of the focus [5] is that the plasma pressure varies severalfold near the shock front and varies only slightly far from the shock. As a result, behind the front of the reflected shock wave, the density of the current driven by the pressure gradient is small and does not substantially contribute to the total discharge current. It is believed that a small charge separation (such that the plasma remains quasineutral) gives rise to the Hall electric field [7]

$$\begin{aligned} \mathbf{E}_H &= (1/eN_e c)\mathbf{J} \times \mathbf{H} = (\omega_{ce}\tau_e/\sigma_{\perp 0}H)\mathbf{J} \times \mathbf{H}, \\ |E_H| &= \omega_{ce}\tau_e E_0, \end{aligned} \quad (3)$$

where $E_0 = J_0/\sigma_{\perp 0}$ and $\sigma_{\perp 0} = 10^7 T_e^{3/2}$ is the transverse conductivity of the unperturbed plasma. This field balances the electrodynamic force acting on the electrons and restores the plasma current across the magnetic field. In this case, the ions are assumed to be in equilibrium with the Hall field, which is not obvious *a priori* for a high-temperature plasma.

Hence, we can expect that the propagation of the reflected shock wave is accompanied by both a wave in which the conduction current is suppressed and the generation of the displacement current. However, because of the high value of the discharge current, the induced electric field is fairly high, so that the ions and the Hall electric field begin to play an important role in the electric conduction. Indeed, at $\omega_{ce}\tau_e \gg 1$ the electrons drift in the radial direction under the action of the crossed electric and magnetic fields and only rare collisions lead to their displacement along the vector \mathbf{E} , contributing insignificantly to the longitudinal current density. Over a time $t < T_{ci} = 2\pi/\omega_{ci}$ (where $\omega_{ci} = eH/m_i c$ is the ion cyclotron frequency), the ions do not yet drift and the electrons drift with respect to the ions, thus producing the radial electric current $J_r = eN_e(cE_z/H_\phi - U_{ir})$, where U_{ir} is the radial ion velocity. This current results in charge separation. The produced space charge density is equal to

$$\rho = -\int \nabla J_r dt. \quad (4)$$

Substituting this expression into the equation $\nabla \cdot \mathbf{E} = 4\pi\rho$, we obtain for the radial electric field

$$E_r = 4\pi e N_e \Delta r, \quad (5)$$

where

$$\Delta r = \int (cE_z/H_\phi - U_{ir}) dt.$$

In principle, this field can restore the plasma current across the electric field. Evidently, the electric field in this case must be equal to the Hall field.

To analyze the further evolution of the system, we use the equation of ion motion

$$\begin{aligned} N_e m_i d\mathbf{U}_i/dt &= eN_e \mathbf{E} + (eN_e/c)\mathbf{U}_i \times \mathbf{H} \\ &- \nabla P_i - (m_e N_e/\tau_e)(\mathbf{U}_i - \mathbf{U}_e). \end{aligned} \quad (6)$$

In a sufficiently powerful discharge, the plasma behind the front of the reflected shock wave is strongly heated and the inequality $\omega_{ce}\tau_e \gg 1$ is satisfied in this region. As a result, the Hall electric field (3) is high. At the same time, we can assume that the pressure gradient and the friction force are small as compared to the electric force and, thus, can neglect them. In experiments, the following condition is also satisfied: $\omega_{pi} \gg \omega_{ci}$, where $\omega_{pi} = (4\pi e^2 N_e/m_i)^{1/2}$ is the ion plasma frequency. Then, using Eq. (5), we obtain for the radial component of the ion velocity

$$dU_{ir}/dt = \omega_{pi}^2 \Delta r. \quad (7)$$

According to this equation, the ions are continuously accelerated by a strong electric field. As a result, the degree of charge separation, the strength of the radial electric field, and, consequently, the longitudinal current decrease, which is equivalent to an increase in the plasma resistivity across the magnetic field. Owing to self-induction, the longitudinal electric field will be generated in the plasma; this field will keep the constant longitudinal current. However, since the radial ion velocity in this case increases with time, the electron drift velocity also should increase in order to keep the constant degree of charge separation. The latter is possible only if the radial electric field decreases and E_z increases. It is obvious that this process will be non-steady and the ions will substantially contribute to it.

Since the induced longitudinal electric field is considerably higher than the critical Dricer field $E_d = 4\pi e^3 N_e \ln \lambda/T_e$, almost all the plasma electrons in the region of a low magnetic field near the plasma column axis should become runaway. However, since the acceleration occurs in a dense quasineutral plasma, the maximum current produced by the beam of runaway electrons is determined by the condition of self-limitation of the current by its proper magnetic field and is equal to the limiting Alfvén current [8]

$$I_a = m_e c^3/e\beta\gamma = 17\beta\gamma \text{ kA}, \quad (8)$$

where $\gamma = (1 - \beta^2)^{-1/2}$ is the relativistic factor and $\beta = v/c$.

If we equate the classical expression for the current $I_k = \pi r^2 e N_e v$ to the Alfvén current, then we obtain the following expression for the beam radius:

$$r_a = 2c/\omega_{pe}\gamma^{1/2}, \quad (9)$$

where $\omega_{pe} = (4\pi e^2 N_e / m_e)^{1/2}$ is the electron plasma frequency.

The limiting beam current at electron energies of several hundreds of kiloelectronvolts is equal to several tens of kiloamperes and is significantly less than the pinch current. However, its cross size is very small and, consequently, the azimuthal magnetic field of the beam $H_\phi = 0.2I_a/r_a$ is high; it is much higher than the pinch magnetic field in this region. As a result, the magnetic field profile change significantly, which makes the problem more complicated.

Depending on the radial profiles of the current and magnetic field, the evolution of the system can substantially differ from that discussed above. In the case of strong skinning ($\delta \ll r$), the skin depth increases because the plasma resistance in the skin layer increases as the reflected shock wave propagates through it. Because of the suppression of the longitudinal current, a strong electric field is generated in the layer. The electrons are accelerated by this field and are deflected by the magnetic field of the pinch toward the plasma column axis. As a result, a hollow electron beam with a relatively low energy is formed. The current of this hollow beam is on the order of $I_p = r/\Delta r_0 I_a$, where Δr_0 is the thickness of the beam wall.

Approximate equations describing the processes occurring in the plasma focus after the plasma implosion at the axis can be derived from Maxwell's equations and the equations of ion motion. In cylindrical coordinates, these equations can be written in the form

$$\partial E_z / \partial r = 1/c \partial H_\phi / \partial t, \quad (10)$$

$$\begin{aligned} & 1/r \partial / \partial r (r H_\phi) \\ & = 4\pi/c \{ \sigma_\perp / (1 + \omega_{ce}^2 \tau_e^2) (E_z + \omega_{ce} \tau_e \mathbf{h} \times \mathbf{E}_r) \\ & + e N_e U_{iz} + 1/4\pi \cdot \partial E_z / \partial t + J_a \}. \end{aligned} \quad (11)$$

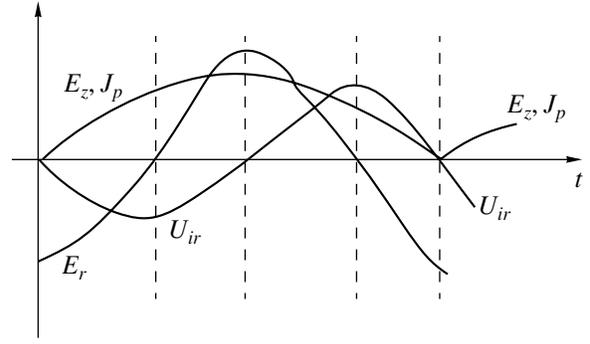
The longitudinal current density is equal to the sum of the densities of the electron conduction current, the ion current, the displacement current, and the electron beam current J_a . Differentiating Eq. (5) with respect to time, we obtain for the radial current density

$$J_r = e N_e U_{ir} - e N_e c E_z / H_\phi + 1/4\pi \partial E_z / \partial t = 0. \quad (12)$$

Neglecting the pressure gradient and the friction force behind the shock front, we obtain for the ion velocity components

$$\begin{aligned} dU_{ir}/dt &= e/m_i E_r - \omega_{ci} U_{iz}, \\ dU_{iz}/dt &= e/m_i E_z + \omega_{ci} U_{ir}. \end{aligned} \quad (13)$$

The initial conditions can be obtained taking into account the fact that, after the suppression of the con-



Qualitative picture of the time behavior of the current, electric field components, and the radial ion velocity behind the front of the reflected shock wave in the plasma focus for the case in which the electron conduction current exceeds the ion current.

duction current, it is substituted with sum of the displacement current and the Hall current:

$$\begin{aligned} & (-e N_e c E_r / H_\phi + 1/4\pi \partial E_z / \partial t)_{t_1} = \sigma_{\perp 0} E_{z0}, \\ & E_a(r_1, t_1) = 0, \quad U_{iz}(r_1, t_1) = U_{ir}(r_1, t_1) = 0, \quad (14) \\ & t_1 r_1 / v_d, \end{aligned}$$

where v_d is the velocity of the diverging shock wave.

Since the shock wave is weak, its velocity is close to the speed of sound. The exact initial distributions of the plasma density and the azimuthal magnetic field are unknown. The magnetic field distribution can be approximated by either a function increasing linearly toward the boundary of the plasma column or a function depending more strongly on the radius. The plasma density in the diverging shock wave is minimum at the axis and decreases with time.

An analysis of the obtained set of equations allows us to draw a number of qualitative conclusions concerning the character of the processes occurring in the plasma focus. Since the radial drift current is high, we can assume that the Hall electric field $E_r = \omega_{ce} \tau_e E_0$ arises in the plasma column almost instantaneously and the pinch current is restored. However, the ions are then accelerated by this field and, consequently, the radial electric field and the longitudinal Hall current decrease. The resistance of the current channel increases. As a result, the longitudinal electric field E_z is generated. The maximum value of this field E_{zm} is limited by a value at which the total longitudinal pinch current (without taking into account the Hall current) is equal to its initial value. If the conduction electron current $J_p = \sigma_\perp E_z / \omega_{ce}^2 \tau_e^2$ is higher than the ion current, then we have $E_{zm} = 2\omega_{ce}^2 \tau_e \tau_{e0} E_0$. The time behavior of the currents and electric fields for this case is qualitatively shown in the figure. As a result, the ions execute radial oscillations. Since the radial electron drift velocity is proportional to E_z , the electrons are displaced toward

the interior of the current column at a velocity that varies with time and, consequently, the ion oscillation half-periods are different.

If the ion current is higher than the electron longitudinal current, then the Hall electric field disappears in a certain time on the order of the ion cyclotron period. The ions drift begins, and the polarization current

$$\mathbf{J}_n = (N_e m_i c^2 / H^2) \partial \mathbf{E} / \partial t \quad (15)$$

makes a major contribution to the pinch current. It is assumed that, since the plasma flows freely out of the constriction region, there are no conditions for the appearance of an appreciable pressure gradient that is sufficient for the longitudinal pinch current to be restored. This is equivalent to a significant increase in the longitudinal resistance of the pinch. In the general case, since the plasma in the current channel is radially nonuniform and the processes are nonlinear, the onset of turbulent motion with the generation of a broad spectrum of high-frequency pulsation at frequencies on the order of the ion plasma frequency is expected.

Similarly to the electrons, a fraction of the ions in the plasma column execute betatron oscillations near the pinch axis, thereby continuously gaining longitudinal energy. Since the magnetic field increases strongly in this region, the ion beam is produced in the channel of the accelerated electron beam. The generated field can vary periodically, so that the energy of the accelerated particles is modulated and the high-energy particles propagate in the form of bunches with a duration on the order of the ion oscillation period [9].

One of the most important characteristics of the plasma focus is a fairly high hard X-ray and neutron yield. It is found experimentally that hard X radiation is produced when the accelerated electron beam interacts with the anode material. The delay time of radiation with respect to the instant of the maximum pinch compression (20–30 ns) agrees satisfactorily with the propagation time of the reflected shock wave.

The neutron yield per pulse in a deuterium discharge is $\sim 10^{10}$ at a discharge current on the order of 1 MA. It is found experimentally that neutron radiation is spatially anisotropic. In the proposed physical model, two groups of ions contribute to neutron radiation. The ions of the first group gain their energy in the course of implosion and in the generated electric field. They interact with the pinch plasma over a long period until they lose their energy via Coulomb collisions. The ions of the second group form a high-energy ion beam flowing out of the constriction region; the energy of these ions is $eE_z L$, where L is the constriction length. This ion flow generates radiation by interacting with a target formed by both the cumulative plasma jet and the residual gas. Qualitatively, this model provides a satisfactory explanation for the radiation anisotropy and the large length of the neutron source [1].

ENHANCEMENT OF THE LONGITUDINAL MAGNETIC FIELD AND FORCE-FREE CONFIGURATIONS IN THE PLASMA FOCUS

It is evident that a weak longitudinal magnetic field on the order of several oersted cannot substantially affect the final plasma parameters in the plasma focus. However, after the collapse of the current-carrying shell, the longitudinal magnetic field increases and becomes on the order of the azimuthal magnetic field [10–12]. For this reason, the role of the magnetic field in the further evolution of the plasma focus is still poorly understood and a study of the possible mechanisms for this process is of considerable interest.

If the condition $4\pi\Delta x^2/c^2 < r(t)/v(t)$ (where Δx is the plasma layer thickness and $v(t)$ is the velocity of the current-carrying shell) is satisfied in the phase of the shell collapse, then the longitudinal magnetic field is frozen in the plasma and the magnetic flux in the plasma column is conserved. The strength of the longitudinal magnetic field at the instant of maximum compression is $H_z = H_0(r_0/r_1)^2$, where H_0 is the initial magnetic field strength and r_0 and r_1 are the initial and final radii of the plasma column, respectively. When the current-carrying shell is compressed starting from a large radius, the resulting magnetic field can be fairly strong. Thus, for $r_0 = 5$ cm, $r_1 = 0.2$ cm, and $H_0 = 1$ Oe, the final magnetic field strength is equal to $H_z = 625$ Oe.

After the instant of implosion in a high-temperature plasma, the pressure in the shock wave reflected from the axis levels off and the poloidal magnetic field continues to increase. Indeed, if we neglect the inertia, then the vector of the current density in a magnetized plasma is determined by the expression

$$\mathbf{J} = \mathbf{J}_\parallel + \mathbf{J}_\perp = \sigma(\mathbf{E} \cdot \mathbf{H})/H^2 \cdot \mathbf{H} + c[\mathbf{H} \times \nabla P]/H^2. \quad (16)$$

Since the pressure behind the shock-wave front levels off, it cannot maintain the initial current in the plasma column. Hence, we may assume that the propagation of the reflected shock wave is accompanied by the suppression of the transverse (with respect to the magnetic field) component of the initial conduction current $\mathbf{J} = (1/4\pi)\partial \mathbf{E}_\perp / \partial t$.

As in the case considered above, the induced electric field \mathbf{E}_\perp in a magnetized plasma causes the particles to drift in the radial direction and does not produce an appreciable current along the \mathbf{E}_\perp vector. Since the electrons begin to drift earlier than the ions and the density of the electron drift current $eN_e c E_\perp / H$ is fairly high, the separation of charges in the plasma leads to the generation of the Hall electric field, which restores the pinch current. However, if the ion cyclotron frequency ω_{ci} is much higher than the ion scattering frequency ν_i , then the ions also begin to drift and the Hall electric field disappears.

The main contribution to the transverse pinch current in this stage comes from the polarization current,

whose density is determined by expression (15). The polarization current is high only if the electric field is strong enough, which is equivalent to a substantial increase in the pinch resistance across the magnetic field produced by the transverse current. In this case, the plasma drifts in the crossed electric and magnetic fields. Since the plasma flows freely out of the constriction region, the drift cannot give rise to an appreciable pressure gradient in this region and, thus, the transverse current cannot be restored. The plasma tends to a force-free state. It follows from expressions (15) and (16) that, in this case, the poloidal current component decreases and the azimuthal component increases because the contributions from the polarization current to these components have opposite signs. Consequently, the poloidal magnetic field behind the front of the reflected shock wave continues to increase. From expression (16) for the parallel current density, it follows that the longitudinal and azimuthal magnetic fluxes are related to each other and, under certain conditions, either magnetic flux can be sustained at the expense of the energy of the other.

Hence, in a high-temperature plasma, the transverse current in the constriction is restored because of the rotation of the magnetic field vector and the transition of the plasma to the force-free state.

During the transition, a fairly strong electric field is induced in the constriction. The plasma particles are accelerated in this field and can gain a substantial energy over one cyclotron period. As a result, a fraction of the magnetic field energy dissipates and transforms into the thermal energy of the plasma particles.

CONCLUSION

According to the predictions of our model, which are in qualitative agreement with experimental data, the following basic effects should be observed after the pinch plasma implosion: an increase in the ohmic resistance of the pinch, resulting in the generation of a strong longitudinal electric field; the generation of superthermal electromagnetic radiation at frequencies on the order of the ion plasma frequency; the formation of an electron beam and the acceleration of the plasma ions, accompanied by the emission of neutron and hard

X radiation; and the modulation of the energy and density of the accelerated particles at the frequencies of electromagnetic radiation. If a weak poloidal magnetic field is initially present in the constriction, then this field should increase in magnitude to a value comparable to the azimuthal magnetic field of the pinch.

The results of this study confirm the acceleration mechanism for neutron generation in pinches proposed by Trubnikov [3].

REFERENCES

1. V. A. Burtsev, V. A. Gribkov, and T. I. Filippova, *Itogi Nauki Tekh., Ser. Fiz. Plazmy* **2**, 80 (1981).
2. V. V. Vikhrev, *Fiz. Plazmy* **12**, 454 (1986) [*Sov. J. Plasma Phys.* **12**, 262 (1986)].
3. B. A. Trubnikov, *Fiz. Plazmy* **12**, 468 (1986) [*Sov. J. Plasma Phys.* **12**, 271 (1986)].
4. V. V. Vikhrev and S. I. Braginskii, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Atomizdat, Moscow, 1980; Consultants Bureau, New York, 1986), Vol. 10.
5. J.P. Somon, in *The Physics of High Energy Density*, Ed. by P. Caldirola and H. Knoepfel (Academic, New York, 1971; Mir, Moscow, 1974).
6. S. I. Anisimov, A. L. Velikovich, N. G. Koval'skiĭ, *et al.*, *Pis'ma Zh. Éksp. Teor. Fiz.* **41**, 191 (1985) [*JETP Lett.* **41**, 231 (1985)].
7. S. I. Braginskii, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1963), Vol. 1.
8. F. Benford and D. L. Book, in *Advances in Plasma Physics*, Ed. by A. Simon and W. B. Thompson (Wiley, New York, 1969, 1971; Mir, Moscow, 1974), Vols. 3, 4.
9. D. A. Frank-Kamenetskiĭ, *Course on Plasma Physics* (Atomizdat, Moscow, 1964).
10. G. Herziger, in *Proceedings of the 16th International Conference on Phenomena in Ionized Gases, Dusseldorf, 1983*, p. 259.
11. R. K. Rout and A. Shyam, *Plasma Phys. Controlled Fusion* **31**, 873 (1989).
12. A. Sestere, R. Robouch, and S. Podda, *Plasma Phys.* **22**, 1039 (1980).

Translated by N. Larionova

GAS DISCHARGES, PLASMA

Unsteady Motions with a Uniform Deformation in Magnetohydrodynamics

N. D. Naumov

Received March 23, 2001

Abstract—The conditions for the existence of unsteady motions with a uniform deformation are analyzed for straight and toroidal plasma columns with elliptical cross sections. The time dependence of the semiaxes of the elliptical cross sections of these plasma configurations are determined. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

There is keen interest in the development of methods for constructing unsteady solutions to the MHD equations. A very efficient instrument for this purpose is a self-similar approach [1], which makes it possible to reduce the problem of solving a set of partial differential equations to that of integrating a set of ordinary differential equations. Self-similar solutions to the MHD equations have been obtained earlier [2–6] for unstable plasma motions that belong to the class of such motions of continuous media whose velocities are proportional to the distance from the center of symmetry.

However, in all of the cited papers, studies were made of one-dimensional plasma motions. Here, a self-similar approach is applied to solve the MHD equations for two-dimensional motions with a uniform deformation. The time-dependent problem for a straight plasma column with an elliptical cross section is solved exactly, and an analogous problem for a thin toroidal plasma column with an elliptical cross section is solved approximately. The solutions obtained describe how the transverse dimensions of the plasma column change with time as both the external magnetic field and the electric current flowing through the plasma change.

BASIC EQUATIONS

In an ideal conducting fluid model, macroscopic plasma motions are described by the MHD equations [7]

$$\nabla \cdot \mathbf{B} = 0, \quad \frac{\partial \mathbf{B}}{\partial t} = \nabla \times [\mathbf{V} \times \mathbf{B}], \quad (1)$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{V} = 0, \quad (2)$$

$$\frac{d\mathbf{V}}{dt} \equiv \frac{\partial \mathbf{V}}{\partial t} + (\mathbf{V}\nabla)\mathbf{V} = -\frac{1}{\rho}\nabla p - \frac{1}{4\pi\rho}\mathbf{B} \times [\nabla \times \mathbf{B}], \quad (3)$$

where ρ , \mathbf{V} , and p are the density, velocity, and pressure of the plasma, respectively. Let S be a quantity that

remains constant during plasma motion. We multiply Eq. (1) by ∇S and transform the resulting equation to

$$\begin{aligned} \nabla S \frac{\partial \mathbf{B}}{\partial t} &= \nabla S \nabla \times [\mathbf{V} \times \mathbf{B}] = -\nabla \cdot [\nabla S \times [\mathbf{V} \times \mathbf{B}]] \\ &= \mathbf{B} \cdot \nabla (\mathbf{V} \cdot \nabla S) - \mathbf{V} \cdot \nabla (\mathbf{B} \cdot \nabla S) - (\mathbf{B} \cdot \nabla S) \nabla \cdot \mathbf{V}. \end{aligned}$$

Taking into account the relationships

$$\mathbf{V} \cdot \nabla S = -\frac{\partial S}{\partial t}, \quad \nabla \cdot \mathbf{V} = -\frac{1}{\rho} \frac{d\rho}{dt},$$

we can arrive at the following consequence of the condition for the magnetic field lines to be frozen in the plasma:

$$\frac{d}{dt} \left(\frac{\mathbf{B} \cdot \nabla S}{\rho} \right) = 0, \quad (4)$$

which implies that the quantity $\mathbf{B} \cdot \nabla S / \rho$ also remains unchanged during the plasma motion.

This result is very useful in analyzing the conditions under which the plasma experiences motions with a uniform deformation. In particular, it is obvious that, for an axisymmetric plasma column, $S = \varphi$ is a conserved quantity. Since motions with a uniform deformation are such that $\rho = \rho_0 a_0^2 / a^2$ (where a is the radius of the column), we have

$$\frac{\mathbf{B} \cdot \nabla S}{\rho} = \frac{a^2 B_\varphi}{r \rho_0 a_0}. \quad (5)$$

One can readily verify that the quantity r/a is conserved, because, for the class of motions in question, the plasma velocity is equal to $V_r = \dot{a}r/a$. As a result, from Eqs. (4) and (5), we can see that the quantity aB_φ is also a constant of motion. This condition is satisfied for $B_\varphi = 2Ir/ca^2$. Hence, we can conclude that a cylindrical plasma column in which the current density is uniformly distributed over the cross section can experience one-dimensional motions with a uniform deformation.

A STRAIGHT PLASMA COLUMN WITH AN ELLIPTICAL CROSS SECTION

A steady solution to Eqs. (1)–(3) for a plasma cylinder with an elliptical cross section over which the current density j is distributed uniformly was obtained by Gajewski [8]:

$$\mathbf{B} = \frac{4\pi j}{c(\lambda^2 + 1)}(-y, \lambda^2 x, 0), \quad (6)$$

$$p = \frac{1}{2}\rho Q \left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}\right), \quad (7)$$

where $\lambda = b/a$ is the ratio of the semiaxes b and a of the elliptical cross section and Q is a constant quantity. In this solution, the magnetic field is a superposition of the external-quadrupole confining field and the field of the plasma column [9]:

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1, \quad \mathbf{B}_0 = k(y, x, 0), \quad (8)$$

$$\mathbf{B}_1 = \frac{4\pi j}{c(\lambda + 1)}(-y, \lambda x, 0). \quad (9)$$

Since expressions (6) and (9) give a different representation of the external magnetic field,

$$\mathbf{B}_0 = \frac{4\pi j\lambda(\lambda - 1)}{c(\lambda + 1)(\lambda^2 + 1)}(y, x, 0), \quad (10)$$

we can compare expressions (8) and (10) to obtain the relationship

$$4\pi j\lambda(\lambda - 1) = ck(\lambda + 1)(\lambda^2 + 1). \quad (11)$$

If we treat relationship (11) as an equation for λ , we find that it has positive roots under the condition $j \geq j_{cr} = ck\sqrt{22 + 10\sqrt{5}}/4\pi$. Consequently, the plasma column can be in a steady state only when the plasma current density is no less than the critical density j_{cr} , which is determined by the gradient of the external magnetic field. Otherwise, the external field would disrupt the column.

Substituting expressions (6) and (7) into Eq. (3) yields the relationship

$$a^2 + b^2 = 2h^2, \quad h^2 = \frac{2I^2}{\pi\rho c^2 Q}, \quad (12)$$

where I is the electric current flowing in the plasma column.

Using relationships (11) and (12), we can determine the semiaxes of the elliptical cross section of the plasma column:

$$a = h\frac{1 - v}{\sqrt{1 + v^2}}, \quad b = a\frac{1 + v}{1 - v}, \quad (13)$$

where we introduced the notation $v = kI/\pi\rho cQ$ and, for definiteness, assumed that $\lambda > 1$.

SELF-SIMILAR SOLUTION

In the case of motion with a uniform deformation, the density and velocity of the plasma of a straight column with an elliptical cross section have the form

$$\rho = \rho_0 \frac{a_0 b_0}{ab} H(1 - \xi^2 - \eta^2), \quad \mathbf{V} = (a\dot{\xi}, b\dot{\eta}, 0),$$

where $H(x)$ is the Heaviside step function and $\xi = x/a$ and $\eta = y/b$ are self-similar variables.

For the class of motions under consideration, these variables are conserved, $d\xi/dt = d\eta/dt = 0$. In relationship (4), we successively set $S_1 = \eta$ and $S_2 = \xi$ in order to arrive at the following conditions for the current density to be uniform:

$$a^2 \left[k + \frac{4I}{ac(a+b)} \right] = C_1, \quad (14)$$

$$b^2 \left[\frac{4I}{bc(a+b)} - k \right] = C_2,$$

where C_i are constants to be determined from the initial conditions.

Conditions (14) imply that, in the general case, the plasma can experience motions with a uniform deformation when the gradient of the external magnetic field and the current in the column,

$$k = \frac{1}{a+b} \left(\frac{C_1}{a} = \frac{C_2}{b} \right), \quad I = \frac{c}{4} \left(C_1 \frac{b}{a} + C_2 \frac{a}{b} \right) \quad (15)$$

are both time-dependent.

In order to determine these time dependences explicitly, we must integrate the following set of ordinary differential equations, which follow from the Euler equation (3):

$$a\ddot{a} = F, \quad b\ddot{b} = G, \quad (16)$$

where the current I is given by expression (15) and the rest of the notation is $F = Q - \mu IC_1$ and $G = Q - \mu IC_2$, with $\mu = 1/\pi c a_0 b_0 \rho_0$. An analytic solution to Eqs. (16) can be constructed when F and G are time-independent, i.e., when the current flowing in the plasma column is constant. From expressions (15), we can see that this is the case if the ratio $\lambda = b/a$ of the semiaxes of the cross section of the plasma column remains unchanged during the plasma motion.

MOTIONS WITH A UNIFORM DEFORMATION

We assume that, at $t = 0$, the plasma column is in a steady state, i.e., the semiaxes a_s and b_s of the cross section of the column are given by expressions (13) with the replacement $k = k_0$ and $\rho = \rho_0$. For $F = G = 0$, we arrive at the simplest time dependence of the semiaxes:

$$a = a_s(1 + wt), \quad b = b_s(1 + wt), \quad (17)$$

where w is a constant.

Solution (17) describes the self-similar plasma motion when the gradient of the external magnetic field changes according to the law $k = k_0/(1 + wt)^2$. This time dependence stems from the fact that the condition $C_1 = C_2$ leads to relationship (11). Accordingly, for $w = 0$, solution (17) passes over to the above steady solution to Eqs. (1)–(3).

One can readily see that, under the condition $F = \lambda^2 G$, the nonlinear law of the evolution of the transverse dimensions of the plasma column can be evaluated analytically. Under this condition, the ratio λ is time-independent when $a = a_0 f$ and $b = b_0 f$, where the function f satisfies the following equation, which is a consequence of Eq. (16) and conditions (14):

$$f \dot{f} = \Gamma. \tag{18}$$

Here, we have used the notation

$$\Gamma = \frac{1}{2} \left[\frac{Q}{\lambda a_0 b_0} (1 + \lambda^2) - \kappa \right], \quad \kappa = \frac{1}{\pi \rho_0} \left(\frac{2I}{a_0 b_0 c} \right)^2.$$

For positive values of Γ , Eq. (18) describes a self-similar expansion of the plasma column; for negative values, it describes the compression of the column, which may be preceded by a preexpansion, provided that $\dot{f}_0 = u > 0$:

$$t = \tau A \sqrt{\pi} [\operatorname{erf}(u\tau) - \operatorname{erf}(\sqrt{u^2 \tau^2 - \ln f})].$$

This result can be obtained from Eq. (18) by setting $\Gamma = -1/2\tau^2$. By the time $t = t_k = \tau \sqrt{\pi} A \operatorname{erf}(u\tau)$, the plasma column expands to the largest transverse dimensions characterized by the quantity $f_k = A = \exp(u^2 \tau^2)$. At later times ($t \geq t_k$), the plasma column is compressed according to the law

$$t = t_k + \tau A \sqrt{\pi} \operatorname{erf}(\sqrt{u^2 \tau^2 - \ln f}).$$

When $\dot{f}_0 = -u \leq 0$, the plasma column immediately begins to be compressed:

$$t = \tau A \sqrt{\pi} [\operatorname{erf}(\sqrt{u^2 \tau^2 - \ln f}) - \operatorname{erf}(u\tau)].$$

In these formulas, $\operatorname{erf}(x)$ is the probability integral [10]. The solution obtained corresponds to the following time dependence of the gradient of the external magnetic field:

$$k = \frac{\sqrt{\pi \kappa \rho_0}}{f^2} \left[\left(1 - \frac{1}{\kappa \tau^2} \right) \frac{\lambda^2 - 1}{\lambda^2 + 1} - \frac{\lambda - 1}{\lambda + 1} \right].$$

For arbitrary initial conditions, the set of differential equations (16) can be solved numerically. The characteristic feature of these equations is that its solution is unstable (even in the linear stage of the motion). We set

$a = a_s(1 + \alpha)$ and $b = b_s(1 + \beta)$ and retain terms up to first order in α and β in expressions (15) to find

$$k = k_0 \left[1 - \alpha - \beta - \frac{\gamma}{2\nu} (1 - \nu^2) \right], \tag{19}$$

$$I = I_0 \left(1 - \frac{2\nu\gamma}{1 + \nu^2} \right),$$

where $\gamma = \alpha - \beta$ and $\nu = k_0 I_0 / \pi \rho_0 c Q$.

After linearizing Eqs. (16), we obtain the following equation for the difference between the relative changes in the semiaxes:

$$\dot{\gamma} = \frac{4\gamma k_0^2}{\pi c \rho_0 (1 - \nu^2)^2}.$$

This result shows that, in the model adopted here, the plasma column is unstable against motions with a uniform deformation.

A TOROIDAL PLASMA COLUMN WITH AN ELLIPTICAL CROSS SECTION

For a spatially bounded, axisymmetric plasma configuration, steady solutions to Eqs. (1)–(3) can be constructed by using the Shafranov equation for the magnetic surfaces [11]. A specific example of a steady solution is the toroidal configuration [12]

$$\Psi = \frac{1}{2} \Psi_0 \left[r^2 z^2 + \frac{\Lambda - 1}{4} (r^2 - R^2)^2 \right],$$

where R is the radius of the magnetic axis and Ψ_0 and Λ are constants such that

$$B_r = -\frac{1}{2\pi r} \frac{\partial \Psi}{\partial z}, \quad B_\phi = 0, \tag{20}$$

$$B_z = \frac{1}{2\pi r} \frac{\partial \Psi}{\partial r}, \quad 16\pi^3 \frac{dp}{d\Psi} = -\Lambda \Psi_0.$$

For a thin plasma column, this solution coincides (to within first-order terms) with the above solution for a straight column with an elliptical cross section. In fact, near the magnetic axis, we have

$$\Psi \approx \frac{1}{2} \Psi_0 R^2 [z^2 + (\Lambda - 1)q^2],$$

where $q = r - R$ (in this case, the small parameter is the ratio of the transverse dimension of the column to the radius of the magnetic axis).

We set $\Psi_0 = 8\pi^2 j / c R \Lambda$ and $\Lambda = \lambda^2 + 1$. Then, within the adopted accuracy, we obtain from expressions (20)

$$B_r = -\frac{4\pi}{c} \frac{jz}{\lambda^2 + 1}, \quad B_z = \frac{4\pi}{c} \frac{j\lambda^2 q}{\lambda^2 + 1}, \tag{21}$$

$$p = \frac{2I^2}{\pi c^2(a^2 + b^2)} \left(1 - \frac{q^2}{a^2} - \frac{z^2}{b^2} \right). \quad (22)$$

Taking into account relationship (12), we can see that expressions (21) and (22) are equivalent to expressions (6) and (7).

Note that the magnetic field (21) corresponds to a uniform distribution of the density of the current flowing in the plasma column in the direction opposite to that of the vector \mathbf{e}_ϕ . It is easy to find the components of the external magnetic field:

$$B_{0r} = kz \frac{R}{r}, \quad B_{0z} = kR \ln \frac{r}{R}.$$

To within first-order terms, these expressions coincide with expressions (21) for the confining magnetic field.

SELF-SIMILAR APPROXIMATION

The self-similar approach may be applied to construct an approximate unsteady solution to Eqs. (1)–(3) for a toroidal plasma column with an elliptical cross section. In my earlier paper [13], an analogous procedure was used to describe the dynamics of an electron ring.

In the absence of the longitudinal motion of the plasma and for $B_\phi = 0$, Eqs. (2) and (3) for an axisymmetric plasma distribution take the form

$$\frac{\partial \rho}{\partial t} + \frac{1}{r} \frac{\partial r \partial \rho V_r}{\partial r} + \frac{\partial \rho V_z}{\partial z} = 0, \quad (23)$$

$$\frac{dV_r}{dt} = \frac{1}{4\pi\rho} \left(B_z \frac{\partial B_r}{\partial z} - \frac{1}{2} \frac{\partial B_z^2}{\partial r} \right) - \frac{1}{\rho} \frac{\partial p}{\partial r}, \quad (24)$$

$$\frac{dV_z}{dt} = \frac{1}{4\pi\rho} \left(B_r \frac{\partial B_z}{\partial r} - \frac{1}{2} \frac{\partial B_r^2}{\partial z} \right) - \frac{1}{\rho} \frac{\partial p}{\partial z}. \quad (25)$$

For the class of motions under analysis, the density and velocity of the plasma of a toroidal column have the form

$$\rho = \rho_1 \frac{R}{r} H(1 - \xi^2 - \eta^2), \quad V_r = a\xi, \quad V_z = b\eta, \quad (26)$$

where $\rho_1 = \rho_0 a_0 b_0 / ab$ and $\xi = q/a$ and $\eta = z/b$ are self-similar variables.

One can readily see that, to first-order accuracy, the results obtained above for a straight plasma column with an elliptical cross section also apply to a thin toroidal column with a uniform current density distribution. However, in the model of an unstable plasma column, these results are applicable only to the initial stage of the motion.

In order to construct a model of a stable toroidal column, we choose the following expression for the

plasma pressure:

$$p = \frac{1}{2} \rho_1 Q \left(1 - q^2 \frac{a_s^2}{a^4} - z^2 \frac{b_s^2}{b^4} \right). \quad (27)$$

In the steady state, this expression coincides with expression (22). Note that, in the familiar Kapchinskii–Vladimirskii model, the pressure of a charged particle beam is described by a similar expression.

We substitute expressions (26) and (27) into Eqs. (24) and (25) and take into account conditions (14). Thus, we arrive at the following set of equations describing the evolution of the semiaxes of the elliptical cross section of the toroidal plasma column:

$$\ddot{a} = Q \frac{a_s^2}{a^3} - \mu I \frac{C_1}{a}, \quad \ddot{b} = Q \frac{b_s^2}{b^3} - \mu I \frac{C_2}{b}, \quad (28)$$

in which, as before, the current I is given by expression (15).

Analytical results can be obtained for small oscillations of the column about the steady state. Linearizing Eqs. (28) and using expressions (19), we find

$$\ddot{\alpha} + \frac{\omega^2}{(1-v)^2} [(1-v+v^2)\alpha + v\beta] = 0, \quad (29)$$

$$\ddot{\beta} + \frac{\omega^2}{(1+v)^2} [(1+v+v^2)\beta - v\alpha] = 0, \quad (30)$$

where we introduced the notation $\omega^2 = \pi\rho_0(cQ/I_0)^2$.

Equations (29) and (30) can be solved in a standard way. First, we calculate the eigenfrequencies $\omega_1 = \omega$ and $\omega_2 = \omega(1+v^2)/(1-v^2) \equiv \Omega$. Then, we derive the following expressions for the relative changes in the semiaxes:

$$\alpha = A_1(1+v)^2 \sin \Omega t - A_2 \sin \omega t,$$

$$\beta = A_2 \sin \omega t - A_1(1-v)^2 \sin \Omega t.$$

Here, the constant quantities A_i have the form

$$A_1 = \frac{1}{4v\Omega} (\dot{a}_0 + \dot{\beta}_0),$$

$$A_2 = \frac{1}{4v\Omega} [\dot{a}_0(1-v)^2 + \dot{\beta}_0(1+v)^2].$$

According to expressions (19), the motion with a uniform deformation of interest to us is governed by the time evolution of both the external magnetic field gradient and the current flowing in the plasma column:

$$k = k_0 \left[1 - A_1 \left(4v + \frac{1}{v} - v^3 \right) \right]$$

$$\times \sin \Omega t - A_2 \left(v - \frac{1}{v} \right) \sin \omega t,$$

$$I = I_0 \left[1 - 4v \left(A_1 \sin \Omega t - \frac{A_2}{1 + v^2} \sin \omega t \right) \right].$$

Obviously, the results obtained in this section provide the basis for constructing a model of a straight plasma column with an elliptical cross section that is stable against motions with a uniform deformation.

REFERENCES

1. L. I. Sedov, *Similarity and Dimensional Methods in Mechanics* (Nauka, Moscow, 1981).
2. A. G. Kulikovskii, Dokl. Akad. Nauk SSSR **114**, 984 (1957) [Sov. Phys. Dokl. **2**, 269 (1958)]; Dokl. Akad. Nauk SSSR **120**, 984 (1958) [Sov. Phys. Dokl. **3**, 507 (1959)].
3. I. M. Yavorskaya, Dokl. Akad. Nauk SSSR **114**, 988 (1957) [Sov. Phys. Dokl. **2**, 273 (1957)].
4. V. P. Korobeinikov, Dokl. Akad. Nauk SSSR **121**, 613 (1958) [Sov. Phys. Dokl. **3**, 739 (1958)].
5. E. V. Ryazanov, Prikl. Mat. Mekh. **23**, 187 (1959).
6. Yu. P. Ladikov, Dokl. Akad. Nauk SSSR **137**, 303 (1961) [Sov. Phys. Dokl. **6**, 198 (1961)].
7. N. A. Krall and A. W. Trivelpiece, *Principles of Plasma Physics* (Academic, New York, 1973; Mir, Moscow, 1975).
8. R. Gajewski, Phys. Fluids **15**, 70 (1972).
9. L. E. Zakharov and V. D. Shafranov, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich and B. B. Kadomtsev (Énergoatomizdat, Moscow, 1982; Consultants Bureau, New York, 1987), Vol. 11.
10. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
11. V. D. Shafranov, Zh. Éksp. Teor. Fiz. **33**, 710 (1957) [Sov. Phys. JETP **6**, 545 (1958)].
12. V. D. Shafranov, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1966), Vol. 2.
13. N. D. Naumov, Zh. Tekh. Fiz. **67** (7), 103 (1997) [Tech. Phys. **42**, 814 (1997)].

Translated by O. Khadin

A Theory of Sharp Yield Point in Low-Dislocation Crystals

B. V. Petukhov

Shubnikov Institute of Crystallography, Russian Academy of Sciences,
Leninskiĭ pr. 59, Moscow, 117333 Russia

e-mail: petukhov@ns.crys.ras.ru

Received January 18, 2001

Abstract—The condition for yield drop existence in stress–strain curves for crystals with an initially low dislocation density is derived. A new theory that radically differs from the well-known Alexander–Haasen model is proposed to describe the stressed state of the crystal near the lower yield point. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

The plastic deformation of low-dislocation crystals, primarily of semiconductors, is characterized by kinetic curves with the nonmonotonic behavior at the transition to the developed plastic flow. The peak in the stress–strain curve (the so-called sharp yield point) may cause the brittle fracture of the material. Because of this, the prediction of such behavior has received much attention.

To describe the initial stage of plastic deformation, various models of elastic–plastic transition that are based on the kinetics of dislocation multiplication have been proposed. Most commonly used is the linear multiplication law [1, 2]

$$\frac{d\rho}{dt} = A\rho\tau_e^n \quad (1)$$

Here, ρ is the dislocation density, τ_e is the effective stress ($\tau_e = \tau - \alpha\sqrt{\rho}$), τ is the applied stress, α is the coefficient of Taylor hardening, and A is the proportionality coefficient dependent on the mechanism of dislocation multiplication. In the model [1], the exponent n in (1) coincides with that in the stress dependence of the dislocation velocity V , $V = B\tau_e^m$ ($n = m$), while in model [2], it differs by unity ($n = m + 1$). According to the experimental data for semiconductor crystals [3–8], B has the Arrhenius form, $B = V_0\exp(-E/kT)$, and $m \approx 1$ for silicon [6].

In what follows we will discuss the results of active stressing tests (at a constant strain rate, $\dot{\epsilon} = \text{const}$). The construction of stress–strain curves, i.e., the dependences of the applied stress on the strain ϵ (or on the time t , $\epsilon = \dot{\epsilon}t$), with the use of Eq. (1) does predict, under certain conditions, the nonmonotonic behavior of

$\tau(\epsilon)$ in the transition region between the elastic and plastic branches of the stress–strain curve, which is in qualitative agreement with experimental data. A representative curve is shown in Fig. 1. Once the stress peak at the end of the elastic region (the upper yield point τ_u) has been passed and a sufficiently high density of mobile dislocations has been generated, intense plastic flow starts and the stress decreases to the lower yield point τ_l . Then, the stress slowly grows due to hardening by the fields $\tau_l = \alpha\sqrt{\rho}$ of newly born dislocations. The upper and lower yield points are characteristic parameters that are convenient to measure. Their dependences on temperature T , strain rate $\dot{\epsilon}$, initial dislocation density ρ_0 , etc., are of great interest [2, 5, 9]. Theoretical analysis of these dependences can be useful for the identification of plastic flow mechanisms.

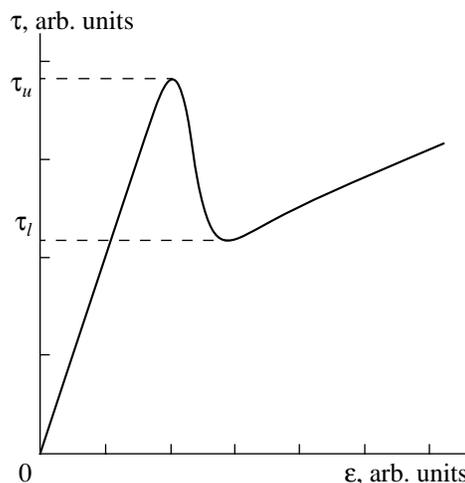


Fig. 1. Typical stress–strain curve.

Recently, promising techniques for obtaining the stress–strain curves by simulating the dynamics of dislocation ensembles on the mesoscopic level have appeared [10]. Nevertheless, the approach based on phenomenological equation (1) is still of interest because of its clearness. Since comparatively simple models, such as Eq. (1), provide the adequate qualitative description of many (but by no means all) features of plastic deformation in real materials, they have been widely used as the basis for interpreting experimental data for as long as forty years. However, the complete analysis of the stress–strain curves by varying the whole set of the parameters near the yield drop is still lacking within these models. The calculations available [1, 11] are of great importance, but they are basically unable to comprehensively solve the problem. In this study, we bridge this gap with new analytical solutions for the type of models mentioned above. Among the early approaches, the classical Alexander–Haasen [2] analysis of the upper yield point with neglect of the effect of hardening and their model for the lower yield point should be mentioned. As will be demonstrated below, the result of Alexander and Haasen for the lower yield point is valid only in a narrow range of low hardening. This, in particular, makes the determination of conditions for yield drop existence impossible. We will find these conditions in our study.

The emphasis will be on the simplest case ($n = m = 1$) of the Johnston–Gilman dislocation multiplication model [1], which, according to [5], fits the experimental data for silicon best of all. The technique is then generalized to models with arbitrary n and m , while the qualitative conclusions following from this study remain valid in an even more general case.

SCALE ANALYSIS OF THE MODEL

We will proceed from Eq. (1), which describes the variation of the dislocation density ρ , and from the expression for the total strain rate $\dot{\epsilon}$, which is the sum of the elastic and plastic contributions:

$$\frac{1}{S} \frac{d\tau}{dt} = \dot{\epsilon} - \rho b V. \quad (2)$$

Here, S is the combined elastic modulus of a sample and a testing machine and b is the Burgers vector of dislocations. We assume that the coefficient of dislocation multiplication w is related to the coefficient A in Eq. (1) as $A = wB$. For the sake of simplicity, let us turn to dimensionless variables and represent the dependences on most parameters as scale factors. To do this, we will express the stress in units of τ_1 ; dislocation density, in units of ρ_1 ; time, in units of t_1 ; and hardening, in units of a :

$$\tau_1 = \left(\frac{\dot{\epsilon} S}{w B} \right)^{1/(n+1)},$$

$$\rho_1 = \frac{1}{b} \left(\frac{w}{B} \right)^{m/(n+1)} \left(\frac{\dot{\epsilon}}{B} \right)^{(n-m+1)/(m+1)},$$

$$t_1 = (w B)^{-1/(n+1)} (\dot{\epsilon} S)^{-n/(n+1)},$$

$$a = \frac{\alpha}{\sqrt{b}} \left(\frac{B}{\dot{\epsilon}} \right)^{(m+1-n)/(2(n+1))} \left(\frac{w}{S} \right)^{(m+2)/(2(n+1))}.$$

To simplify the formulas, we will use the previous notation for the dimensionless variables, since only they appear in the subsequent discussion. The equations will then take the following form:

$$\dot{\rho} = \rho (\tau - a \sqrt{\rho})^n, \quad (3)$$

$$\dot{\tau} = 1 - \rho (\tau - a \sqrt{\rho})^m. \quad (4)$$

The scale transformation performed provides in simple terms the first approximation of the dependence of the plastic flow characteristics on most parameters, since their number has been reduced to two: a and the dimensionless initial dislocation density ρ_0 . Moreover, as will be shown below, many of the parameters, for instance, the upper yield point, depend on the parameters a and ρ_0 only slightly (logarithmically). Thus, the very separation of the scale factors in the explicit form explains much of the deformation kinetics, for example, the similar temperature dependences of τ_u and τ_l , which are defined by the Arrhenius factor B involved in τ_1 , and also the dependence of τ_u and τ_l on the strain rate $\dot{\epsilon}$:

$$\tau_u \sim \tau_l \sim \dot{\epsilon}^{1/(n+1)} \exp\left(-\frac{E}{kT(n+1)}\right).$$

Now we embark on studying the deformation kinetics predicted by Eqs. (3) and (4), starting from the simplest yet illustrative case.

$m = n = 1$: LOW HARDENING

As was noted, the experimental data [3–6] indicate that for silicon (the best-studied semiconductor material), the stress dependences of the dislocations velocity and rate of their multiplication are nearly linear. Therefore, the case $m = n = 1$ deserves particular consideration. The scale factors are then given by

$$\tau_1 = \sqrt{\frac{S \dot{\epsilon}}{w B}}, \quad t_1 = \frac{1}{\sqrt{S w \dot{\epsilon} B}}, \quad \rho_1 = \frac{1}{b} \sqrt{\frac{w \dot{\epsilon}}{S B}},$$

and

$$a = \frac{\alpha}{\sqrt{b}} \left(\frac{B}{\dot{\epsilon}} \right)^{1/4} \left(\frac{w}{S} \right)^{3/4}.$$

In the case under study, Eqs. (3) and (4) can be reduced to one closed equation. To do this, we combine these equations to get the relationship $\dot{\tau} + \dot{\rho} = 1$,

which, after integrating with respect to time, becomes $\tau = t - \rho + \rho_0$. Substituting this expression into (3), we arrive at the equation with excluded τ :

$$\dot{\rho} = \rho(t - \rho + \rho_0 - a\sqrt{\rho}). \quad (5)$$

Let us first consider the low-hardening range. With $a = 0$, Eq. (5) takes the form

$$\dot{\rho} = \rho(t - \rho + \rho_0). \quad (6)$$

The solution of this equation is

$$\rho(t) = \frac{\rho_0 \exp\left(\frac{1}{2}t^2 + \rho_0 t\right)}{1 + \rho_0 J(t)}, \quad (7)$$

where

$$J(t) = \int_0^t \exp\left(\frac{1}{2}t'^2 + \rho_0 t'\right) dt'.$$

In initially low-dislocation crystals, $\rho_0 \ll 1$. In that event, we can disregard the term $\rho_0 t$ in the exponents in (7). Assuming that, in view of small ρ_0 , the denominator in (7) noticeably deviates from unity only at $t \gg 1$, we can approximate the integral $J(t)$ by the asymptotic expression $J(t) \approx \exp(t^2/2)(1 + 1/t^2)/t$.

Since the exponential $\exp(t^2/2)$ strongly varies with t at $t \gg 1$ (it changes by a factor of e when t is changed by $\delta t = 1/t \ll 1$), the denominator in (7) changes from nearly unity to a large value within a narrow range in the vicinity of $t = t_0$, where t_0 is defined by the condition $\rho_0 J(t_0) = 1$. As a result, $\rho(t)$ abruptly changes from

$$\rho(t) \approx \rho_0 \exp(t^2/2), \quad t < t_0 \quad (8)$$

to

$$\rho(t) \approx t(1 - 1/t^2), \quad t > t_0. \quad (9)$$

The change in the character of the solution at $t > t_0$ can qualitatively be explained by the fact that it asymptotically tends to the quasi-stationary value at large t , when the time derivative in Eq. (6) is small and $\rho \approx t$ in the principal order and, according to (9), $\rho \approx t - \dot{\rho}/\rho \approx t - 1/t$ in the next order. Nontrivial here is only the circumstance that the transition range is narrow, on the order of $1/t_0 \ll 1$, and can be considered as a kink in the kinetic curves at $t = t_0$.

The time dependence of stress $\tau(t) = t - \rho(t)$ shows a peak followed by a monotonic decrease. The vicinity of the peak up to $t = t_0$ will be called the region of sharp yield point. Beyond this range, the stress drops smoother: $\tau(t) \approx 1/t$. In the region of sharp yield point,

the stress can be approximated with the use of (8) for $\rho(t)$:

$$\tau(t) \approx t - \rho_0 \exp(t^2/2). \quad (10)$$

The position t_u of the peak is given by the equation

$$\dot{\tau} \approx 1 - t_u \rho_0 \exp(t_u^2/2) = 0. \quad (11)$$

For $\ln(1/\rho_0) \gg 1$, Eq. (11) can be solved by iterations with any required accuracy. Restricting ourselves to double iteration, we have

$$t_u \approx \left[2 \ln \left(\frac{1}{\rho_0 \sqrt{2 \ln(1/\rho_0)}} \right) \right]^{1/2}. \quad (12)$$

The substitution of t_u into (10) yields the upper yield point:

$$\begin{aligned} \tau_u &= t_u - 1/\tau_u \\ &\approx \sqrt{2 \ln(1/\rho_0)} - \frac{\ln \sqrt{2 \ln(1/\rho_0)} + 1}{\sqrt{2 \ln(1/\rho_0)}}. \end{aligned} \quad (13)$$

Let us trace how the results change at the presence of hardening. As long as the hardening coefficient is small, the change is noticeable only at comparatively large t ($t > t_0$) in the region of quasi-stationary behavior. In Eq. (5) rewritten as

$$t = \rho + a\sqrt{\rho} + \dot{\rho}/\rho, \quad (14)$$

the two last terms are small and can be substituted by the solution of the zeroth approximation $\rho \approx t$. This yields $\rho \approx t - a\sqrt{t} - 1/t$ and

$$\tau = t - \rho \approx a\sqrt{t} + 1/t. \quad (15)$$

Unlike the case of zero hardening, dependence (15) has a minimum and grows as $t \rightarrow \infty$. The minimum is located at $t = t_l = (2/a)^{2/3}$ and corresponds to $\tau_l = 3(a/2)^{2/3}$. This is the solution for the lower yield point similar to that obtained by Alexander and Haasen [2]. Its specific feature is the independence of τ_l from the initial dislocation density ρ_0 . The properties of the lower yield point have been thoroughly discussed elsewhere (see, e.g., [5, 11]). However, the applicability range of the Alexander–Haasen solution has not been adequately studied.

From the derivation of the solution, it is evident that it is valid only if the minimum falls into the quasi-stationary region $t_l > t_0$, i.e., at $a < 2/t_0^{3/2} \ll 1$. With higher hardening coefficients, the Alexander–Haasen approach is inapplicable and the description of the lower yield point should be revised.

The bound of the quasi-stationary region also depends on a , while slightly, and will be denoted as t_a instead of t_0 . We will proceed from the circumstance that only the quasi-stationary region of the stress–strain curve, $t > t_a$, is severely distorted by hardening. The

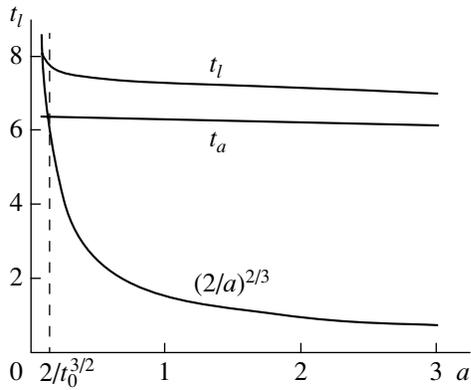


Fig. 2. Location of the lower yield point vs. hardening coefficient for an initial dislocation density $\rho_0 = 10^{-8}$. The Alexander–Haasen solution $t_l = (2/a)^{2/3}$ and the yield drop limit t_a are also shown.

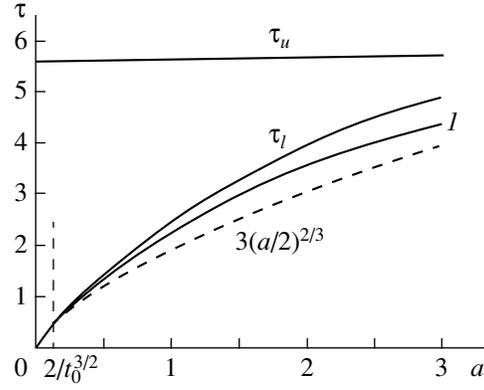


Fig. 3. Upper and the lower yield points vs. hardening coefficient, $\rho_0 = 10^{-8}$. The Alexander–Haasen solution $\tau_l = 3(2/a)^{2/3}$ and approximation (18) are also shown.

effect of hardening on the stress–strain curve, and in particular on the shape of the yield drop, at $t < t_a$ is substantially weaker and becomes essential at considerably higher a (specifically, this shows up in the slight a dependence of τ_u ; see below). Therefore, $\tau(t)$ monotonically decreases between t_u and t_a , whereas at $a > 2/t_0^{3/2}$ and $t > t_a$, it monotonically grows; hence, the stress–strain curve passes through a minimum in the vicinity of t_a .

Thus, the position of the minimum on the stress–strain curve is defined by the position of the boundary of the quasi-stationary deformation region. Taking into account the term with $\dot{\rho}$ in (14) in terms of the perturbation theory, one obtains the approximate time dependence of the dislocation density in the quasi-stationary mode:

$$\rho(t) \approx (\sqrt{1 + a^2/4} - a/2)^2 - \frac{1}{t + a^2/4}. \quad (16)$$

The boundary t_a can approximately be found as the point of intersection between time dependences of ρ (8) and (16):

$$\rho_0 \exp(t^2/2) \approx (\sqrt{t + a^2/4} - a/2)^2 - \frac{1}{t + a^2/4}. \quad (17)$$

In passing from the Alexander–Haasen solution $t_l = (2/a)^{2/3}$ to the solution $t_l \approx t_a$ defined by Eq. (17), the effect of the hardening coefficient a on the location of the minimum becomes much weaker. Thus, for the hardening coefficient $a > 2t_0^{-3/2}$, the lower yield point can be estimated by the substitution of the value $t = t_a$ which is minimal for the quasi-stationary region, into $\tau(t)$. Eventually,

$$\tau_l \approx t_a - (\sqrt{t_a + a^2/4} - a/2)^2 + \frac{1}{t_a + a^2/4}. \quad (18)$$

Since t_a is ρ_0 -dependent, the lower yield point [see (18)] also depends on the initial dislocation density, thereby qualitatively differing from the solution obtained by Alexander and Haasen [2]. To illustrate this, Fig. 2 shows the a dependence of t_l obtained by the numerical solution of Eq. (5). It is evident that the boundary t_a of the yield drop is the lower barrier for t_l within a wide range of a . As a result, the minimum in this range lies close to t_a , so that t_a is a satisfactory approximation for the location of the minimum.

Figure 3 depicts the a dependences of the upper and lower yield points τ_u and τ_l obtained by numerically solving Eq. (5). It is clearly seen that, beyond the narrow applicability range of the Alexander–Haasen dependence, solution (18) (curve 1) better fits the lower yield point. Note also the aforementioned weak dependence of the upper yield point τ_u on a .

$m = n = 1$: CRITERION FOR YIELD DROP EXISTENCE

As a grows, the maximum τ_u and the minimum t_l approach each other and the height $\tau_u - \tau_l$ of the yield drop decreases as shown in Fig. 3. At a critical value $a = a_c$, the maximum and minimum merge together and the stress–strain curve becomes monotonic. Thus, the criterion for yield drop existence is the inequality $a < a_c$. Let us calculate the value of a_c . It is convenient to proceed from the approximated variant of Eq. (5) where the term $a\sqrt{\rho}$ dominates: $\dot{\rho} = \rho(t - a\sqrt{\rho})$. Its solution can be written as

$$\rho = \frac{\rho_0 \exp(t^2/2)}{\left[1 + \frac{a}{2}\sqrt{\rho_0} J_a(t)\right]^2}, \quad (19)$$

where

$$J_a(t) = \int_0^t \exp(t'^2/4) dt' \approx \exp(t^2/4) \frac{2}{t} \left(1 + \frac{2}{t^2} \right).$$

This solution, in particular, indicates that, as in the case of low hardening, the time dependence abruptly changes in the vicinity of the t_a point, given by the condition $(a/2)\sqrt{\rho_0}J_a(t_a) = 1$. At $t < t_a$, $\rho(t) \approx \rho_0 \exp(t^2/2)$, as before in the case $a = 0$, and is nearly independent of the hardening. At $t > t_a$, $\rho(t) \approx (t^2 - 4)/a^2$, which agrees with (16) in the limit $1 \ll t \ll a^2$ considered. We can thus conclude that, as long as the stress-strain curve peaks within the region $t_u < t_a$, its height and position are affected by hardening only slightly. However, with an increase in a , the quasi-stationary region approaches t_u and the peak may disappear. Let us trace how this occurs.

The equation defining the extreme points of $\tau(t)$, $\dot{\tau} = 1 - \rho(t - \rho - a\sqrt{\rho}) = 0$, suggests that $\dot{\rho} = 1$ at the extrema. This equation can be rewritten as

$$t = \rho(t) + a\sqrt{\rho(t)} + 1/\rho(t). \quad (20)$$

At $a < a_c$, the curves corresponding to the right- and the left-hand sides of the equation intersect at two points representing the maximum, t_u , and the minimum, t_l , of stress-strain curves. With an increase in a , t_u and t_l approach and merge together when the t curve becomes tangent to the t dependence in the right-hand side of Eq. (20):

$$1 = \left(1 + \frac{a}{2\sqrt{\rho}} - \frac{1}{\rho^2} \right) \frac{d\rho}{dt} = 1 + \frac{a}{2\sqrt{\rho}} - \frac{1}{\rho^2}. \quad (21)$$

Here, we take into account that $\dot{\rho} = 1$ at the extreme points. Expressing ρ_c in terms of a_c from (21), $\rho_c = (2/a_c)^{2/3}$, and using Eq. (20), we obtain $t_c = 3(a_c/2)^{2/3} + (2/a_c)^{2/3}$. Substituting these relationships into (19) yields the critical condition in the form

$$\ln(1/\rho_{0c}) \approx \frac{9}{2} \left(\frac{a_c}{2} \right)^{4/3} + \frac{2}{3} \ln(a_c/54) + 3. \quad (22)$$

Figure 4 illustrates the agreement between this formula and the numerical solution. It is evident that the lower the initial dislocation density, the wider the domain of yield drop existence. In order to trace the dependences on the other parameters, one should return to the dimensional variables and write the critical condition in the form

$$\frac{\alpha}{\sqrt{b}} \left(\frac{w}{S} \right)^{3/4} \left(\frac{B}{\dot{\epsilon}} \right)^{1/4} < a_c.$$

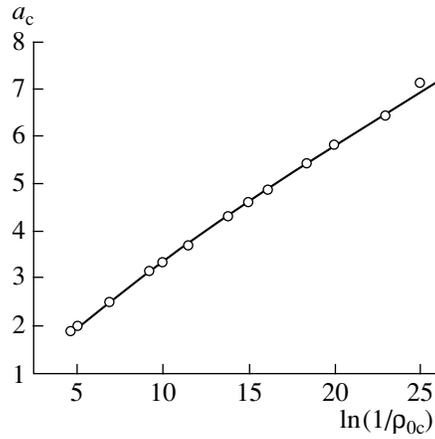


Fig. 4. Yield drop existence limit [expression (22)] for $a_c(\rho_0)$; circles show the results of numerical simulation.

Disregarding the logarithmic dependence, which is conditioned by the normalizing factor ρ_1 included in ρ_0 , one can conclude that a decrease in the temperature and an increase in the strain rate $\dot{\epsilon}$ favor the yield drop appearance.

GENERALIZATION

Qualitatively, the stress-strain curves retain their features in the more general case of arbitrary m and n . Let us briefly list them bearing in mind the complete analogy with the case of $m = n = 1$ considered in detail. To describe the vicinity of the upper yield point, it will suffice to calculate the first correction to elastic behavior. In the zeroth approximation, $\rho \approx 0$ and $\tau \approx \tau_e \approx t$. In the next approximation,

$$\rho \approx \rho_0 \exp(t^{n+1}/(n+1)), \quad (23)$$

$$\tau \approx t - \rho_0 t^{m-n} \exp(t^{n+1}/(n+1)). \quad (24)$$

The extreme points are found from the equation $\dot{\tau} = 0$; i.e.,

$$\tau = a\sqrt{\rho} + 1/\rho^{1/m}. \quad (25)$$

When the hardening is low and the stress-strain curve peaks within the region of low dislocation density, one can omit the term $a\sqrt{\rho}$ in Eq. (25), substitute t for τ , and express $\rho(t)$ by (23). As a result, Eq. (25) becomes

$$\rho_0 \exp(t_u^{n+1}/(n+1)) \approx 1/t_u^m. \quad (26)$$

The equation can be solved by iterations. To find the first approximation, we substitute $t_u = 1$ into the right-hand side and obtain $t_1 = [(n+1)\ln(1/\rho_0)]^{1/(n+1)}$. Then, substituting $t_u = t_1$ into the right-hand side of (26) yields the next approximation, $t_u \approx t_1 [1 - m(n+1)/t_1^{n+1}]^{1/(n+1)}$,

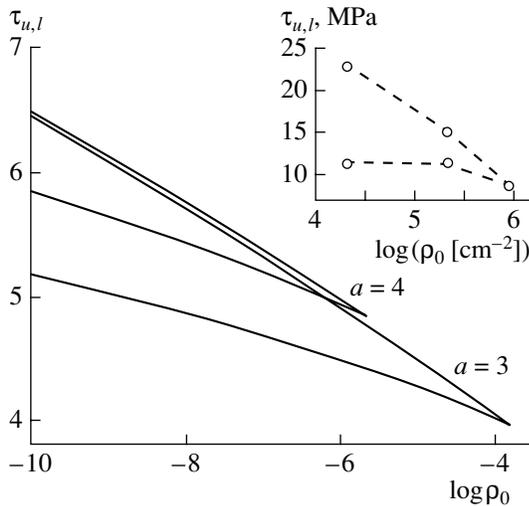


Fig. 5. Upper and lower yield points vs. initial dislocation density. Hardening coefficient $a = 3$ and 4 . The inset shows the experimental dependences for silicon [12] ($T = 800^\circ\text{C}$ and $\dot{\epsilon} = 1.2 \times 10^{-4} \text{ s}^{-1}$).

to which we restrict ourselves. The insertion of this solution into (25) specifies the upper yield point as

$$\begin{aligned} \tau_u &= t_u - t_u^{m-n} \rho(t_u) \approx t_u - 1/t_u^n \\ &\approx t_1 - [m \ln(t_1) + 1]/t_1^n. \end{aligned} \quad (27)$$

It is also not difficult to obtain the generalized criterion for yield drop existence. For this purpose, we will solve the equation for extreme points, $\tau = a\sqrt{\rho} + 1/\rho^{1/m}$, combined with the bifurcation condition

$$\dot{\tau} = 0 = \frac{a}{2\sqrt{\rho}} - \frac{1}{m\rho^{\frac{m+1}{m}}},$$

which generalizes (21).

Hence,

$$\rho_c = \left(\frac{2}{a_c m}\right)^{\frac{2m}{m+2}}, \quad \tau_c \approx t_c \approx \frac{m+2}{m} \left(\frac{a_c m}{2}\right)^{\frac{2}{m+2}}. \quad (28)$$

To yield the criterion in the basic logarithmic approximation, i.e., up to terms on the order of unity, which are small compared with the major contributions, the time dependence of ρ can iteratively be approximated by expression (23) with the substitution of quantities (28). Finally, we have

$$\begin{aligned} \ln(1/\rho_{0c}) &\approx \frac{1}{n+1} \left(\frac{m+2}{m}\right)^{n+1} \left(\frac{a_c m}{2}\right)^{\frac{2(n+1)}{m+2}} \\ &+ \frac{2m}{m+2} \ln\left(\frac{a_c m}{2}\right). \end{aligned} \quad (29)$$

In addition to the ρ_0 dependence of a_c , this formula also relates the criterion for yield drop existence to the exponents m and n , which characterize the stress sensitivity of the dislocation velocity and the kinetics of their multiplication.

CONCLUSIONS

Let us outline some qualitative conclusions common to the models of the type considered above.

(1) In crystals with an initially low dislocation density, the stress–strain curves display a sharp transition to the quasi-stationary behavior. The transition region is so narrow that one can speak of the transition “point” (t_0 or t_a). This point is as important for characterizing the deformation kinetics as the upper and lower yield points are.

(2) At some critical hardening coefficient $a = a_c$, the maximum and the minimum on the stress–strain curve merge together and disappear, rendering the curve monotonic. The lower the initial dislocation density, the larger the value of a_c . Experimentally, it may be more convenient to vary the initial dislocation density ρ_0 with the other parameters fixed. The yield drop decreases with increasing ρ_0 and disappears as the critical value ρ_{0c} is exceeded. The qualitative correlation of the behavior predicted by the model under discussion with the experimental data for silicon [12] is evident from Fig. 5.

(3) The quasi-stationary region of the stress–strain curve is independent of the initial dislocation density ρ_0 . If the lower yield point falls into this region, it is also ρ_0 -independent. This is true, however, only for the limited interval of hardening coefficient. When the upper limit of this interval is exceeded, the lower yield point approaches t_a —the point of transition to the quasi-stationary mode—and begins to depend on the initial dislocation density. Such a situation takes place near the region of yield drop existence, i.e., when the minimum of the stress–strain curve is near the yield drop. This is confirmed by the experimental data for silicon [12] (Fig. 5), which show that τ_l depends on ρ_0 near the point where τ_u and τ_l merge.

The shift of the lower yield point towards larger times and, therefore, higher dislocation densities can also be responsible for the experimentally observed fact [12] that the internal stresses exceed their effective value near the lower yield point vastly larger than predicted by the expression $\tau_l = 2\tau_c/m$, following from the Alexander–Haasen theory [2].

REFERENCES

1. W. G. Johnston and J. J. Gilman, *J. Appl. Phys.* **30**, 129 (1959).
2. H. Alexander and P. Haasen, *Solid State Phys.* **22**, 22 (1968).

3. V. I. Nikitenko, *Dislocation Dynamics* (Naukova Dumka, Kiev, 1975), pp. 7–26.
4. A. George and J. Rabier, *Rev. Phys. Appl.* **22**, 941 (1987).
5. H. Alexander, in *Dislocations in Solids*, Ed. by F. R. N. Nabarro (North-Holland, Amsterdam, 1986), Vol. 7, Chap. 35, pp. 113–234.
6. K. Sumino and M. Imai, *Philos. Mag. A* **47**, 753 (1983).
7. Yu. L. Iunin, V. I. Nikitenko, V. I. Orlov, and B. Ya. Farber, *Fiz. Tverd. Tela (Leningrad)* **33**, 1262 (1991) [*Sov. Phys. Solid State* **33**, 715 (1991)].
8. T. Suzuki, H. Koizumi, and H. O. K. Kirchner, *Philos. Mag. A* **71**, 389 (1995).
9. J. Rabier and A. George, *Rev. Phys. Appl.* **22**, 1327 (1987).
10. A. Noulin, M. Condat, and L. P. Kubin, *Acta Mater.* **47**, 2879 (1999).
11. M. Suezawa, K. Sumino, and I. Yonenaga, *Phys. Status Solidi A* **51**, 217 (1979).
12. I. Yonenaga and K. Sumino, *Phys. Status Solidi A* **50**, 685 (1978).

Translated by A. Sidorova-Biryukova

Dynamics of Photon Emission Due to Strains in Metals

K. B. Abramova, A. A. Semenov, and I. P. Shcherbakov

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

e-mail: abramova.mhd@pop.ioffe.rssi.ru

Received October 9, 2000

Abstract—Calculations of a photon emission pulse generated on the back side of a metallic sample laser-irradiated on its front side are performed. A detailed comparison with experimental data is made. © 2001 MAIK “Nauka/Interperiodica”.

In this work, we pursue studies of mechanoluminescence in metals [1–12]. In experiments [8–12], mechanical stresses were induced by a single laser shot with a beam diameter much smaller than the diameter of the sample. The irradiated spot on the surface rapidly heats up, and the temperature distribution varying in time and space arises because of the finite thermal conductivity of the sample. As a result, mechanical stresses that also vary in space and time appear. Processes taking place on the back (nonirradiated) side of the sample were considered. It has been found that luminescence from the metal surface is excited when the stresses approach the yield point but the temperature remains nearly unchanged [11, 12]. However, as follows from experiments [10–12], the dynamics of a luminescence pulse differs from that of the stresses and the rate of their change; namely, the luminescence decays when the stresses and the rate of their rise continue increasing.

According to the dislocation mechanism of photon emission, when metals deform, the rate of mechanoluminescence must depend on the initial dislocation density and the number of dislocations crossing the surface due to strain [13–15]. Both statements have been confirmed experimentally [9, 16, 17]. It has been shown that the deformation or failure of a sample having a high dislocation density is attended by intense emission and the topography of its surface considerably changes, because many dislocations cross the surface. If the dislocation density in a sample being deformed is small (all other things being equal), the emission is weak or is not detected at all and the surface relief changes insignificantly.

The aforesaid allows us to assume that the dynamics of an emission pulse correlates with the dynamics of dislocations crossing the surface. The latter can be calculated by assuming that dislocations crossing the surface change the emitting area of the sample surface.

Then, we have

$$\frac{dN}{dt} = -A\dot{\epsilon}N. \quad (1)$$

Here, N is the dislocation density, ϵ is the relative change in the back surface area due to thermal stresses, $\dot{\epsilon}$ is the rate of change of the back surface area, t is time, and A is a constant. In our early experiments [11, 16], we studied luminescence from the back side of copper samples irradiated by a laser shot. The samples were 0.5-mm-thick copper disks with a diameter of 30 mm. The 10.6- μm 1.5-ms laser shot was focused into a spot of diameter 1 mm. The minimal laser energy was selected so that luminescence was reliably detected. The energy applied to the sample was 0.6 J. At the center of the spot, the maximal stresses exceeded the yield point of copper only slightly, while at its edge, they only approached it by the end of the luminescence pulse [11]. Such conditions allow us to calculate ϵ and $\dot{\epsilon}$ in the linear approximation by invoking equations of the classical elasticity theory, which gives averaged values of strains and stresses. Integrating Eq. (1) yields

$$N = N_0 \exp(-A\epsilon), \quad (2)$$

where N_0 is the initial density of dislocations. We assume that the strains are symmetric about the laser beam center. Let us express the changes in the surface area through displacements:

$$r' = r + U_r, \quad z' = z + U_z. \quad (3)$$

Here, r and z are the positions of surface points before deformation, U_r and U_z are the positions of the displacement vector, and r' and z' are the positions of points on the surface being deformed:

$$dL' = \sqrt{(dr')^2 + (dz')^2} \approx \left(1 + \frac{\partial U_r}{\partial r}\right) dr; \quad (4)$$

$$dS' = dL'r'd\phi \approx \left(1 + \frac{\partial U_r}{\partial r}\right)(r + U_r)drd\phi$$

$$\approx \left(1 + \frac{1}{r}\frac{\partial}{\partial r}(rU_r)\right)rdrd\phi. \tag{5}$$

Since $rdrd\phi = dS$, the relative change in the surface area takes the form

$$\epsilon = \frac{dS' - dS}{dS} = \frac{1}{r}\frac{\partial}{\partial r}(rU_r). \tag{6}$$

The displacement \mathbf{U} is found from equations derived in [11]:

$$\Delta U_r - \frac{U_r}{r} + \frac{1}{1-2\mu}\frac{\partial}{\partial r}(\text{div}\mathbf{U}) = \frac{2(1+\mu)}{1-2\mu}\frac{\partial(\alpha T)}{\partial r},$$

$$\Delta U_z + \frac{1}{1-2\mu}\frac{\partial}{\partial z}(\text{div}\mathbf{U}) = \frac{2(1+\mu)}{1-2\mu}\frac{\partial(\alpha T)}{\partial z}, \tag{7}$$

where μ is the Poisson's ratio, α is the thermal expansion coefficient, r and z are cylindrical coordinates, $T(z, r, \tau)$ is the temperature distribution in the sample, $\tau = kt/c\rho a^2$ is dimensionless time, k is the thermal conductivity, t is time, c is the heat capacity, ρ is the density of the material, and a is the sample thickness.

From Eq. (7), it follows that the time and spatial temperature distributions in the sample set under irradiation are needed in order to calculate the displacements and, accordingly, the change in the back surface area. The temperature distribution within the laser spot is

$$f(\tau) = \begin{cases} \sin \frac{\pi\tau}{\tau_0} & \tau \leq \tau_0 \\ 0 & \tau > \tau_0, \end{cases} \tag{8}$$

where $\tau_0 = kt_0/c\rho a^2$ and t_0 is the shot duration. The temperature distribution in the sample is given in [1]. For $\tau < \tau_0$, it has the form

$$T(r, z, \tau) = \frac{q_0 r_0}{k} \int_0^\infty J_1\left(\lambda \frac{r_0}{a}\right) J_0\left(\lambda \frac{r}{a}\right)$$

$$\times \left\{ \frac{\lambda^2 \sin \frac{\pi\tau}{\tau_0} - \frac{\pi a^2}{\tau_0} \left(\cos \frac{\pi\tau}{\tau_0} - \exp\left(-\lambda^2 \frac{\tau}{a^2}\right) \right)}{\lambda^4 + \frac{\pi a^4}{\tau_0^2}} \right.$$

$$\left. + 2 \sum_{n=1}^\infty \frac{\cos \frac{n\pi z}{a}}{(n^2 \pi^2 + \lambda^2)^2} \right\} (n^2 \pi^2 + \lambda^2) \sin \frac{\pi\tau}{\tau_0} \tag{9}$$

$$\left. - \frac{\pi a^2}{\tau_0} \left(\cos \frac{\pi\tau}{\tau_0} - \exp\left[-(n^2 \pi^2 + \lambda^2) \frac{\tau}{a^2}\right] \right) \right\} d\lambda,$$

for $\tau > \tau_0$,

$$T(r, z, \tau) = \frac{q_0 r_0 \pi a^2}{k \tau_0} \int_0^\infty J_1\left(\lambda \frac{r_0}{a}\right) J_0\left(\lambda \frac{r}{a}\right)$$

$$\times \left\{ \frac{\exp\left(-\lambda^2 \frac{\tau}{a^2}\right) \left(\exp\left(\lambda^2 \frac{\tau_0}{a^2}\right) + 1 \right)}{\lambda^4 + \frac{\pi a^4}{\tau_0^2}} \right.$$

$$\left. + 2 \sum_0^\infty \frac{\cos\left(\frac{n\pi z}{a}\right)}{(n^2 \pi^2 + \lambda^2)^2 + \frac{\pi a^4}{\tau_0^2}} \exp\left[-(n^2 \pi^2 + \lambda^2) \frac{\tau}{a^2}\right] \right.$$

$$\left. \times \left[\exp\left[(n^2 \pi^2 + \lambda^2) \frac{\tau}{a^2}\right] + 1 \right] \right\} d\lambda. \tag{10}$$

Here, q_0 is the energy flux absorbed by the sample and r_0 is the radius of the laser spot on the surface. The boundary conditions for Eqs. (7) are

$$\sigma_{ik} n_k = 0, \tag{11}$$

where σ_{ik} is the stress tensor and n is the normal to the surface.

Set (7) with boundary conditions (11) was solved by using the thermoelastic potential of displacements and Love functions [18]. Eventually, the desired solution for U_r on the back side of the sample is obtained in the form

$$U_r = \beta \int_0^\infty J_1\left(\lambda \frac{r_0}{a}\right) J_1\left(\lambda \frac{r}{a}\right)$$

$$\times (F_1(\lambda, \tau, \tau_0) - F_2(\lambda, \tau, \tau_0)) \lambda d\lambda. \tag{12}$$

Here, $\beta = 2\alpha(1 + \mu)(q_0 r_0 a/k)$. For $\tau \leq \tau_0$,

$$F_1(\lambda, \tau, \tau_0) = \frac{\sinh \lambda}{\sinh \lambda + \lambda} \left\{ \frac{1}{\lambda^2 \left(\lambda^4 + \frac{\pi^2}{\tau_0^2} \right)} \right.$$

$$\begin{aligned} & \times \left(\lambda^2 \sin \frac{\pi \tau}{\tau_0} - \frac{\pi}{\tau_0} \left(\cos \frac{\pi \tau}{\tau_0} - \exp(-\lambda^2 \tau) \right) \right) \\ & + 2 \sum_{k=1}^{\infty} \frac{1}{(4k^2 \pi^2 + \lambda^2) \left((4k^2 \pi^2 + \lambda^2)^2 + \frac{\pi^2}{\tau_0^2} \right)} \\ & \times \left((4k^2 \pi^2 + \lambda^2) \sin \frac{\pi \tau}{\tau_0} - \frac{\pi}{\tau_0} \left(\cos \frac{\pi \tau}{\tau_0} \right. \right. \\ & \left. \left. - \exp(-(4k^2 \pi^2 + \lambda^2) \tau) \right) \right) \end{aligned}$$

$$F_2(\lambda, \tau, \tau_0) = \frac{2 \sinh \lambda}{\sinh \lambda - \lambda}$$

$$\begin{aligned} & + \sum_{k=1}^{\infty} \frac{1}{((2k+1)^2 \pi^2 + \lambda^2) \left(((2k+2)^2 \pi^2 + \lambda^2)^2 + \frac{\pi^2}{\tau_0^2} \right)} \\ & \times \left(((2k+1)^2 \pi^2 + \lambda^2) \sin \frac{\pi \tau}{\tau_0} \right. \\ & \left. - \frac{\pi}{\tau_0} \left(\cos \frac{\pi \tau}{\tau_0} - \exp(-((2k+1)^2 \pi^2 + \lambda^2) \tau) \right) \right), \end{aligned}$$

for $\tau > \tau_0$,

$$\begin{aligned} F_1(\lambda, \tau, \tau_0) &= \frac{\sinh \lambda}{\sinh \lambda + \lambda \tau_0} \left\{ \frac{1}{\lambda^2 \left(\lambda^4 + \frac{\pi^2}{\tau_0^2} \right)} \right. \\ & \times (\exp(-\lambda^2 \tau) + \exp(-\lambda^2 (\tau - \tau_0))) \\ & + 2 \sum_{k=1}^{\infty} \frac{1}{(4k^2 \pi^2 + \lambda^2) \left((4k^2 \pi^2 + \lambda^2)^2 + \frac{\pi^2}{\tau_0^2} \right)} \\ & \times (\exp(-(4k^2 \pi^2 + \lambda^2) \tau) \\ & \left. + \exp(-(4k^2 \pi^2 + \lambda^2) (\tau - \tau_0))) \right\}, \end{aligned}$$

$$F_2(\lambda, \tau, \tau_0) = \frac{2 \sinh \lambda}{\sinh \lambda - \lambda \tau_0}$$

$$\begin{aligned} & + \sum_{k=1}^{\infty} \frac{1}{((2k+1)^2 \pi^2 + \lambda^2) \left(((2k+1)^2 \pi^2 + \lambda^2)^2 + \frac{\pi^2}{\tau_0^2} \right)} \\ & \times (\exp(-((2k+1)^2 \pi^2 + \lambda^2) \tau) \\ & + \exp(-((2k+1)^2 \pi^2 + \lambda^2) (\tau - \tau_0))). \end{aligned}$$

Thus, substituting (12) into (6), we obtain

$$\begin{aligned} \varepsilon(r, \tau) &= \beta \int_0^{\infty} J_1 \left(\lambda \frac{r_0}{a} \right) J_0 \left(\lambda \frac{r}{a} \right) (F_1(\lambda, \tau, \tau_0) \\ & - F_2(\lambda, \tau, \tau_0) \lambda^2 d\lambda). \end{aligned} \tag{13}$$

To continue calculations, we express $\partial \varepsilon / \partial t$, or $\dot{\varepsilon}$, as

$$\begin{aligned} \dot{\varepsilon}(r, \tau) &= \beta_1 \int_0^{\infty} J_1 \left(\lambda \frac{r_0}{a} \right) J_0 \left(\lambda \frac{r}{a} \right) \left(\frac{\partial}{\partial \tau} F_1(\lambda, \tau, \tau_0) \right. \\ & \left. - \frac{\partial}{\partial \tau} F_2(\lambda, \tau, \tau_0) \right) \lambda^2 d\lambda, \end{aligned} \tag{14}$$

where $\beta_1 = \beta k / c \rho a^2$. The derivatives of $F_1(\lambda, \tau, \tau_0)$ and $F_2(\lambda, \tau, \tau_0)$ with respect to τ_0 are not given here, since they are derivatives of the elementary functions. Figures 1a and 1b show the experimental curves $I_{\text{las}}(t)$ and $I_{\text{lum}}(t)$, as well as the dependences $\varepsilon(t)$ and $\dot{\varepsilon}(t)$ calculated from (13) and (14), respectively, and averaged over a circle of radius $r = 1$ mm. The radius on the back side is taken to be equal to the radius of the laser spot, because the material structure changes within this radius (as follows from the examination of the metallographic sections [9]). The constant A in expressions (1) and (2) is determined in the following way. According to the dislocation mechanism, mechanoluminescence on the metal surface is initiated when mobile dislocations emerge on the surface. Hence, the maximal luminescence intensity will be observed at the instant the greatest number of the dislocations appear on the surface. At this time instant, the second-order time derivative of expression (2) must vanish. Mathematically, this condition is written as

$$A = \frac{\ddot{\varepsilon}}{\dot{\varepsilon}^2}, \tag{15}$$

where $\ddot{\varepsilon}$ and $\dot{\varepsilon}$ are taken at the instant of maximal luminescence intensity and averaged over a circle of radius $r = 1$ mm.

In our experiments, the mechanoluminescence intensity peaks at $t = 0.7$ ms. For the sample under

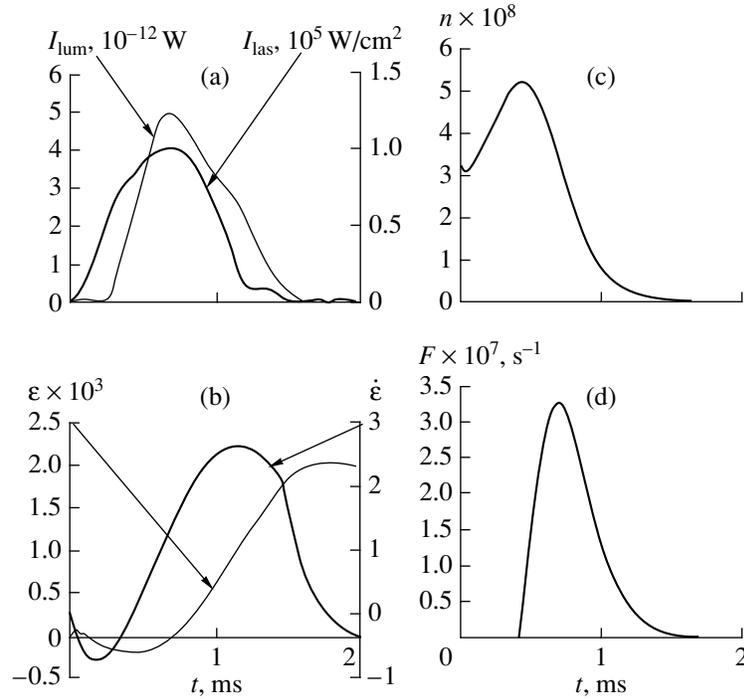


Fig. 1. (a) Waveforms of the irradiating pulse (the numbers indicate the power absorbed by the sample) and the pulse of mechanoluminescence initiated on the back side of the sample, (b) calculated relative stress-induced change in the back surface area and the rate of its change vs. time, (c) calculated time variation of the number of dislocations emerging on the surface within a circle of radius $r = 1$ mm, and (d) calculated photon emission pulse.

study,

$$A = 5.00 \times 10^6. \quad (16)$$

Figure 1c shows the time variation of the number n of dislocations in the plastic zone within a circle of radius $r = 1$ mm. The plot is obtained from (1) in view of (14) and (16). It follows that, as long as $\dot{\epsilon} < 0$, the number of the dislocations grows, while at $\dot{\epsilon} > 0$, it decreases (the dislocations emerge on the surface). We assume that for $\dot{\epsilon} > 0$, the number of newly generated defects is negligible compared with the number of dislocations crossing the surface; then, the continuity condition

$$\frac{\partial N}{\partial t} + \text{div} \mathbf{j} = 0, \quad (17)$$

(\mathbf{j} is the flux density of dislocations emerging on the surface) must be met.

The second term in Eq. (17) can be estimated as

$$\text{div} \mathbf{j} \cong \frac{j_n}{\delta_t}, \quad (18)$$

where δ_t is the plastic zone depth.

Thus, the flux density of dislocations emerging on the surface can be represented, with regard for (17) and

(18), in the form

$$j_n = -\delta_t \frac{\partial N}{\partial t}. \quad (19)$$

When moving, each of the dislocations interacts with $1/a$ atoms within a unit length (a is the lattice constant). The expression for the photon flux from a unit surface area has the form

$$N_{\text{ph}} = \eta j/a$$

where η is the probability of emitting a photon. In view of (1), (2), and (19),

$$N_{\text{ph}} = \eta \delta_t a^{-1} N_0 \dot{\epsilon} \exp(-A\epsilon). \quad (20)$$

For copper, $a = 3.61 \text{ \AA}$ [19], $N_0 = 10^{14} \text{ m}^{-2}$, $\delta_t = 2.5 \times 10^{-4} \text{ m}$ [20], and $\eta = 10^{-9}$ [14].

The calculated time dependence of the photon flux is depicted in Fig. 1d. It should be noted that knowing the sensitivity of the optical detector and its position relative to the light-emitting surface and bearing in mind that the mechanoluminescence spectrum of copper has two lines with peaks at ≈ 1.7 and ≈ 2.1 eV [1], one can experimentally estimate the emission pulse intensity. From such estimates, the peak of the luminescence intensity I_{lum} (Fig. 1a) equals $\approx 1.5 \times 10^7$ ph/s, or $\approx 5 \times 10^{-12} \text{ W}$. Thus, the experimental and analytic spectra (Figs. 1a, 1d) can be compared by shape, duration, delay relative to the beginning of the irradiation (stress-

ing), and amplitude. It is seen that the spectra are in fairly good agreement. Thus, we may draw the conclusion that the dynamics of a luminescence pulse from the back side of a metal sample irradiated by a laser shot on its front side correlates with the dynamics of dislocations emerging on the surface.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 97-02-18097) and the Federal Program "Integration" (project no. KO854).

REFERENCES

1. K. B. Abramova and I. P. Shcherbakov, *Zh. Tekh. Fiz.* **64** (9), 76 (1994) [*Tech. Phys.* **39**, 901 (1994)].
2. A. A. Tupik and N. P. Valuev, *Pis'ma Zh. Tekh. Fiz.* **6** (2), 82 (1980).
3. V. T. Sotnikov, V. A. Gritsan, and A. A. Nechiporenko, *Pis'ma Zh. Tekh. Fiz.* **13**, 1291 (1987) [*Sov. Tech. Phys. Lett.* **13**, 539 (1987)].
4. E. A. Borisova, R. D. Glebova, A. A. Platonov, and N. M. Sklyarov, *Dokl. Akad. Nauk SSSR* **222**, 807 (1975) [*Sov. Phys. Dokl.* **20**, 435 (1975)].
5. V. I. Veksler, *Zh. Tekh. Fiz.* **52**, 1273 (1982) [*Sov. Phys. Tech. Phys.* **27**, 775 (1982)].
6. S. V. Vlasov, O. V. Farberovich, A. M. Bugakov, *et al.*, *Opt. Spektrosk.* **58**, 815 (1985) [*Opt. Spectrosc.* **58**, 499 (1985)].
7. N. L. Coleburn, M. Solow, and R. C. Wiley, *J. Appl. Phys.* **36**, 507 (1965).
8. A. M. Kondyrev, I. P. Shcherbakov, K. B. Abramova, and A. E. Chmel', *Zh. Tekh. Fiz.* **62** (1), 206 (1992) [*Sov. Phys. Tech. Phys.* **37**, 110 (1992)].
9. K. B. Abramova, I. P. Shcherbakov, A. M. Kondyrev, and I. Ya. Pukhonto, *Zh. Tekh. Fiz.* **66** (5), 190 (1996) [*Tech. Phys.* **41**, 511 (1996)].
10. A. F. Banishev, A. Ya. Panchenko, and A. V. Shishkov, *Fiz. Tverd. Tela (St. Petersburg)* **41**, 1538 (1999) [*Phys. Solid State* **41**, 1408 (1999)].
11. K. B. Abramova, A. I. Rusakov, A. A. Semenov, and I. P. Shcherbakov, *Fiz. Tverd. Tela (St. Petersburg)* **40**, 957 (1998) [*Phys. Solid State* **40**, 877 (1998)].
12. K. B. Abramova, A. I. Rusakov, A. A. Semenov, and I. P. Shcherbakov, *J. Appl. Phys.* **87**, 3132 (2000).
13. M. I. Molotskiĭ, *Fiz. Tverd. Tela (Leningrad)* **23**, 2171 (1981) [*Sov. Phys. Solid State* **23**, 1269 (1981)].
14. M. I. Molotskiĭ, *Fiz. Met. Metalloved.* **55**, 43 (1983).
15. B. P. Chandra, M. S. Khan, R. Singh Seema, and M. H. Ansari, *Cryst. Res. Technol.* **31**, 495 (1996).
16. K. B. Abramova, V. I. Vettegren', I. P. Shcherbakov, *et al.*, *Zh. Tekh. Fiz.* **69** (12), 102 (1999) [*Tech. Phys.* **44**, 1491 (1999)].
17. K. B. Abramova, V. I. Vettegren', I. P. Shcherbakov, and V. N. Svetlov, *Zh. Tekh. Fiz.* **71** (10), 112 (2001) [*Tech. Phys.* **46** (2001) (in press)].
18. H. Parkus, *Thermoelasticity* (Blaisdell, Waltham, 1968; Mir, Moscow, 1963).
19. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. S. Meilikhov (Energoatomizdat, Moscow, 1991; CRC Press, Boca Raton, 1997).
20. P. Lucas, M. Klesnil, and R. Fiedler, *Philos. Mag.* **20**, 799 (1969).

Translated by V. Isaakyan

OPTICS,
QUANTUM ELECTRONICS

The Bistable Mode of Second Harmonic Generation by Femtosecond Pulses

T. M. Lysak and V. A. Trofimov

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

Received November 24, 2000

Abstract—The possibility of controlling the efficiency of second harmonic generation by femtosecond pulses is demonstrated. By the phase shift in a certain cross section of the medium, switching from the low-efficient to the high-efficient generation state (and vice versa) is realized. The switching contrast in terms of the conversion efficiency is up to 4 (or more) and in terms of the peak intensity at the center of the pulse it is as high as 9. A new class of optical switches based on this effect is suggested. © 2001 MAIK “Nauka/Interperiodica”.

The problem of efficient second harmonic generation (SHG) by femtosecond pulses has remained unresolved to date [1] in spite of the long-standing efforts in this field [2–4]. The reason is known to be the self-action of waves that is caused by the cubic nonlinearity of a medium. For femtosecond pulses, the effect of this nonlinearity becomes comparable to that of quadratic nonlinearity. The self-action produces an additional phase shift of the interacting waves, violating the optimal phase relations. As a result, energy transfer to the second harmonic does not occur.

By using the SHG invariants (the quantities that retain their value during the wave interaction), it was shown [5] that the generation efficiency in the case of phase and group matching is defined by the ratio of the dimensionless constants characterizing the quadratic and cubic nonlinearity. The efficiency depends on this ratio monotonically: the higher the effect of the cubic nonlinearity, the lower the generation efficiency.

A qualitatively new feature of the SHG process appears if the generation proceeds in the absence of phase matching. It has been shown [6] that, in this case, two stable generation modes are possible. Consequently, wave self-action at certain conditions causes the spontaneous periodic switching of the generation process to the high-efficient state (branch), where a generation efficiency as high as 60% is achieved [7]. In this paper, we show that SHG can be controlled by switching the process from one stable state to the other by inducing a phase shift between the interacting waves in a certain cross section of the medium. It is important that at specific conditions, the contrast between these states is high. This offers the possibility of implementing an optical switch (processor) with a potentially small switching time, which is determined by the transition time of cubic nonlinearity. Thus, with the electronic mechanism of nonlinearity, the switching between the states may take from 10 to 50 fs (or less).

If it is assumed that the diffraction length far exceeds the length of a nonlinear medium, the set of dimensionless equations describing the process of SHG by a femtosecond pulse with regard to its self-action has the form [2]

$$\begin{aligned} \frac{\partial A_1}{\partial z} + iD_1 \frac{\partial^2 A_1}{\partial \eta^2} + i\gamma A_1^* A_2 e^{-i\Delta k z} \\ + i\alpha_1 A_1 (|A_1|^2 + 2|A_2|^2) = 0, \quad 0 < z \leq L_z, \\ \frac{\partial A_2}{\partial z} + \nu \frac{\partial A_2}{\partial \eta} + iD_2 \frac{\partial^2 A_2}{\partial \eta^2} + i\gamma A_1^2 e^{i\Delta k z} + i\alpha_2 A_2 \\ \times (2|A_1|^2 + |A_2|^2) = 0, \quad \alpha_2 = 2\alpha_1 = 2\alpha. \end{aligned} \quad (1)$$

Here, η is the dimensionless time in the coordinate system related to the basic wave pulse; z is the normalized longitudinal coordinate;

$$D_j \approx -0.5 \frac{\partial^2 \bar{k}_j}{\partial \bar{\omega}_j^2}$$

are the coefficients describing the second-order dispersion; \bar{k}_j and $\bar{\omega}_j$ are the dimensional wave number and the angular frequency of j th wave, respectively; γ is the coefficient of nonlinear coupling between the interacting waves; $\Delta k = k_2 - 2k_1$ is the dimensionless mismatch between the wave numbers; α_j are the wave self-action coefficients; and A_j are the complex amplitudes of the harmonics ($j = 1, 2$) normalized to the maximal amplitude of the first harmonic in the initial section of the medium ($z = 0$). The parameter ν is proportional to the difference of the reciprocals of the group velocities for the second and first harmonics, and L_z is the length of the nonlinear medium.

At the entrance into the nonlinear medium, the initial distribution of the fundamental frequency pulse is

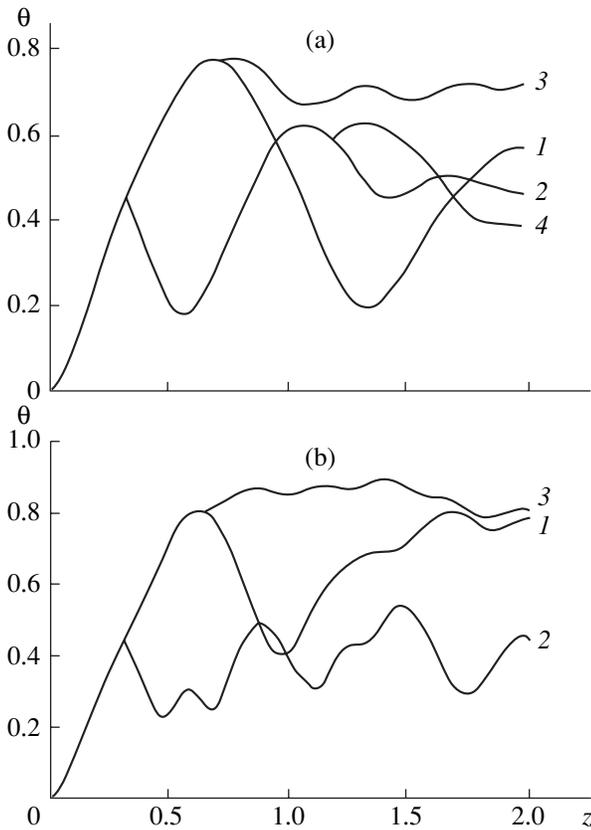


Fig. 1. Variation of the SHG efficiency (θ) without and (2–4) with switching at different points of the medium at $\alpha = 8$, $\gamma = 4$, $\Delta k = -2.35$, and $\nu = 0$ for (a) $D_1 = D_2 = 0.0000625$, and $\tau = 4$ and (b) $D_1 = D_2 = 10^{-3}$ and $\tau = 0.1$.

set in the form

$$A_1(z = 0, \eta) = A_0(\eta), \quad 0 \leq \eta \leq L_t, \quad (2)$$

where L_t is the dimensionless time for which the pro-

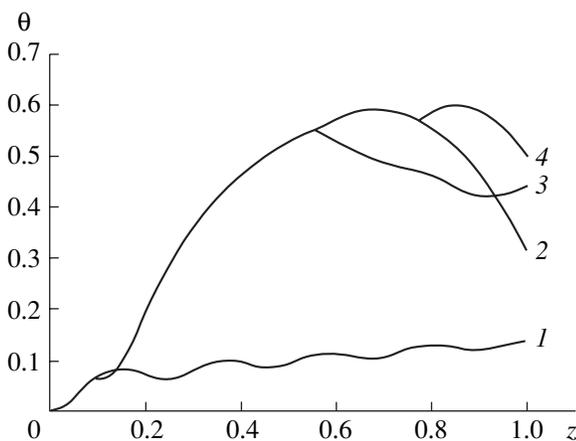


Fig. 2. SHG efficiency for $\alpha = 16$, $\gamma = 4$, $\Delta k = -2.0$, $D_1 = D_2 = 0.001$, $\nu = 0$, and $\tau = 10$ (θ) without switching into the high-efficient state and with the introduction of the phase shift $\Delta\phi =$ (2) -3 , (3) -1.5 , and (4) -2.5 in appropriate sections of the medium. The sections where the phase shift is introduced correspond to the bifurcation of the curves.

cess is considered; the amplitude of the second harmonic in this section is equal to zero: $A_2(z = 0, \eta) = 0$.

In physical experiments, the initial distribution of the fundamental wave pulse $A_0(\eta)$ is usually approximated by the Gaussian pulse:

$$A_0(\eta) = A_{10} \exp(-((\eta - L_t/2)/\tau)^2/2). \quad (3)$$

For our parameter normalization, the dimensionless amplitude at the fundamental frequency is $A_{10} = 1$.

The SHG process in the presence of wave self-action has a number of invariants [5]. Their values should be controlled upon the computer simulation of SHG. We used conservative difference schemes that retain the values of the invariants.

The efficiency of energy transfer from the fundamental harmonic to the second harmonic was evaluated as

$$\Theta(z) = \int_0^{L_1} |A_2(z, \eta)|^2 d\eta \bigg/ \int_0^{L_1} |A_1^0(\eta)|^2 d\eta. \quad (4)$$

We carried out numerical experiments including group matching between the interacting waves ($\nu = 0$).

It is known that the transfer (conversion) efficiency $\Theta(z)$ first smoothly increases from zero to a certain maximal value. In this case, the generation occurs within the first matching region where the phase difference between the interacting waves is less than π . Then, the efficiency oscillates near its mean value. The mean, in turn, is defined by the values of the problem parameters and rapidly drops with an increase in the self-action (the parameter α).

A considerable increase, as well as a considerable decrease, in the conversion efficiency within a specified distance from the entrance into the medium can be

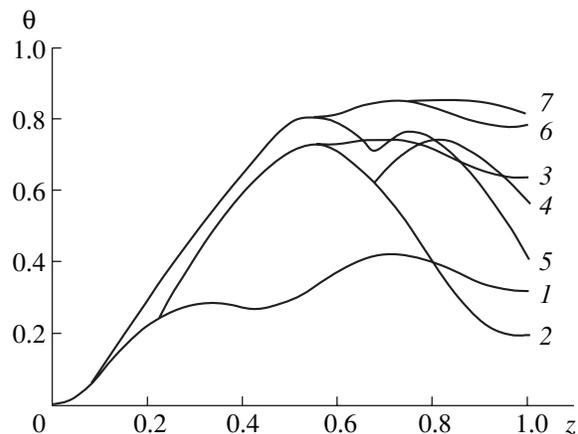


Fig. 3. SHG efficiency for $\alpha = 10$, $\gamma = 4$, $\Delta k = -2.35$, $\nu = 0$, $D_1 = D_2 = 0.0000625$, and $\tau = 4$ (θ) without switching into the high-efficient state and upon introducing (2) the phase shift $\Delta\phi = -1$; (3–6) double phase shift $\Delta\phi = -1$ and π ; and (7) triple phase shift $\Delta\phi = -1$, π , and π in appropriate sections. The sections where the phase shift is introduced correspond to the bifurcation of the curves.

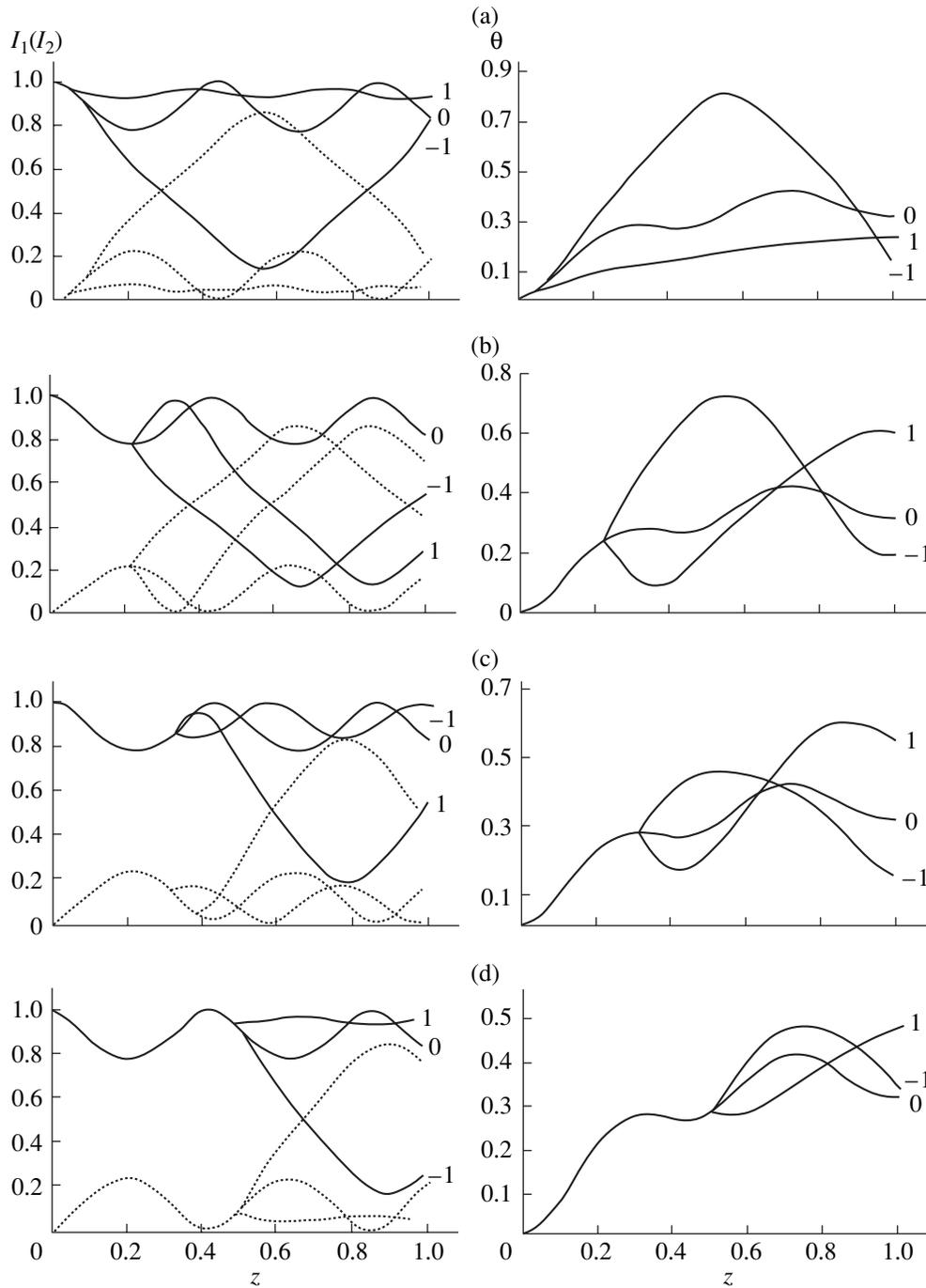


Fig. 4. Dependences of the spatial distributions of the efficiency (on the right) and of the intensity of the first, $I_1(z) = |A_1(z, \eta = L_r/2)|^2$ (continuous curves), and second, $I_2(z) = |A_2(z, \eta = L_r/2)|^2$ (dotted line), harmonics at the center of the pulse (on the left) on the position of the switching point and the phase shift sign at the switching point. $z =$ (a) 0.05, (b) 0.22, (c) 0.32, and (d) 0.5. Figures near the curves indicate the phase shift at the switching point. $\alpha = 10$, $\gamma = 4$, $\Delta k = -2.35$, $\nu = 0$, $D_1 = D_2 = 0.0000625$, and $\tau = 4$.

achieved by shifting the phases of the interacting waves inside the medium if the cubic nonlinearity dominates over the quadratic one. In this situation, the sign of the shift may sometimes effectively control the conversion efficiency. Note that we accomplished the phase shift between the interacting waves by simultaneously introducing the equal phase $\Delta\phi = \text{const}$ into the first and sec-

ond harmonics. It is this value $\Delta\phi$ that is given in the figure captions.

It should be emphasized that if the dimensionless coefficients characterizing the quadratic and cubic nonlinearities differ insignificantly ($\alpha \sim 2\gamma$), the switching only maintains the conversion efficiency at a high level

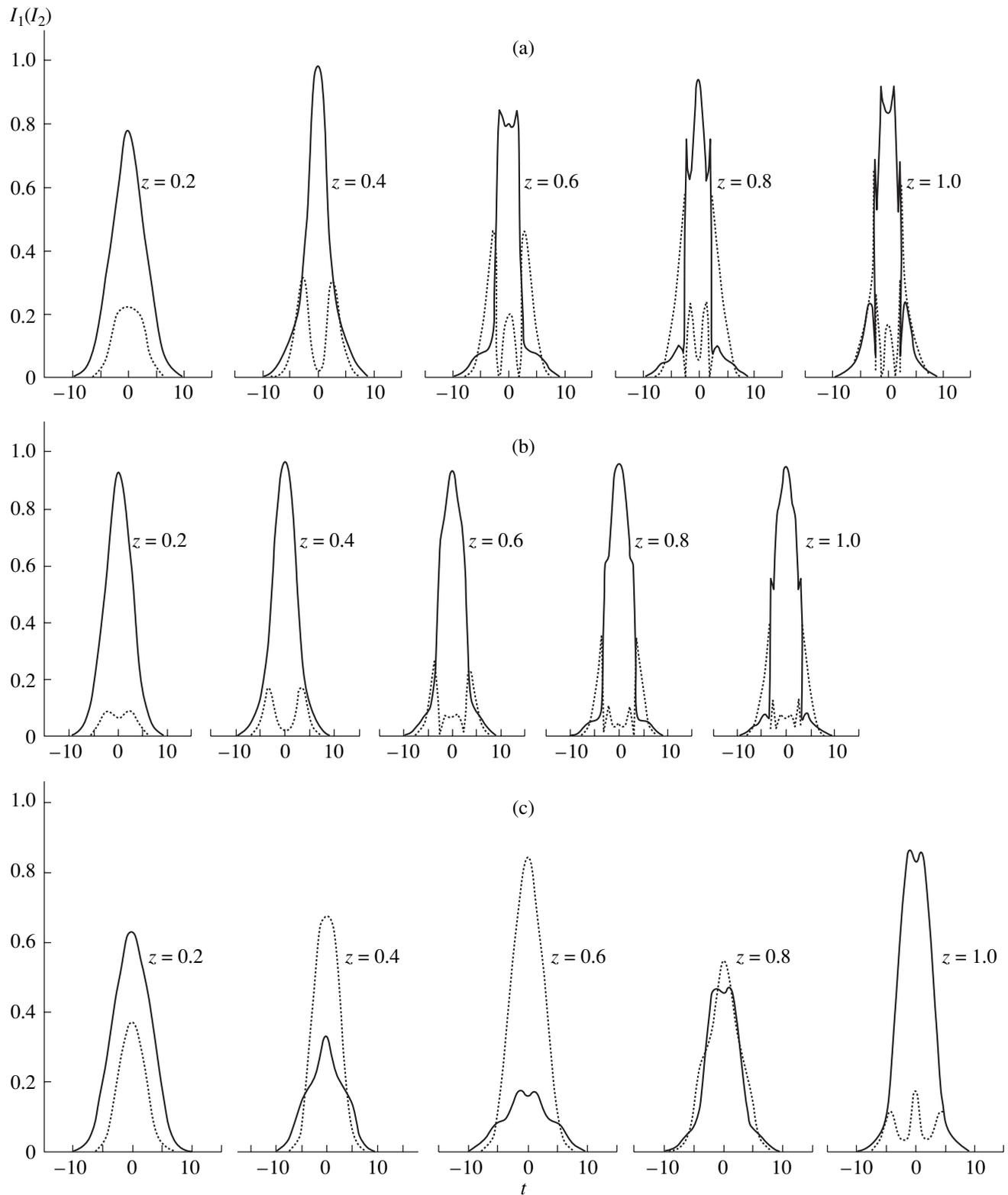


Fig. 5. Variation of the pulse shape during its propagation inward to the medium (a) without switching, with switching at the point $z = 0.05$ for the phase shift equal to (b) 1 and (c) -1 , and with switching at the point $z = 0.5$ for the phase shift equal to (d) 1 and (e) -1 . $\alpha = 10$, $\gamma = 4$, $\Delta k = -2.35$, $\nu = 0$, $D_1 = D_2 = 0.0000625$, and $\tau = 4$.

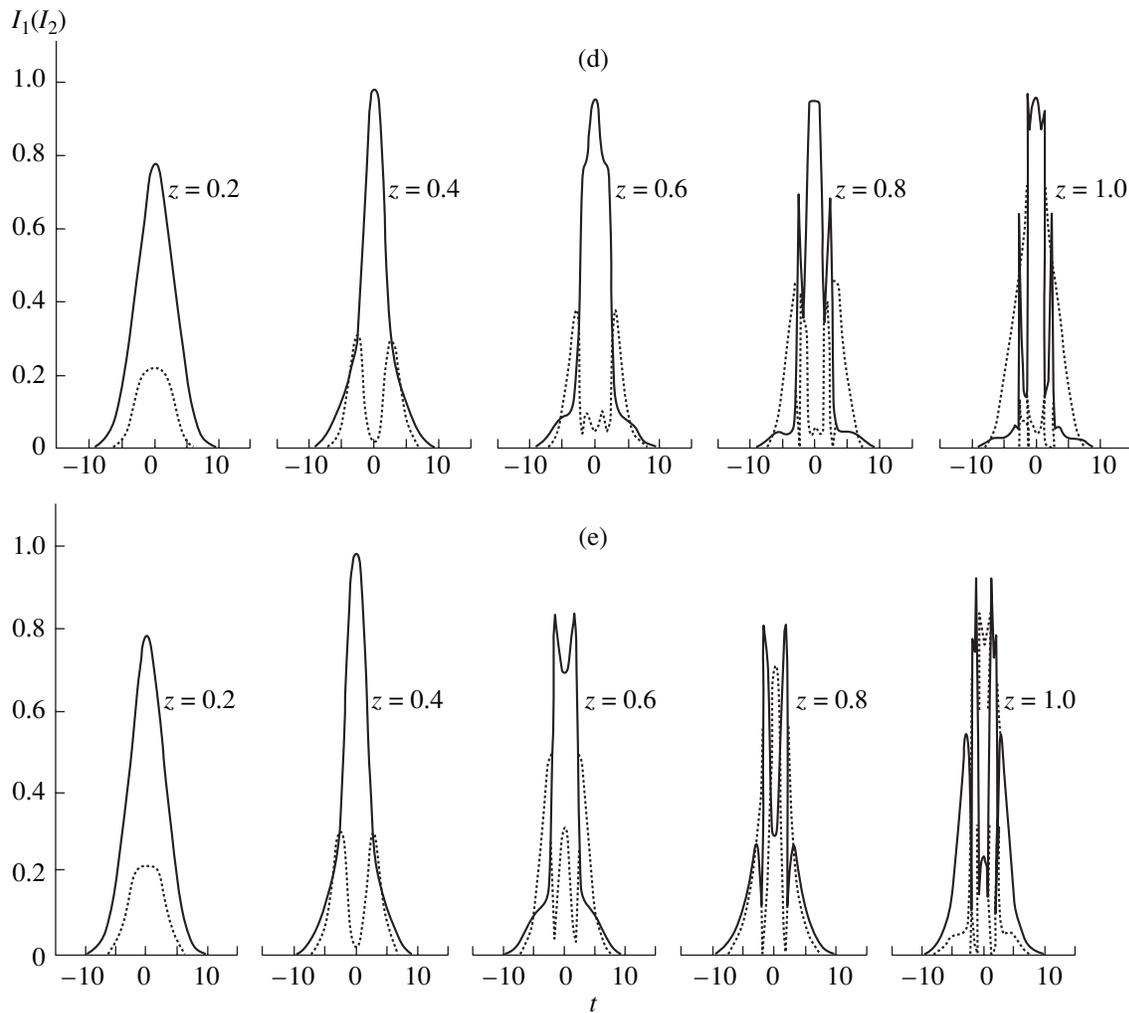


Fig. 5. (Contd.)

without increasing it. This conclusion holds both for sufficiently long pulses (Fig. 1a) and for short pulses (Fig. 1b), which are substantially affected by the second-order dispersion. The advantage of the switching to the high-efficient state, in this case, is that the conversion coefficient depends on the length of the medium only slightly.

The most interesting practical case is when the cubic nonlinearity greatly overrides the quadratic one during second harmonic generation by femtosecond pulses. In this case, the switching provides a high efficiency of generation at $\Delta k \neq 0$. This is illustrated in Fig. 2, where the variation of the conversion efficiency as the pulse propagates along the medium without switching (curve 1), as well as for single and double switching (curves 2–4), is shown. The first switching greatly increases the conversion efficiency (curves 1 and 2 in Fig. 2). Depending on the section taken to perform the second switching, the conversion efficiency can be degraded (curve 3) or kept at the level achieved (curve 4).

Note that similar results are also observed for other α and γ , particularly, for $\alpha = 10$ and $\gamma = 4$. However, the switching efficiency for these parameters is somewhat lower, because a higher energy should be transferred to the second harmonic under the conditions of phase matching. This statement is illustrated by Figs. 3 and 4, where the results of computer simulation for $\alpha = 10$, $\gamma = 4$, $\Delta k = -2.35$, $D_1 = D_2 = 0.0000625$, and $\tau = 4$ are presented. At some points inside the medium (the bifurcation points of the curves), the switching was performed at different phase shifts that maximize the efficiency (Fig. 3) or deviate (increase or decrease, Fig. 4) the efficiency from its nonperturbed value (when the phase shift is absent). The wave phase was varied near the entrance to the medium and in its interior.

As follows from Fig. 3, the conversion efficiency as high as 90% can be achieved by introducing a phase shift between the interacting waves in properly selected sections of the medium. It is significant to emphasize that the maximal difference in the efficiency values upon reversing the phase shift is observed for switching

points within the first quarter of the interval (Fig. 4). The efficiency is the highest for the positive shifts and the lowest, for the negative shifts.

Too early (Fig. 4a) or too late (Fig. 4d) switching considerably reduces the phase-shift effect at the switching point. It is important that in the former case, the conversion efficiency at the end of the interval decreases. In the latter case, conversely, the efficiency increases regardless of the direction of the shift. In both cases, the difference in the efficiency values at the end of the interval for the switched waves is small. Therefore, the switching contrast must be small as well. However, for the early switching, there is a domain in the middle of the interval where the switching contrast is high. When the switching point tends toward the end of the interval, the contrast decreases and becomes insignificant for the latest switching.

However, the variation of the pulse shape during its propagation inside the medium implies that only the early switching (Fig. 4a) can retain the shape of the propagating pulses; as a result, the same generation mode is observed throughout the pulse width (Figs. 5b, 5c). In this case, the essential (several-fold) difference between the peak intensities of the first and second harmonics can be achieved (Figs. 5b, 5c; section $z = 0.6$). In the absence of switching (Fig. 5a), as well as when the switching is late (Figs. 5d, 5e), the uniform generation mode and the appropriate contrast (the ratio of the intensities) are observed only in the area nearest to the pulse center due to self-focusing. The pulses break into several subpulses. As a result, the maximum attainable contrast of generation efficiency is reduced compared with the case of the early switching.

In closing, let us draw some conclusions from the results on SHG by femtosecond pulses with regard to wave self-action. First, to achieve a high generation efficiency, the condition of phase matching, used in the case of SHG by longer (picosecond or still longer) pulses, should be abandoned. Second, by introducing the phase shift between the interacting waves in certain sections, the generation efficiency up to 80% (or more) can be achieved. Third, when providing a high generation efficiency, one can control the second harmonic pulse shape. In particular, both a Gaussian pulse and a pulse that has a dip in the intensity distribution along

the axis can be realized. Hence, different modes of the second harmonic can be generated.

Fourth, by switching generation modes, we can realize bistable generation with the generation efficiencies and peak intensities of the harmonics differing by several times. This enables the development of a new class of optically bistable systems that offer high efficiency and switching rate in the femtosecond range if an additional phase shift is introduced in the same time scale. As a result, an optical processor based on the above principle can be designed.

It is also worth noting that modulation instability may prevent the achievement of high-efficient SHG by femtosecond pulses with a very high intensity. However, this effect is beyond the scope of this study.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (grant no. 99-01-01233) and by the program "Universities of Russia: Basic Research" (grant no. 992409).

REFERENCES

1. G. Steinmeyer, P. H. Sutter, L. Gallmann, *et al.*, *Science* **286**, 1507 (1999).
2. T. B. Razumikhina, L. S. Telegin, A. I. Kholodnykh, and A. S. Chirkin, *Kvantovaya Élektron. (Moscow)* **11**, 2026 (1984).
3. P. P. Ho, Q. Z. Wang, and R. R. Alfano, *Opt. Lett.* **16**, 970 (1991).
4. T. Ditmire, A. M. Rubenchik, D. Eimerl, and M. D. Perry, *J. Opt. Soc. Am. B* **13**, 649 (1996).
5. I. S. Ivanova, S. A. Magnitskiĭ, and V. A. Trofimov, *Vestn. Mosk. Univ., Ser. Vychisl. Mat. Kibern.*, No. 1, 21 (2000).
6. T. M. Lysak and V. A. Trofimov, *Zh. Vychisl. Mat. Mat. Fiz.* **41** (8) (2001) (in press).
7. T. M. Lysak and V. A. Trofimov, in *Technical Program of International Conference LO'2000, St. Petersburg, 2000*, p. 48.

Translated by M. Lebedev

ACOUSTICS,
ACoustoelectronics

Interaction between Light and Acoustic Microwaves Excited by Aperiodic Multielement Transducers. II

M. A. Grigor'ev, A. V. Tolstikov, and Yu. N. Navrotskaya

Chernyshevsky State University, ul. Dvadsatiletiya VLKSM 112a, Saratov, 410071 Russia

Received July 3, 2000; in final form, January 9, 2001

Abstract—Aperiodic multielement piezoelectric transducers operating in the high-frequency part of the microwave range are studied theoretically. The efficiency of interaction between a plane light wave and the acoustic fields excited by the transducers with antiphase adjacent elements is calculated as a function of frequency. Structures where the pitch varies gradually or stepwise and the piezoelectric elements are wide- or narrow-band are analyzed. © 2001 MAIK "Nauka/Interperiodica".

INTRODUCTION

The purpose of this paper (see also part I in [1]) is to compare aperiodic and periodic multielement electroacoustic transducers (MEATs) in terms of the efficiency of acousto-optical (AO) interaction under the condition that tuning to the Bragg angle is provided in a given frequency band. We study antiphase MEATs, in which adjacent elements are excited with a phase shift $\varphi_0 = \pi$ rad. Three types of aperiodic MEATs are considered: (i) MEATs with a gradually varying pitch and wide-band piezoelectric elements, (ii) those with a gradually varying pitch and narrow-band piezoelectric elements, and (iii) sectionalized transducers with narrow-band piezoelectric elements. As a reference periodic MEAT, we use an antiphase transducer. This transducer is known (see Ref. [5] in [1]) to produce a partial acoustic wave (usually employed for the interaction with incident light) that is more intense than that produced by an inphase transducer. We assume that the divergence of the incident light in the plane of AO interaction is significantly smaller than the divergence of the acoustic wave. This situation occurs in deflectors. Acoustic oscillations excited by variable-pitch MEATs in a photoelastic medium have complex spatial amplitude and phase distributions. Therefore, we will first derive a formula for the AO interaction efficiency in the general case of a sectionalized transducer.

AO INTERACTION EFFICIENCY OF AN APERIODIC (VARIABLE-PITCH) TRANSDUCER

As is known, upon AO interaction, light diffraction is due to the variation of the permittivity ε of the medium. For isotropic diffraction, this variation is related to the strain S by the formula $\Delta\varepsilon = -\varepsilon^2 p S$ [2], where p is the photoelastic constant. Therefore, to find the intensity of the diffracted light, one should know the distribution of the strain amplitude and phase on the

surface of the medium. Consider a sectionalized MEAT, which is the most complex device among those addressed in this paper. Figure 1 shows the distribution of the strain amplitude S over the x coordinate in the transducer plane and also the basic parameters. The transducer contains n sections, having M_i elements each; i is the section no.; v_i is the element no. in the i th section; and l_i and L_i are, respectively, the period and the length of the elements in the i th section. The strain amplitude is constant over the surface of each of the piezoelectric elements and equals zero between the elements. The oscillation phases of adjacent elements are shifted by an angle φ . Assume that the power P_{v_i} of the acoustic wave excited by each of the elements is known. Using the Umov vector, we can express the strain on the surface of a piezoelectric element in terms of the emitted acoustic power:

$$S_{v_i}^2 = 2P_{v_i}/(\rho v_{ac}^3 H L_i), \quad (1)$$

where H is the width of the piezoelectric element.

The distribution $S(x)$ over the MEAT plane can be represented through the Fourier integral as a continuous plane-wave spectrum, the amplitude of each of the plane waves being equal to $g(k_x)dk_x$, where k_x is the projection of the wave vector k on the x axis. The density $g(k_x)$ of this spectrum is given by the Fourier transform

$$g(k_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(x) e^{-jk_x x} dx. \quad (2)$$

One can write $k_x = (2\pi/\Lambda)\sin\gamma$, where Λ is the acoustic wavelength and γ is the angle between the normal to the transducer plane and the direction of the wave vector of the elementary wave. Then, expression (2) gives the value of $g(\gamma)$, referred to as the angular spectrum.

Thus, AO interaction between the plane optical wave and the diverging acoustic beam amounts to interaction between the plane waves. The diffracted light is

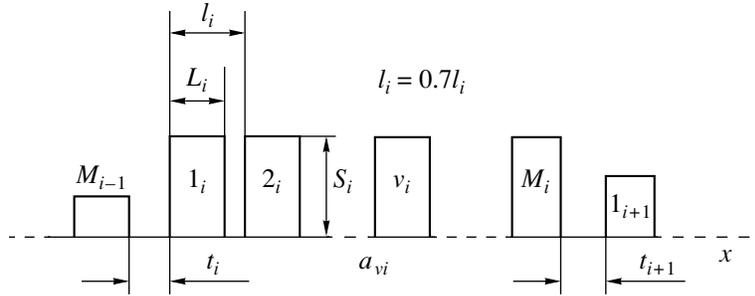


Fig. 1. Strain vs. x coordinate for the sectionalized transducer.

produced only by the elementary component of the angular spectrum that satisfies the Bragg condition. If a weakly diverging optical beam is directed so that it satisfies this condition at the center frequency f_0 at the midpoint of a lobe of the angular spectrum, a diffraction maximum will be observed at this frequency.

The light intensity in the diffraction maximum is actually finite, because the region of AO interaction is limited, which causes the light beam to diverge. The AO interaction thus involves those components of the angular spectrum falling into a finite angular interval $\Delta\gamma$. As a result, the amplitude of the diffracted light becomes finite.

If the frequency differs from f_0 , another component of the acoustic angular spectrum will meet the Bragg condition and the amplitude of the diffraction light will decrease almost in proportion to the decrease in the spectrum density $g(\gamma)$.

Thus, the AO interaction efficiency up to a constant factor can be found from expression (2) as the squared magnitude of the angular spectrum. This expression should use the angle $\gamma_B(f)$ that specifies the direction of the elementary wave satisfying the Bragg condition for given frequency f and direction Θ_0 of the incident light:

$$\gamma_B(f) = \Theta_0 - \Theta_B(f).$$

For the strain distribution $S(x)$ (Fig. 1), formula (2) can be recast as

$$g(\gamma) = \frac{1}{2\pi} \sum_{i=1}^n \sum_{v_i=1}^{M_i} \int_{a_{v_i}}^{(a_{v_i} + L_i)} S_{v_i} e^{-j(\chi_{v_i} - 1)\varphi} e^{-jk \sin \gamma x} dx,$$

where χ_{v_i} is the serial number of the v th element in the i th section; a_{v_i} is the coordinate of the left edge of the v th element; and S_{v_i} is the coordinate-independent strain on the v th element in the i th section, which can be found from formula (1) if the power P_{v_i} of elastic oscillations excited by this element is known.

Integrating the above formula and taking the square of the magnitude, we obtain

$$\eta_{AO} = \frac{I(f)}{I_0} = K \left| \sum_{i=1}^n \left(\sqrt{L_i} \frac{\sin \Gamma_i}{\Gamma_i} e^{-j\Gamma_i} \sum_{v_i=1}^{M_i} \sqrt{P_{v_i}} e^{-j\Phi_{v_i}} \right) \right|^2, \tag{3}$$

where $I(f)$ is the intensity of the diffracted light, I_0 is the intensity of the incident light, K is a proportionality coefficient, and

$$\Gamma_i = \frac{\pi L_i f}{v_{ac}} \sin[\Theta_0 - \Theta_B(f)],$$

$$\Phi_{v_i} = (\chi_{v_i} - 1)\varphi + \frac{2\pi f}{v_{ac}} \sin[\Theta_0 - \Theta_B(f)] a_{v_i}.$$

Under the assumption that the AO interaction is weak ($\eta_{AO} \ll 1$), K can be estimated by comparing formula (3) for a single piezoelectric element ($n = M_i = 1$, $\Theta_0 = \Theta_B$) with the Gordon formula [3]:

$$K = \pi^2 M_2 / 2\lambda_0^2 H \cos^2 \Theta_B.$$

A SINGLE PIEZOELECTRIC ELEMENT AND A PERIODIC MEAT

In order to compare aperiodic MEATs with conventional transducers and to reveal their advantages and disadvantages, we calculated the AO efficiencies of a single piezoelectric element and of a multielement periodic antiphase structure in the frequency range of 8–10 GHz. The lengths of the transducers were chosen such that the AO efficiency varied within ≈ 3 dB in the above range. We assumed that the transducers comprised wide-band piezoelectric elements made of Z-oriented zinc oxide ($0.09 \mu\text{m}$) with copper sublayer ($0.1 \mu\text{m}$) and overlayer ($0.8 \mu\text{m}$). As a photoelastic medium, we took X-cut lithium niobate, in which the transducer excited longitudinal elastic waves. Figure 2 plots the frequency dependence of the conversion factor for the case when the piezoelectric element terminates the optimal transmission line. The conversion factor is

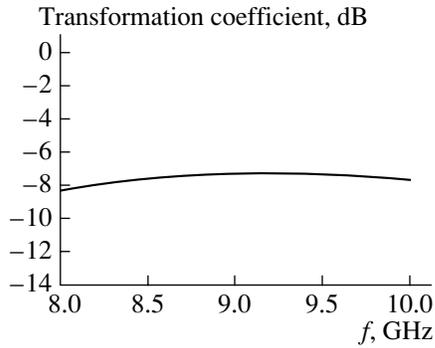


Fig. 2. Conversion factor vs. frequency for a single zinc oxide piezoelement ($h = 0.09 \mu\text{m}$). The wave impedance of the transmission line is $Z_0 = 0.38 \Omega$.

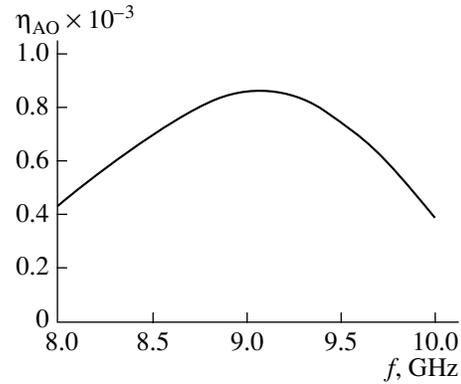


Fig. 3. AO interaction efficiency vs. frequency for a single piezoelectric element at $L = 15 \mu\text{m}$ and $P_{ac} = 0.03 \text{ W}$.

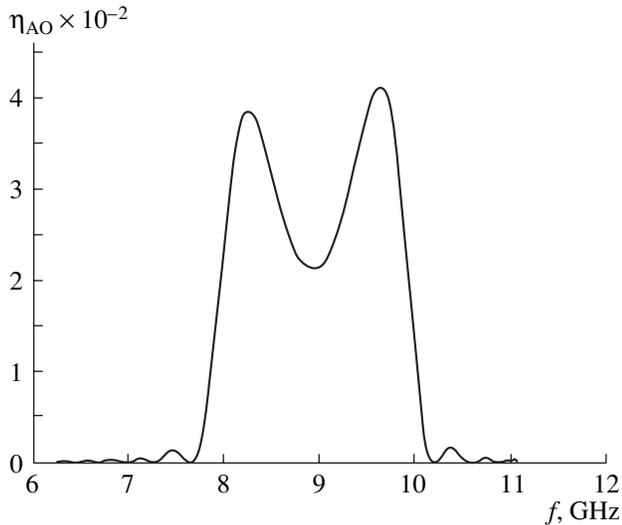


Fig. 4. AO interaction efficiency vs. frequency for a periodic MEAT at $M = 135$ and $l = 1.9 \mu\text{m}$.

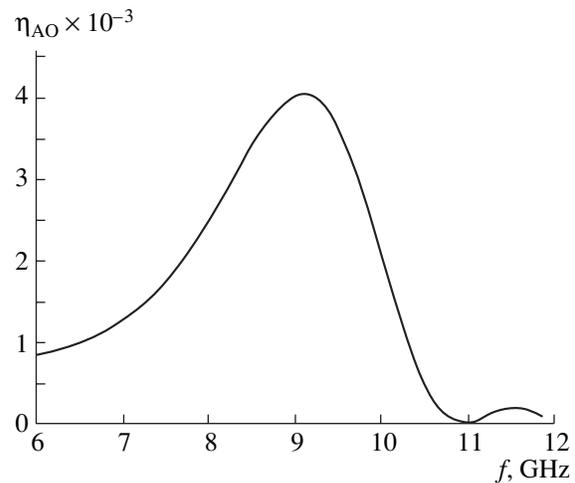


Fig. 5. AO interaction efficiency vs. frequency for a periodic MEAT at $M = 7$, $l = 3.5 \mu\text{m}$, and $Z_0 = 85 \Omega$.

seen to vary from -8.3 to -7.3 dB in the frequency range of 8–10 GHz. In our calculations, the transverse size of the piezoelectric element was $50 \mu\text{m}$ and its longitudinal dimension (in the direction of the incident light) was set equal to $15 \mu\text{m}$. Figure 3 shows the AO efficiency $\eta_{AO}(f)$ of the single piezoelectric element at a microwave power of 1 W and an angle of light incidence $\Theta_0 = \Theta_B(f_0 = 9 \text{ GHz}) = 0.197$ rad. The efficiency η_{AO} reaches its maximum (0.085%) at 9 GHz, decreasing to 0.045% at the bounds of the frequency range.

We considered two versions of periodic antiphase MEATs. The first version had the period of $1.9 \mu\text{m}$; the second, $3.5 \mu\text{m}$. The lengths of the elements were equal to $0.7l$, i.e., 1.33 and $2.45 \mu\text{m}$, respectively. To provide self-tuning near 9 GHz, the angles Θ_0 were set equal to 0.392 and 0.3019 rad, respectively, according to formula (3) in [1]. In the first case, the working point at $f = 9$ GHz was at the minimum of the curve $l(f)$; in the second case, the working point lay at its high-frequency

slope [1, Fig. 1]. For both transducers to provide the above bandwidth, we had to use 135 elements in the first transducer and only seven elements in the second one. The power emitted by the piezoelectric elements was calculated by formula (16) in [1] at $P_{al} = 1 \text{ W}$. We also assumed that the transmission lines have the optimal wave impedances which we estimated at $Z_{0, \text{opt}} = 8.5 \Omega$ and 85Ω . The resulting $\eta_{AO}(f)$ curves are shown in Figs. 4 and 5, respectively. The former has the double-humped shape typical of an antiphase MEAT with the optimal period. The maximal AO efficiency reaches 4.3%. The latter curve has the single-humped shape with $\eta_{AO, \text{max}} = 0.4\%$ at 9 GHz.

Thus, the optimum-period ($1.9 \mu\text{m}$) MEAT proves to be much more efficient for an AO deflector than a single piezoelectric element or the MEAT with the nonoptimal period ($3.5 \mu\text{m}$). The only advantage of the latter MEAT is that it is easier to fabricate because of its longer period and smaller number of elements. How-

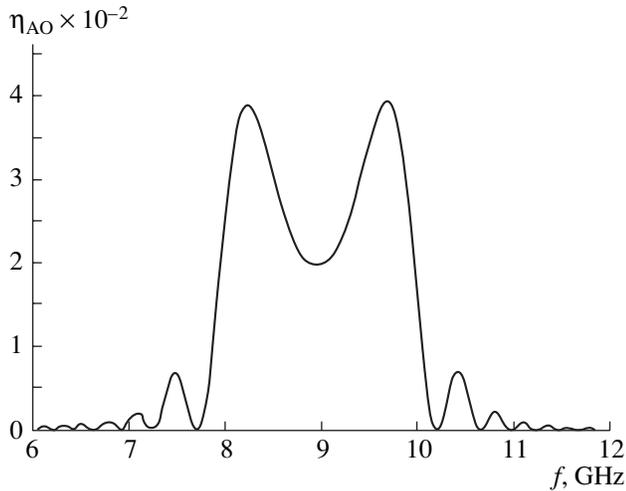


Fig. 6. AO interaction efficiency vs. frequency for an aperiodic MEAT with wide-band piezoelectric elements at $M = 135$, $l = 1.9 \mu\text{m}$, and $P_{ac} = 0.186 \text{ W}$.

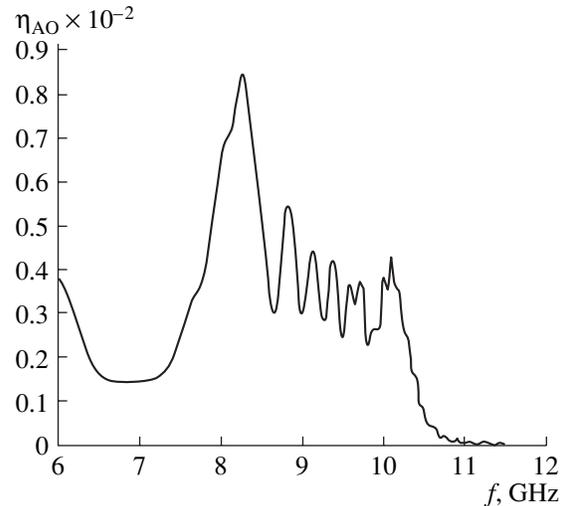


Fig. 7. AO interaction efficiency vs. frequency for an aperiodic MEAT with wide-band piezoelectric elements at $M = 135$, $l = 3.22\text{--}4.15 \mu\text{m}$, and $P_{ac} = 0.186 \text{ W}$.

ever, its low diffraction efficiency makes the application of this MEAT unpromising.

AN APERIODIC MEAT WITH A GRADUALLY VARYING PITCH AND WIDE-BAND PIEZOELECTRIC ELEMENTS

We considered two versions of these aperiodic transducers. They differed in the distribution $l(x)$. Both distributions were obtained by the algorithm described in [1]. For the first version, the angle of light incidence was $\Theta_0 = 0.393 \text{ rad}$; for the second, 0.3019 rad . The pitch versus frequency dependences for the two versions satisfy formula (3) in [1] and are shown in Fig. 1 therein. The lower curve with a minimum at 9 GHz refers to the first version; the upper curve, to the second one. The pitch values are noteworthy. In the first version, the pitch varies insignificantly. It equals $\approx 1.91 \mu\text{m}$ at the ends of the transducer and $1.89 \mu\text{m}$ in its central part. The maximum variation in the pitch is $\approx 1\%$. In the second version, the pitch increases along the transducer from 3.22 to $4.15 \mu\text{m}$. Its maximal variation was found to be $\approx 25\%$ of the average value. The elements' lengths were always equal to $0.7l_i$, where l_i is the pitch next after the i th element. Also note that both transducers involved the wide-band piezoelectric elements similar to those applied previously and that the photoelastic medium was lithium niobate as before. Both MEATs terminated the optimal transmission lines, providing a conversion factor of -7.3 dB . Thus, 0.186 W of 1 W of the electromagnetic power was converted into acoustic waves. We assumed that this power was distributed among the elements in proportion to their area.

Figures 6 and 7 depict the AO interaction efficiency vs. frequency for the two MEATs with a gradually varying pitch. As expected, the curve in Fig. 6, obtained for

135 elements, is double-humped as for the periodic MEAT with the same number of elements (Fig. 4). The AO interaction efficiency decreased by as small as $\approx 0.15\%$ compared with the periodic MEAT at the same bandwidth.

Our calculations show that increasing the number of the elements with the maximum difference in the pitch along the transducer remaining unchanged narrows the $\eta_{AO}(f)$ curve at a level of 3 dB and increases its height. However, these changes are slower than in the periodic MEAT. For example, at $M = 1000$, the transmission bandwidth of the periodic MEAT shrinks to 0.65 GHz at a 32% AO efficiency, while that of the aperiodic MEAT, to 0.95 GHz at a 10% efficiency. Note that the $\eta_{AO}(f)$ curve for the aperiodic MEAT with a large number of elements exhibits strong oscillations on both sides of the central double-humped part.

Thus, at a large number of wide-band elements, the aperiodic (variable-pitch) MEAT operating near the minimum of the $l(f)$ curve (see formula (3) in [1]) has the significantly lower AO interaction efficiency and the slightly wider bandwidth than the periodic MEAT.

Figure 7 shows the same curve for the second version of the aperiodic MEAT considered in this section. The MEAT consists of 135 elements. The curve is asymmetric (single-humped) and strongly oscillates. Its upper part resembles an exponential with anharmonic oscillations imposed on it. The swing of these oscillations is $\approx 0.13\%$ in the central part and exceeds 0.5% at the left edge. The efficiency attains $\approx 0.9\%$ near 8.4 GHz and drops to $\approx 0.32\%$ at 9.5 GHz . The width of the left-most peak at a level of 0.45% is $\approx 0.6 \text{ GHz}$. Increasing the number of the elements, while keeping the maximum pitch variation unchanged (within the range from 3.22 to $4.15 \mu\text{m}$) increases the number of the oscillations and slightly decreases their amplitude.

The left-most peak slightly increases, while the right-most one remains nearly the same. For example, at $M = 300$, the number of the oscillations is 10, while at $M = 1000$, it is greater than 30. The height of the left-most maximum becomes 1.04 and 1.2%, respectively, and its half-height width is 0.45 and 0.3 GHz. The amplitude of the oscillations at $M = 1000$ decreases in the central part of the curve to 0.1%.

Thus, aperiodic MEATs with a gradually varying pitch and wide-band piezoelectric elements cannot improve the AO interaction efficiency if the bandwidth must be kept constant.

A SECTIONALIZED MULTIELEMENT TRANSDUCER

The center frequency, period, and number of the elements in each of the sections were chosen as follows. First, the material and the thicknesses of the layers incorporated into the piezoelectric elements were selected so that a $\approx 3\%$ operating band was provided. Then we used the program for calculating the AO interaction efficiency of the periodic MEAT. Initially, the center frequency of the first subband was tentatively specified, the period was estimated by formula (3) in [1], and $\eta_{AO}(f)$ for various MEAT lengths was constructed. The total acoustic power emitted by the transducer was set equal to 0.1 W. The relative half-height width of the resulting curve was $\approx 3\%$. Then, we refined the center frequency of the first subband so that its lower limit was placed at 8 GHz. The final values of the center frequency, period, and number of the elements in the first section were 8.12 GHz, 3.27 μm , and 112, respectively. The subband width and the maximal AO efficiency appeared to be equal to 275 MHz and 0.037, respectively.

When finding the center frequency, period, and number of elements for the next subband, we selected its length in such a way as to provide a 0.037 AO efficiency, which automatically defines the subband width. The resulting $\eta_{AO}(f)$ curve was then matched to that for the previous section at the half-height level. As a result, the range 8–10 GHz was covered by a 13-section transducer with a total of 1345 elements.

After finding the center frequencies of the subbands, the overlayer thicknesses were chosen such that the piezoelectric elements can operate at these frequencies. To this end, we used the program based on formulas (6)–(9) in [1] to calculate the conversion factor. We assumed that the piezoelectric elements terminate optimal transmission lines. The parameters of all the 13 sections are summarized in the table in [1].

To calculate the AO interaction efficiency versus frequency, we developed the following algorithm. Given the material constants of the piezoelectric element layers and their dimensions, the impedances of the piezoelectric elements were calculated for each of the frequencies and then the power emitted by each of the ele-

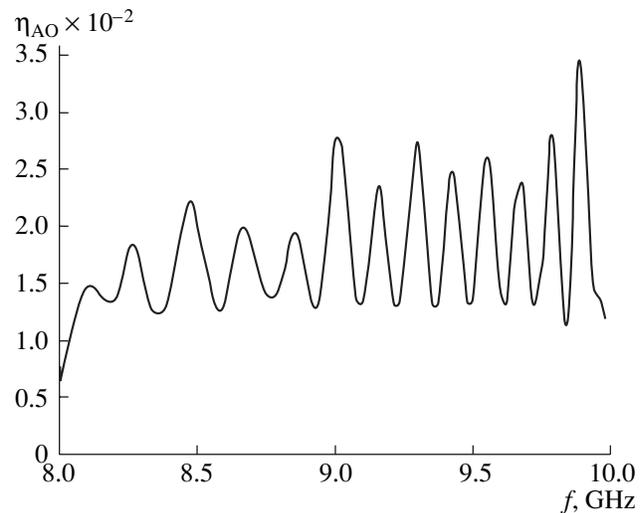


Fig. 8. AO interaction efficiency vs. frequency for a 13-section transducer. The sections are spaced at 1 μm .

ments was evaluated. We assumed that the sectionalized MEAT terminates the optimal transmission line. The output power at the matched load was assumed to be 1 W. The AO interaction efficiency was calculated from formula (3). The program generated a plot of η_{AO} versus frequency.

First of all, we studied the effect of the intersection spacing t_i (Fig. 1). To this end, t_i was varied from 0 to 1000 μm . The spacing was assumed either to be constant or to vary along the MEAT by the rule similar to that governing the pitch variation. In both cases, the efficiency $\eta_{AO}(f)$ exhibited significant oscillations with the amplitude increasing with frequency. The number of the oscillations was equal to the number of the sections, and their swing depended on the distance t_i . Initially, at small t_i , the swing decreased with increasing t_i , achieving its minimal value at $t_i \approx 1 \mu\text{m}$. Subsequently, it steadily grew as the dips between the maxima became deeper. Even at $t_i = 5 \mu\text{m}$, the minimal efficiency was close to zero. Figure 8 shows the curve at $t_i = 1 \mu\text{m}$. It can be seen that the first maximum of η_{AO} reaches $\approx 1.5\%$; the third, $\approx 2.2\%$; the fourth and fifth, $\approx 1.9\%$; the sixth, $\approx 2.5\%$; the seventh, $\approx 2.2\%$; the eighth, $\approx 2.5\%$; the ninth, tenth, eleventh, and twelfth slightly oscillate near $\approx 2.4\%$; and the thirteenth is $\approx 3.5\%$. All the minima are close to $\approx 1.3\%$. All the oscillations except for the last maximum fall into the 3-dB-wide range. The mean value of η_{AO} appears to be small: even smaller than that of a single section. This fact can be attributed to the shunting effect of idle elements, which do not generate elastic waves at the given frequency. They act as an additional spurious capacitor connected to the active elements. This effect also decreases the electroacoustic conversion factor even if the optimal transmission line is used.

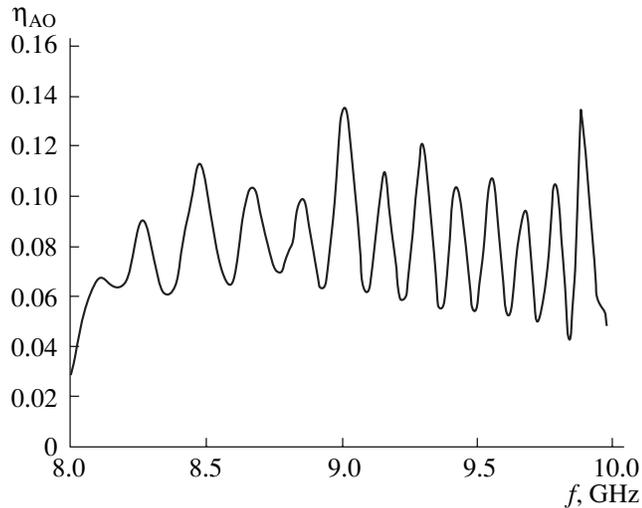


Fig. 9. AO interaction efficiency vs. frequency for the 13-section transducer. The intersection spacings are equal to the gaps between the preceding elements. The additional inductance $L_{\text{add}} = 0.03$ nH, and $Z_0 = 0.5$ Ω .

The efficiency η_{AO} can significantly be improved by connecting an additional inductance coil with an inductance L_{add} in series with the sectionalized transducer in order to at least partially compensate for the capacitance of the transducer and/or by using a transmission line with a lower wave impedance Z_0 . For example, at $L_{\text{add}} = 0.035$ nH and $Z_0 = 0.35$ Ω , the AO efficiency increases to $\sim 15\%$ at 8.2–8.4 GHz and to $\sim 6\%$ at 9.8 GHz. One can vary the shape of the $\eta_{\text{AO}}(f)$ curve by varying L_{add} and Z_0 , i.e., by moving one or another part of the curve up or down. Figure 9 shows the AO efficiency calculated at $L_{\text{add}} = 0.03$ nH and $Z_0 = 0.5$ Ω . The plot contains 13 peaks as before, the highest ones (the sixth and thirteenth) reaching 12.8%; the third and eighth, $\sim 12\%$. The deepest minimum, $\sim 4.5\%$, lies between the twelfth and the thirteenth peaks.

Emphasize the basic result of the above analysis: the sectionalized MEAT allows one to increase the AO efficiency, while retaining the bandwidth wide. This conclusion is valid even if the pitches in the sections are far away from the minimum of expression (3).

A significant disadvantage of the derived function $\eta_{\text{AO}}(f)$ is its strongly oscillating behavior. The oscillations near the upper limit of the frequency range considered are higher than 3 dB. One can attempt to suppress the oscillations by selecting appropriate center frequencies of the subbands and numbers of the elements in the sections. This point, however, will be omitted in this article.

AN APERIODIC MEAT WITH NARROW-BAND PIEZOELECTRIC ELEMENTS

We analyzed the case when $l(x)$ is described by the lower curve in Fig. 1 in [1], for which $l = 1.9$ μm at $f_0 =$

9 GHz. The value of l for a particular element number χ was calculated by the algorithm described in [1]. The angle of light incidence was assumed to equal 0.393 rad, the frequency was varied from $f_1 = 7.5$ GHz to $f_2 = 10.5$ GHz, and the number of elements M was between 6 and 2802. The program used for calculating η_{AO} could also compute the pitch versus element number χ , the impedances Z_χ , and the powers $P_{\text{ac},\chi}$. The thicknesses of the ZnO piezoelectric ($h_1 = 0.3193$ μm) and of the Al sublayer ($h_2 = 0.15965$ μm) were the same for all the piezoelectric elements. The Cu overlayer thickness h_3 was chosen for each of the elements so that their frequencies were distributed within the above range. As a preliminary, we calculated the frequency dependencies of the conversion factor $\eta_{\text{ac}}(f)$ for various thicknesses of the overlayer when the piezoelectric element was connected with the oscillator via the optimal transmission line. These dependences were used to find the frequencies f_{h_3} at which η_{ac} acquire the maximal values and the bandwidths where these values dropped by no more than 3 dB. Using the frequencies, we constructed a piecewise-linear approximation for the thickness h_3 versus f_{h_3} dependence: $h_3 = A_j + B_j f_{h_3}$, where A_j and B_j are coefficients and j is the number of the frequency interval to which they refer. The range 7.5–10.5 GHz was divided into six intervals. The entry to the program were thus six pairs of coefficients A_j and B_j . In this frequency range, the overlayer thickness changed from ~ 0.38 to ~ 0.2 μm . For each of the six intervals to contain an integer number of elements, their total number must be a multiple of 6. Our calculations also showed that the bandwidth of the piezoelectric elements varies from 898 to 776 MHz as h_3 grows in this thickness range.

Thus, only some of the piezoelectric elements could operate at a given frequency. At the lower frequency limit, acoustic waves were excited by $\sim 24\%$ of the elements; at the higher frequency limit, by $\sim 28\%$. The acoustic power was calculated from formula (16) on assumption that $P_{\text{al}} = 1$ W and the wave impedance Z_0 is equal to the magnitude of the MEAT impedance.

The AO interaction efficiency was shown to gradually increase with M from 0.26% at $M = 12$ ($Z_{0,\text{opt}} = 352$ Ω) to 21% at $M = 2802$ ($Z_{0,\text{opt}} = 1.52$ Ω). Accordingly, the bandwidth of AO interaction decreases from 3 to ~ 0.7 GHz. At 138 elements ($Z_{0,\text{opt}} = 30.77$ Ω), the maximal η_{AO} was found to be 2.74% and the bandwidth was 2.45 GHz. Only at $M = 252$ ($Z_{0,\text{opt}} = 16.85$ Ω) did the bandwidth decrease to 2 GHz and the maximal efficiency reach 4.6%. Figure 10 shows the $\eta_{\text{AO}}(f)$ curve for the last case. The curve is asymmetric and double-humped, which is associated with the lower electroacoustic conversion factor and the wider bandwidth of the piezoelectric elements designed for the higher frequency portion of the range.

Thus, we can conclude that, if the bandwidth of AO interaction is ~ 2 GHz, the application of narrow-band

piezoelectric elements in the aperiodic MEAT is to little avail. For example, for a MEAT with wide-band piezoelectric elements, we have $\eta_{AO, \max} = 4.0\%$ at $M = 138$, while for a similar MEAT with narrow-band elements, $\eta_{AO, \max} = 4.6\%$ at $M = 252$.

COMPARISON OF PERIODIC AND APERIODIC MEATs

In order to correctly compare the performance of different MEATs, one should keep in mind that, in the above analysis, the electromagnetic power is applied to the transducer via a transmission line with the optimal wave impedance equal to the magnitude of the MEAT impedance. When narrow-band piezoelectric elements are used, their center frequencies are distributed in the specified frequency range. As a result, only few elements operate at a particular frequency, the remaining elements acting as a shunt capacitor, which decreases the electroacoustic conversion factor. This decrease occurs, because the capacitive contribution to the magnitude of the impedance increases. Therefore, the magnitude of the reflection coefficient in the optimal transmission line increases, causing the electromagnetic power consumed by the transducer and, hence, the acoustic power emitted by its elements to drop.

It should be noted that the only valid approach to separating the effect of MEAT aperiodicity and the frequency properties of the piezoelectric elements from other factors affecting the AO interaction efficiency is to compare the performance of the transducers matched to the transmission line.

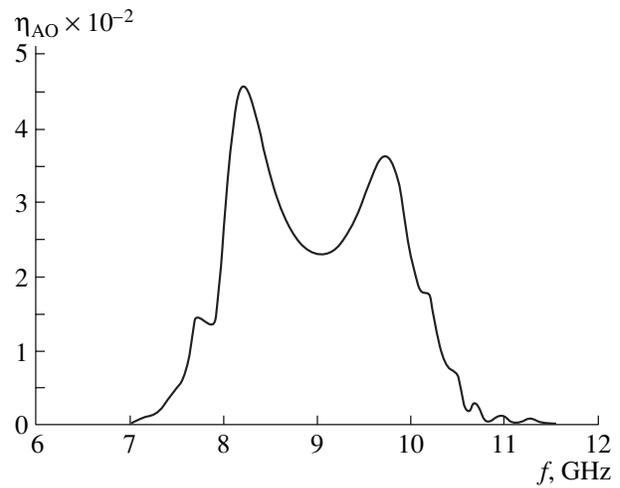


Fig. 10. AO interaction efficiency vs. frequency for an aperiodic MEAT with narrow-band piezoelectric elements at $M = 252$ and $Z_0 = 16.85 \Omega$.

If an inductance is series-connected to the MEAT to compensate for its capacitance, the coefficient η_{ac} increases almost to unity (assuming that loss is negligible). The resulting RLC circuit has a low unloaded Q factor: $Q_0 = \omega L_{add}/R_{\Sigma}$, where L_{add} is the resonant inductance and R_{Σ} is the MEAT total radiation resistance. As is known, at $VSWR = 1$, the loaded Q factor Q_L equals $Q_0/2$. Having calculated the MEAT impedance, we can estimate Q_L . In our case, it approaches 4 or 5. Thus, one might expect that a series inductance can provide the wide-band (2.25–1.8 GHz) matching of the transducer to the optimal transmission line at $f_0 =$

Multielement transducers at $f_0 = 9$ GHz: periodic and with gradually varying pitch

MEAT type	Number of elements, M	Additional inductance L , nH	Wave impedance Z_0 of the transmission line, Ω	VSWR	AO efficiency $\eta_{AO, \max}$, %
2-GHz-wide AO interaction band					
PW	132	0	8.67	~19	4.2
PW	132	0.15	0.9	1.0	18
APW	138	0	8.21	19.5	4.0
APW	132	0.15	1.07	1.0	16
APN	252	0	16.85	18	4.6
APN	354	0.21	1.34	1.0	18.4
0.7-GHz-wide AO interaction band					
PW	1002	0	1.14	~20	10.5
PW	1002	0.02	0.12	1.0	170*
APW	1700	0	0.67	19.5	8.5
APW	1700	0.012	0.07	1.0	43
APN	2802	0	1.52	17.5	21
APN	2802	0.03	0.22	1.0	100*

Note: For the abbreviations, see the text.

* Data obtained in the small-signal approximation.

9 GHz. Clearly, the resulting decrease in the AO efficiency at the bounds of the operating frequency range and the decrease in the AO interaction bandwidth can easily be avoided by changing the angle Θ_0 of light incidence.

The table summarizes $\eta_{\text{AO, max}}$ calculated for the periodic MEATs with the wide-band piezoelectric elements (PW) and for the aperiodic MEATs with the wide-band (APW) and narrow-band (APN) piezoelectric elements. The each of these MEATs, two ways to apply the electromagnetic power to the transducer are indicated. In the first case, the MEAT terminates the optimal transmission line. In the second one, the MEAT is series-connected to resonant inductance that completely compensates for the MEAT capacitance at a frequency f_0 . In the first case, the VSWR is within 17–20; in the second case, it approaches unity. The table also lists the resonant inductances and the respective wave impedances of the transmission lines for which the calculations were performed. In the matched variant, the wave impedance equals the sum of the MEAT radiation resistance and the loss resistance (the latter was taken to be 0.1Ω). The table compares the MEATs that provide AO interaction in the 2- and 0.7-GHz-wide frequency ranges.

Of the wide-band MEATs considered here, the periodic transducer, in both the matched and mismatched versions, seems to be the best. This design requires the least number of the elements, when providing almost the same AO efficiency as the other two. Also, the matched version of the periodic transducer offers the highest efficiency among the transducers that provide AO interaction in the 0.7-GHz-wide frequency range. The higher AO efficiency of the mismatched aperiodic MEAT with the narrow-band piezoelectric elements is presumably due to the difference in the VSWR, since the wide- and narrow-band piezoelectric elements have different impedances.

CONCLUSION

This paper theoretically analyzes the possibility of using aperiodic multielement piezoelectric transducers in Bragg acousto-optical devices with a weakly diverging light beam.

A sectionalized antiphase transducer with narrow-band piezoelectric elements is capable of significantly increasing the AO efficiency in a wide frequency band even if the periods in the sections are far from optimal. However, their $\eta_{\text{AO}}(f)$ curves exhibit oscillations, which may exceed 3 dB. Moreover, increasing the number of the elements in the sectionalized transducer decreases the magnitude of its electrical impedance, making its matching to the transmission line difficult. The situation is aggravated by the fact that the piezoelectric elements do not all operate at a given frequency and the idle elements introduce a reactance.

The analysis of the antiphase MEAT with the pitch gradually varying along the transducer by the law that is linearly related to the required frequency dependence of the period gave the following results. If the acoustic power is emitted by the elements simultaneously and uniformly throughout the frequency range, the transducer's aperiodicity only decreases the AO efficiency at a given bandwidth. As for the aperiodic MEAT with the narrow-band piezoelectric elements, at a given band of AO interaction, it requires a much (2–3 times) greater number of the elements than the periodic MEAT, the AO efficiency being almost the same. An increase in the number of the elements by an order of magnitude for a given frequency range of the elements decreases the AO interaction bandwidth. The resulting increase in the AO efficiency does not exceed the associated increase observed for the periodic MEAT with the same AO bandwidth. However, the number of the elements is much smaller.

REFERENCES

1. M. A. Grigor'ev, A. V. Tolstikov, and Yu. P. Navrotskaya, *Zh. Tekh. Fiz.* **71** (10), 70 (2001) [*Tech. Phys.* **46**, 1274 (2001)].
2. R. W. Damon, W. T. Maloney, and D. H. McMahon, in *Physical Acoustics: Principles and Methods*, Ed. by W. P. Mason and R. N. Thurston (Academic, New York, 1970; Mir, Moscow, 1974), Vol. VII.
3. H. G. Cohen and E. I. Gordon, *Bell Syst. Tech. J.* **44**, 693 (1965).

Translated by A. Khzmalyan

A High-Power Vircator Based on an Ironless Linear Induction Accelerator of Electrons

V. D. Selemir, A. E. Dubinov, B. G. Ptitsyn, A. A. Evseenko, V. A. Letyagin,
R. K. Nurgaliev, A. V. Sudovtsov, and A. V. Yachny

Sarov Institute of Experimental Physics, Sarov, Nizhegorodsk oblast, 607190 Russia

e-mail: dubinov@ntc.vniief.ru

Received December 4, 2000

Abstract—A high-power vircator that is based on the Korvet ironless linear induction electron accelerator built around radial pulse-forming lines is implemented for the first time. Results on the computer simulation and experimental optimization of the vircator are discussed. The experimentally found parameters of the vircator are the following: cathode current 35 kA (at a limit current of 19 kA), microwave pulse base widths 40 and 18 ns, and peak microwave power more than 500 MW. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

Microwave oscillators with a virtual cathode (VC), including vircators, form a basic class of oscillators in ultra-high-power relativistic high-current microwave electronics. They have been reviewed in [1]. Current vircators are based on high-voltage nanosecond oscillators that are built around single and dual pulse-forming lines [2, 3], inductive energy storage devices with plasma current interrupters [4, 5], or magnetic explosion generators [6, 7].

Ironless linear induction accelerators (LIAs) that exploit radial pulse-forming lines [8] are more promising for feeding the vircators than the above-mentioned sources. The advantage of such LIAs is that they can generate high-voltage pulses of both polarities and operate in the pulse-train mode. Also, in these LIAs, the cathode and the anode can be grounded simultaneously. This enables new types of vircators to be designed that cannot be implemented with conventional feeders, for example, a cyclotron-resonance vircator [8].

In Section 1 of this work, we for the first time describe a conventional vircator based on the Korvet ironless LIA. In Section 2, we report the results on the computer simulation of this vircator. Section 3 is devoted to measuring the microwave characteristics of the vircator and its optimization.

1. KORVET LIA AND VIRCATOR ON ITS BASIS

The Korvet ironless LIA, earlier used as an injector for the well-known LIA-10 accelerator [9], consists of four modules, each representing three functionally coupled units: an inductor unit, a GIN-500 pulse voltage generator that charges the capacitor of the inductor unit, and a generator that forms pulses triggering the

switches of the inductors (switch-triggering generator, STG). The modules have their own designs and electrical circuits. The inductor unit, in its turn, comprises three series-connected sections built on radial lines that are insulated by high-resistivity deionized water. The lines are parallel-connected to one GIN-500 through the charge circuit. A separate switch is made by the torus-shaped grounded body, which has a break in its inner diameter with a circular high-voltage electrode inside. Along the perimeter of one of the exit gaps, the radial line of the inductor is closed on its inner radius by means of a multichannel circular switch. The other exit gap is connected to a load through the same switch. The multichannel switch is formed by ten separate gas-filled 500-kV gaps of the trigatron type. The gap operation time spread is no more than 2 ns. Near the axis of the inductor unit, a polyethylene accelerating tube separating the vacuum space of the accelerating path from the deionized-water-filled inductor space is placed. The inner diameter of the accelerating tube is 380 mm.

Ideally, the unloaded inductor with uniform lines generates rectangular pulses of the accelerating voltage whose amplitude is close to the value of the charging voltage. The base width of the first pulse was 20 ns. The second and third pulses with alternating polarity had 40-ns-wide bases. The parameters of the inductor unit were as follows: voltage 500 kV, charging pulse with 520 ns, peak amplitude of the accelerating voltage 1.5 MV, and short-circuit current 180 kA.

The capacitor of the inductor unit was charged by the modified Arkad'ev–Marx GIN-500 generator [10]. Five cascades of the GIN-500 included IK-100-0.25 capacitors, gas-filled 100-kV trigatron switches, liquid resistors, and conductive lines. Transformer oil was used as an insulator. The time spread of GIN switching-on was about 3 ns at a constant voltage of 100 kV. The

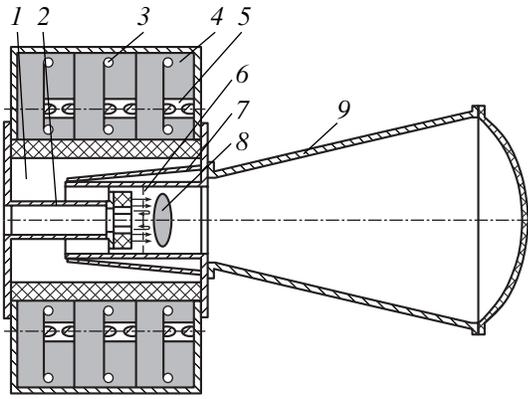


Fig. 1. Vircator based on one module of the Korvet LIA. 1, vacuum space of the accelerator; 2, cathode; 3, high-voltage electrode of the inductor unit; 4, deionized water; 5, gas-filled controllably gap; 6, anode grid; 7, anode; 8, virtual cathode; 9, microwave horn.

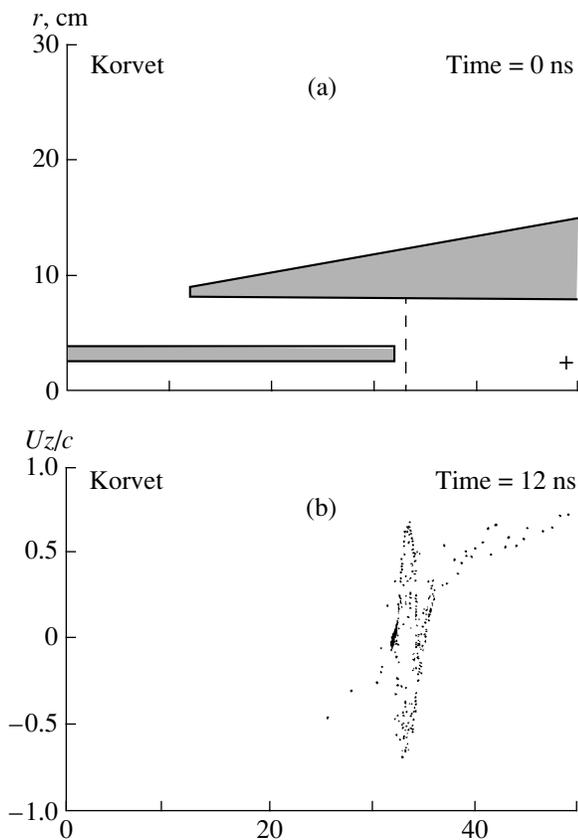


Fig. 2. Computer simulation of the vircator: (a) geometry of the region stimulated and (b) instantaneous phase portrait of the beam.

dielectric strength safety margin of the gas-filled gap was 2. The electrical loop of the GIN was screened by its metallic body, which was connected to the inductor unit by the 500-kV coaxial lead provided with the filter to suppress pulses that run back through the lead upon

switching the inductor wires. The energy content of one GIN-500 is 6.25 kJ; pulse current, 50 kA; and charging voltage, 100 kV.

An STG consists of a dual pulse-forming line, a controllable gap, a sharpener, and a charging capacitor. An external electrode acts also as the case of the dual forming line.

The vircator was built around one module of the Korvet LIA (Fig. 1). The device integrates coaxial cathode and anode, the latter enclosing the former. The cathode represents the cathode holder (thin-walled (0.2 mm) stainless steel tube), to which a steel wire was welded on the anode side. Graphite cylinders 20 mm in diameter and 40 mm long acting as emitters were screwed on the substrate. A total of 8, 10, or 12 closely spaced emitters were arranged on the substrate, producing a cathode with an outer diameter of 72, 92, or 112 mm, respectively.

The anode was a 380-mm-long hollow cylinder with a diameter of 160 mm. A metallic ring was inserted in the cylinder, and a square-mesh (3 mm on a side) anode grid made of nichrome wire (0.3 mm in diameter) was stretched over it by electric arc welding. Moving the ring with the anode grid along the axis of the vircator, one can vary the anode-cathode spacing. In this work, this spacing was varied between 8 and 13 mm.

The anode was terminated by a conical horn antenna with an opening angle of the emitting horn of 10° . The diameter of the exit window that separates the vacuum space of the vircator from the environment was 700 mm. The window was made of sheet organic glass. The residual pressure in the vircator was kept at $(3-5) \times 10^{-5}$ torr.

With the inductor unit charged to the maximal extent, the voltage extracted from the diode of the vircator was 900 kV. In this case, the diode current was 50 kV, which is several times higher than the limit beam current in the anode plane behind the grid (estimated at 15–20 kA).

The currents passing in the vircator were measured with three Rogowski loops, as shown in Fig. 1.

The radiated microwave energy per pulse was measured with a wide-band calorimeter. The radiated microwave power was determined with hot-carrier semiconductor detectors [1].

2. COMPUTER SIMULATION OF THE VIRicator

The computer simulation of the vircator was performed with the software suite based on the Karat 2.5-dimensional PIC code [12]. The geometry of a region being simulated and its dimensions are shown in Fig. 2a: a tubular cathode with outer and inner diameters of 72 and 32 mm, respectively; an anode with a diameter of 160 mm; and the anode-cathode spacing varying from 7 to 18 mm. Such a geometry is roughly identical to the vircator that was based on the Korvet

LIA and studied in this work. It was assumed that a 900-kV 40-ns pulse is applied to the diode, with the pulse shape meeting the previously recorded waveform. Under these conditions, the diode current was as high as ≈ 35 kA.

Figure 2b shows a typical instantaneous phase portrait of a particle ensemble. It is seen that a virtual cathode (VC) is formed in the system and its oscillations generate microwave radiation. Electrons leaving the diode are reflected from the VC space charge and oscillate in the potential well produced by the cathode, grid, and VC. The simulation shows that the appearance of the VC coincides with the generation of microwave radiation and the time instant the VC collapses, with the instant the generation ceases.

Our goal was to compare the results of simulation with experimental data. We believed that, in the case of good agreement between the simulation and experiment, we could optimize the parameters of the vircator, for example, increase the extracted power or obtain the desired frequency.

Figure 3 shows a typical simulated waveform of the extracted power for a cathode–grid spacing of 12 mm. The maximum of the peak power is observed at a spacing between 12 and 14 mm, in agreement with experimental data (see below). The values of the peak power, 400–500 MW, are also consistent with experimental findings.

Thus, the simulation of the vircator based on the Korvet LIA suggests that the device can generate high-power microwave pulses. The agreement between the analytical and experimental results gives promise that the given model can be used to tentatively optimize the system.

3. EXPERIMENTAL STUDY OF THE VIRCATOR

The goal of our early experiments was to find the cathode that provides the maximal energy of the microwave radiation (among three available cathodes consisting of 8, 10, and 12 emitters). The maximal value of the microwave energy detected by the calorimeter was found for the eight-emitter cathode with a diameter of 72 mm.

The next series of experiments was aimed at determining the optimal spacing d between the end face of the cathode and the grid. The spacing was varied between 6 and 14 mm with a step of 1 mm. No radiation was observed for spacing of 6 and 14 mm. For the other spacings, the energy W per microwave pulse measured with the calorimeter is listed in Table 1 (the average energy of three pulses is given for each of the spacings). It follows from Table 1 that the curve $W(d)$ has a narrow peak, which agrees with experimental data obtained elsewhere (see, e.g., [13]).

To determine the peak power of a microwave pulse, we recorded pulse envelopes, using the semiconductor

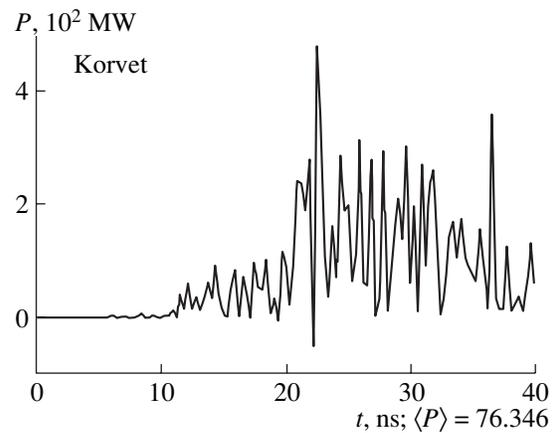


Fig. 3. Stimulated waveform of the microwave power.

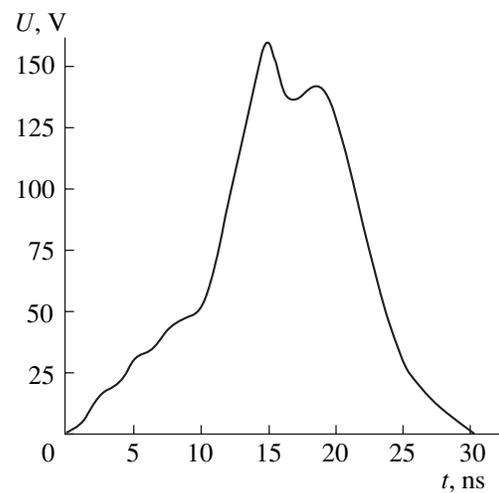


Fig. 4. Experimental envelope of the microwave signals coming from semiconductor detector (short cathode).

detectors, and compared them with the calorimetric data. A typical envelope obtained for the spacing $d = 10$ mm is shown in Fig. 4. The peak microwave power obtained at the optimal spacing $d = 10$ mm was found to be roughly 150 MW.

Table 1. Results of experiments on varying the cathode–grid spacing

d , mm	W , J
7	0.63
8	0.96
9	1.21
10	1.37
11	1.15
12	0.69
13	0.09

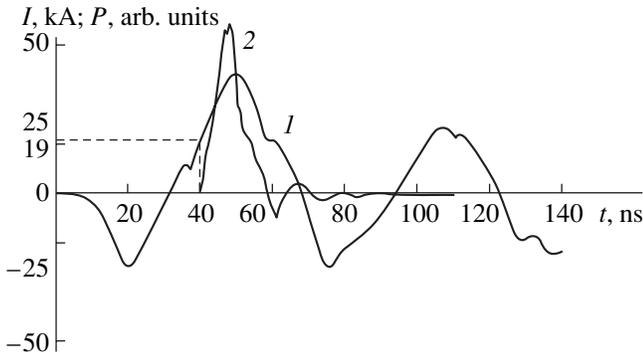


Fig. 5. Synchronized waveforms of the (1) cathode current and (2) microwave pulse for the longer cathode.

In another series of experiments, we tried to reduce losses in the waveguide slots where the Rogowski loops were inserted. The cathode length was increased to 130 mm, all other design features remained the same. As before, we calorimetrically measured the microwave energy, varying the cathode-grid spacing (the average energy of three pulses for each spacing) from 7 to 19 mm. The results are listed in Table 2, where the values of the peak power are also given. Note that the microwave pulses become shorter nearly twice: 20 ns (against 40 ns for the shorter cathode).

From Table 2, it is seen that both the energy absorbed by the calorimeter and the peak power of the pulses are maximal when the cathode-grid spacing equal 12 mm. In this case, the peak power reaches 520 MW.

Table 2. Results of experiments on varying the cathode-grid spacing

d , mm	W , J	P , MW
7	0.92	180
8	1.45	285
9	1.58	310
10	2.16	425
11	2.50	490
12	2.65	520
13	2.48	485
14	2.44	480
15	2.32	455
16	1.79	350
17	1.64	320
18	1.30	255
19	0.73	145

The data obtained lend optimism to the possibility of vircator optimization in terms of the radiation output. One can further extend the cathode and place the anode grid at the very edge of the anode cylinder so as to form the VC immediately in the horn. Here, however, there is a limitation: the extension of the cathode electrode will lead to a considerable increase in the diode inductance. This may decrease the diode current down to the lower limit.

Finally, one more series of experiments were targeted at determining the starting current of microwave generation. It is generally accepted that this current coincides with the ultimate beam current in the drift region of the vircator. To determine the starting current, we synchronized the signals coming from the cathode Rogowski loop and from the semiconductor detectors with regard for the time it takes for the microwave radiation to propagate from the VC to the detector. Superposing the signals, we found that the starting current of microwave generation is $I = 19$ kA (Fig. 5), which is approximately equal to the flight current.

CONCLUSION

Thus, we implemented a high-power vircator based on the ironless LIA using radial pulse-forming lines and studied its performance.

The parameters of the vircator are as follows: cathode current 35 kA, starting (limit) current 19 kA, base width of microwave pulses 18 ns, and peak power more than 500 MW.

The radiation was extracted into the environment. Power measurements were performed with a microwave calorimeter and hot-carrier semiconductor detectors. The results of simulation using the Karat code are in agreement with experimental data.

ACKNOWLEDGMENTS

The authors thank V.P. Tarakanov for the assistance in the simulation.

REFERENCES

1. A. E. Dubinov and V. D. Selemir, *Zarubezhn. Radioelektron.*, No. 4, 54 (1995).
2. R. D. Scarpetti and S. C. Burkhart, *IEEE Trans. Plasma Sci.* **13** (6), 506 (1985).
3. V. D. Selemir, B. A. Alekhin, V. E. Vatrugin, *et al.*, *Fiz. Plazmy* **20**, 689 (1994) [*Plasma Phys. Rep.* **20**, 621 (1994)].
4. A. G. Zherlitsyn, V. S. Lopatin, and O. V. Luk'yanov, *Pis'ma Zh. Tekh. Fiz.* **16** (11), 69 (1990) [*Sov. Tech. Phys. Lett.* **16**, 431 (1990)].
5. V. S. Zhdanov, A. L. Babkin, S. M. Galkin, *et al.*, in *Proceedings of XXV Zvenigorod Conference on Plasma*

- Physics and Controlled Thermonuclear Fusion, Zvenigorod, 1998*, p. 169.
6. E. I. Azarkevich, A. N. Didenko, P. V. Dolgoplov, *et al.*, Dokl. Akad. Nauk SSSR **319**, 352 (1991) [Sov. Phys. Dokl. **36**, 539 (1991)].
 7. E. I. Azarkevich, A. N. Didenko, A. G. Zherlitsyn, *et al.*, Teplofiz. Vys. Temp. **32**, 127 (1994).
 8. A. I. Pavlovskii, V. S. Bosamykin, V. D. Selemir, *et al.*, in *Relativistic High-Frequency Electronics* (Inst. Prikl. Fiz. Ross. Akad. Nauk, Gorki, 1992), No. 7, p. 81.
 9. A. I. Pavlovskii, V. S. Bosamykin, V. A. Savchenko, *et al.*, Dokl. Akad. Nauk SSSR **250**, 1117 (1980) [Sov. Phys. Dokl. **25**, 120 (1980)].
 10. A. I. Gerasimov, A. S. Fedotkin, D. I. Zenkov, and S. T. Nazarenko, Prib. Tekh. Éksp., No. 1, 96 (1998).
 11. M. D. Raizer and L. É. Tsopp, Radiotekh. Élektron. (Moscow) **20**, 1691 (1975).
 12. V. P. Taranov, *User's Manual for Code Karat* (Berkley Research Associate Inc., Springfield, 1992).
 13. M. Yatsuzuka, M. Sato, I. Ohta, *et al.*, *Annual Review* (Inst. of Plasma Physics, Nagoya University, Nagoya, 1990), p. 96.

Translated by V. Isaakyan

Chaotic Generation in a Megawatt Backward-Wave Tube

N. S. Ginzburg, N. I. Zaitsev, E. V. Ilyakov, I. S. Kulagin,
Yu. V. Novozhilova, R. M. Rozental', and A. S. Sergeev

*Institute of Applied Physics, Russian Academy of Sciences,
ul. Ul'yanova 46, Nizhni Novgorod, 603600 Russia*

Received March 1, 2001

Abstract—Experimental and theoretical studies on the self-modulation mode of generation in a high-power BWT with the electrodynamic system representing a slightly corrugated waveguide are presented. The BWT is fed by an electron beam with the energy 150 keV and the current 150 A. The system operates at the E_{01} mode with the mean frequency 8.7 GHz. Dynamic chaos is obtained by a three-fold increase in the length of the interaction space in comparison with the prototype exhibiting stationary generation. The stationary generation was changed to periodic sinusoidal self-modulation and then to chaotic self-modulation as the current increases from 6 to 60 A. The generation mode is simplified when the current ranges from 70 to 90 A and becomes complicated again for the current exceeding 100 A. Experimental observations are in good agreement with the results of simulation predicting a certain simplification of the self-modulation mode at the currents 70–90 A owing to the effect of high-frequency space charge. Under the conditions of chaotic generation, the mean power was as high as 2 MW at the relative spectral width of the signal 4% and the total duration of the microwave pulse 10 μ s. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

The nonlinear dynamics of backward-wave tubes (BWTs) has attracted much attention [1–6], since these devices exemplify a distributed self-oscillation system. Ginzburg *et al.* [7, 8] have demonstrated that complicated (including chaotic) modes of generation in such systems are a possibility when the values of the supercriticality parameter are large. Since a BWT is a conventional device of high-power relativistic electronics, studies of the multifrequency processes in oscillators of this type can be of practical interest for the generation of high-power periodically modulated or noise-like signals. To date, however, self-modulation generation has been observed only in the experiments with special breadboard models of weakly relativistic milliwatt BWTs [9–11]. Recently, the team at the Institute of Applied Physics, Russian Academy of Sciences, has begun to study self-modulation generation in high-power BWTs using a Saturn microsecond accelerator [12]. In these experiments, as an experimental model, we use a BWT of design that is typically used to generate relativistic electrons: a slightly corrugated waveguide with an evanescent taper at the cathodic end where the radiation is extracted at the collector [13].

Self-modulation generation in a high-power BWT was first observed in [1], where the length of the working space was roughly one and a half times larger than the length corresponding to stationary generation. The problem of mode selection by transverse index was solved by taking the lower H_{11} mode as the working one. This made it possible to avoid the synchronous interaction of the electronic beam with other transverse

modes. However, in this case, the working point is in the range of small group velocities ($0.2c$). On the one hand, this provides an increased coupling impedance and favors self-modulation generation; on the other hand, the extraction of the microwave energy from the working space becomes difficult and the risk of reflections from the output of the working space grows. Because of these features, the microwave fields inside the interaction space were high and the output power of self-modulation generation in the BWT was limited at a level of 50–100 kW (for a microsecond pulse duration) by microwave breakdowns [1].

The purpose of our further studies was to obtain self-modulation generation at megawatt levels of the mean output power. The excitation of the E_{01} mode strongly coupled with the electron beam provided the excess of the working current over the starting value that was sufficient to trigger self-modulation generation. Also, the relatively high group velocity of the working mode ($0.4c$) favored the extraction of the radiation from the working space and provided a high output power in the absence of breakdowns. Note that Ilyakov *et al.* [14] studied stationary generation at the E_{01} working mode in a microsecond BWT excited in the Saturn accelerator. The authors of [14] considered the mechanisms limiting the duration of microsecond pulses and applied several techniques (including the use of oxygen-free copper for manufacturing the electrodynamic system) that made it possible to increase the output power to 5 MW at a pulse duration of up to 10 μ s.

Note that the operation at the E_{01} mode poses the problem of mode selection by transverse index. Indeed, the periodic and particularly chaotic self-modulation modes require that the injection currents substantially (ten or more times) exceed the starting one. In the general case, the starting conditions are satisfied for a number of transverse modes other than the working mode. Theoretical analysis of the competing modes with different transverse indices showed the possibility of self-modulation generation at one of the modes if the other are nonlinearly suppressed [15].

In the early BWT model with the E_{01} working mode, where the length of the interaction space was one and a half times larger than the space of stationary generation, deep periodic self-modulation was observed [2]. Under the self-modulation conditions, the radiation power (0.5 MW) was limited by microwave breakdown in the slow-wave structure produced by the electrochemical method. To provide the chaotic generation mode, the length of the slow-wave structure was further increased so that the total length of the interaction space became three times larger than the length of the stationary BWT. In addition the BWT electrodynamic system was made of oxygen-free copper by the special technique [14] to improve the electric strength. Under these conditions, we succeeded in observing both periodic and chaotic self-modulation by varying the injection current. The mean power of the output radiation was no less than 2 MW at a pulse duration to 10 μ s.

This work is devoted to the theoretical and experimental study of self-modulation processes at the E_{01} mode in a high-power BWT. In Section 1, we simulate the nonstationary processes in the model BWT, using the equations for a slowly varying field amplitude and the relativistic equations of motion of electrons. BWT models with a single transverse mode synchronous with the electron beam and those taking into account the excitation and competition of several modes were analyzed. In Section 2, we present the experimental results on the observation of self-modulation (including chaotic) generation in a high-power BWT. These results are in good agreement with the theoretical prediction.

1. NUMERICAL SIMULATION OF NONSTATIONARY PROCESSES IN BWT

A Model of BWT the Slow-Wave Structure of Which Has a Single Mode Synchronous with the Electron Beam

Assuming that the electron beam excites a single mode in the corrugated waveguide, we can describe the multifrequency processes in a BWT by three self-consistent equations: the equation for the synchronous wave amplitude [7, 8],

$$\frac{\partial A}{\partial \tau} - \frac{\partial A}{\partial \eta} = -\frac{J}{\pi} \int_0^{2\pi} e^{-i\vartheta} d\vartheta_0, \quad (1)$$

and the relativistic equations of electron motion taking into account the effect of the beam space charge field:

$$\frac{\partial \vartheta}{d\zeta} = \frac{1}{\sqrt{1-\gamma^{-2}}} - \frac{1}{\sqrt{1-\gamma_0^{-2}}}, \quad (2)$$

$$\frac{\partial}{\partial \zeta} = \text{Re} \left(A(\zeta, \tau) \exp(i\vartheta) + i\sigma \sum_n f_n \rho_n \exp(in\vartheta) \right) \quad (3)$$

with the boundary and initial conditions given by

$$\begin{aligned} A|_{\zeta=l} &= 0, & A|_{\tau=0} &= A_0(\zeta), \\ \vartheta|_{\zeta=0} &= \vartheta_0 \in (0, 2\pi), & \gamma|_{\zeta=0} &= \gamma_0. \end{aligned} \quad (4)$$

Here, $A = eE_z/(mc\omega)$ is the dimensionless amplitude of the synchronous (-1) st harmonic of the radiation field; $\tau = \omega(t - z/v_0)/(1/\beta_0 + 1/\beta_{gr})$ and $\zeta = \omega z/c$ are the dimensionless time and longitudinal coordinate, respectively; l is the dimensionless length of the interaction space, $\vartheta = \omega t - hz$ is the phase of the electrons relative to the synchronous harmonic of the field; $J = eI/Z/(2\beta_0^2 mc^2)$; Z is the coupling impedance for this harmonic [16]; I is the beam current; γ is the relativistic mass factor of the electrons; ω is the synchronism frequency; $v_0 = \beta_0 c$ is the initial translation velocity of the electrons, $v_{gr} = \beta_{gr} c$ is the group velocity of the wave; $\sigma = Ieg/(mc\omega^2 b)$ is the space charge parameter; f_n is the reduction coefficient for the n th harmonic of the space charge field; and $\rho_n = (1/\pi) \int_0^{2\pi} \exp(-n\vartheta) d\vartheta_0$ is the n th harmonic of the space charge density. The reduction coefficient f_n of the space charge was varied so as to provide the best agreement with the experimental results. The optimal value of f_n was found to be $f_{1,2} \approx 0.55$, which is about two times smaller than the values estimated from the formula $f_n = 1 - \exp(-2ngd)$, where d is the distance between the beam and the waveguide wall and $g = \omega/(c\beta_0\gamma)$ is the transverse wave number. We assume that the waveguide is smooth and that the distance d is substantially smaller than the mean radius of the waveguide, so that one can neglect the curvature of the waveguide wall and of the electron beam.

The numerical simulation of Eqs. (1)–(3) was carried out for various values of the beam current for tube parameters close to those used in experiments. The length of the interaction space was 62.3 cm, the mean radius of the waveguide, 1.38 cm; the corrugation period, 1.73 cm; the radius of the electron beam, 0.67 cm; and the energy of the electrons, 150 keV. The coupling impedance and the starting current were calculated for the given geometry of the electrodynamic system using the results in [16]. For the E_{01} mode, the starting current was 6 A at the working frequency 8.7 GHz and the coupling impedance 0.5 Ω . The simulation included the first two harmonics of the space

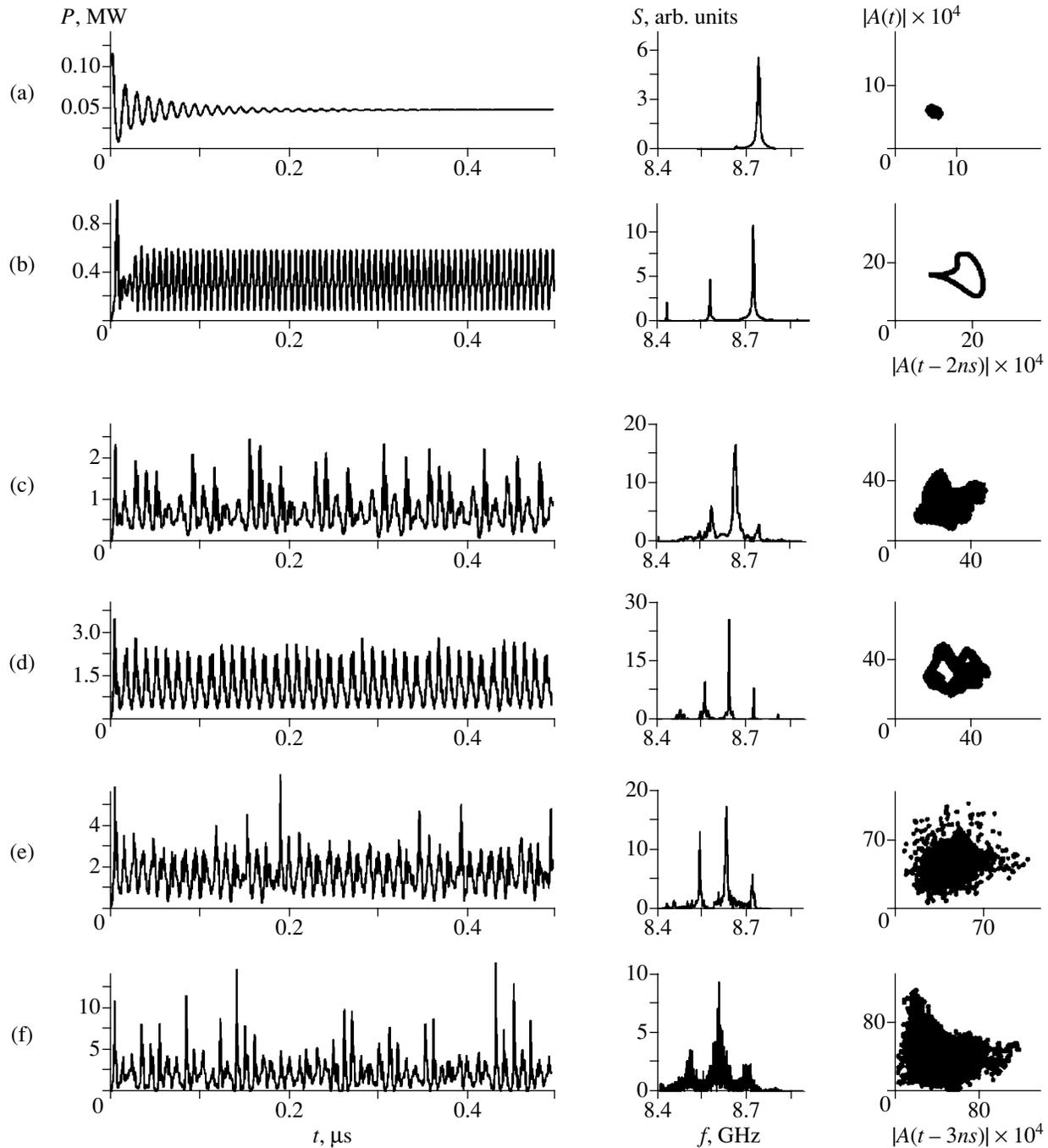


Fig. 1. Simulation of the nonstationary processes in a BWT with the coupling impedance 0.47Ω and the particle energy $\gamma_0 = 1.3$. The plots of the radiation power versus time (left column), the spectra of the output radiation (middle column), and the phase portraits (right column) at the injection currents (a) 7, (b) 30, (c) 55, (d) 70, (e) 90, and (f) 120 A.

charge field, since the allowance for the next harmonics led to insignificant corrections.

Figure 1 shows the time dependences of the output power, the output signal spectra, and the phase portraits for various values of the beam current. To construct the phase portraits, we took the time dependence of the output signal amplitude $|A(0, t)|$ and plot points corresponding to the states of the system at various time

instants on the plane $(|A(0, t)|, |A(0, t - t_{\text{del}})|)$. Here, t_{del} is the delay time, which was set approximately equal to a quarter of the self-modulation period. Figure 1a shows the setting of stationary generation at $I = 7$ A. The increase in the current complicates the self-modulation mode, as well as the spectra and the phase portraits. In particular, self-modulation is periodic at the current 30 A (Fig. 1b) but loses periodicity and

becomes chaotic at the current 55 A (Fig. 1c). The comparison of the corresponding spectra and phase portraits confirms the transition to chaotic generation. In the chaotic mode, the width of the spectrum is about 3% and the mean efficiency is 10%. Note that the further increase in the current with the coupling impedance remaining unchanged suppresses self-modulation owing to the space charge effect (cf. with [8, 9]). Specifically, modulation becomes sinusoidal again at the current 70 A (Fig. 1d). However, as the injection current grows still further at a given reduction coefficient for the space charge parameter, the complicated character of self-modulation returns (Figs. 1e, 1f). The generation mode is quasi-period at the current 90 A (Fig. 1e) and chaotic at the current 120 A (Fig. 1f). At the current 30 A, the self-modulation period is 7 ns or 90 dimensionless units (Fig. 1b). When the current exceeds 70 A and the system passes through the region of chaotic modulation, the period of self-modulation almost doubles (13 ns or 170 dimensionless units) (Fig. 1d). As the characteristic time scale of self-modulation grows, the delay time used in constructing the phase portraits also increases from 2 ns in Figs. 1a–1c to 3 ns in Figs. 1d–1f. Figure 2 shows the autocorrelation function for the injection current 120 A given by

$$K(t_{\text{del}}) = \frac{1}{T} \int_0^T (|A(0, t')| - |\bar{A}|) \times (|A(0, t' + t_{\text{del}})| - |\bar{A}|) dt', \quad (5)$$

where T is the realization time and $|\bar{A}|$ is the modulus of the field amplitude averaged over the realization time [17].

The typical decay time of the autocorrelation function for the chaotic mode at the given current is 50 ns, which corresponds to the width of the fundamental band in the spectrum shown in Fig. 1f (0.02 GHz).

Simulation of Nonlinear Mode Competition

As follows from the dispersion relations for the electrodynamic system with the above parameters, synchronous interaction with three lower waveguide modes, H_{11} , E_{01} , and H_{21} , may occur at a working voltage of 150 kV (Fig. 3). For the H_{11} and E_{01} modes, the points of synchronism correspond to the excitation of the backward waves, whereas the frequency of synchronism for the H_{21} mode is close to the cutoff frequency. The starting currents for the H_{11} and E_{01} modes calculated according to [16] are 6 and 27 A, respectively. The calculation of the starting current for the H_{21} mode within the theory of orotron with a variable longitudinal field structure [18] yielded the value 13 A. Thus, the E_{01} mode exhibits the larger coupling impedance and the smaller starting current in comparison with the other modes.

In the experiments, we observed the excitation of only one of the working modes (E_{01}) with the center fre-

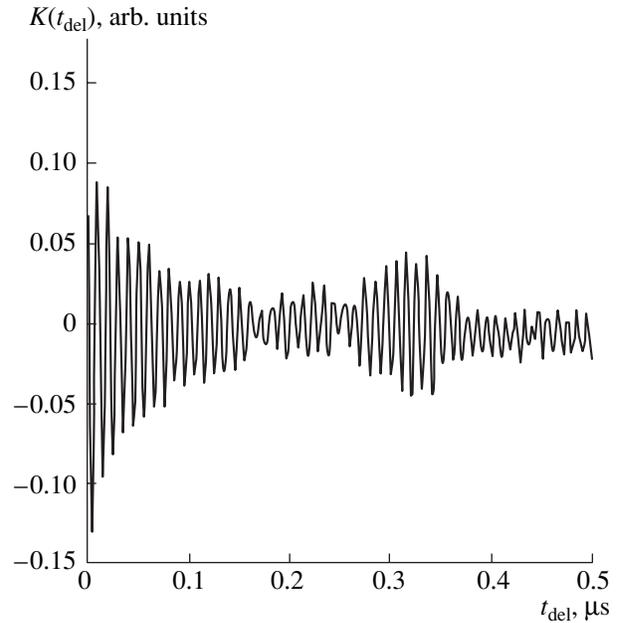


Fig. 2. Autocorrelation function calculated from the results of numerical simulation for the current 120 A.

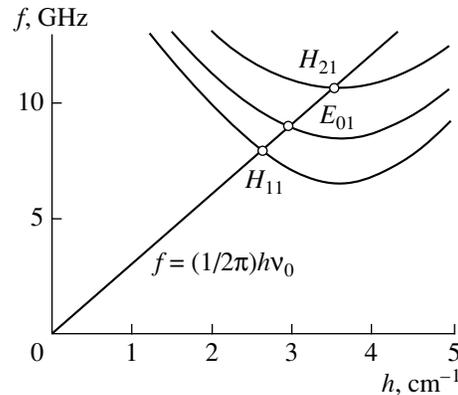


Fig. 3. Dispersion diagram of the BWT with the slow-wave structure parameters close to the experimental values.

quency 8.7 GHz. Nevertheless, the self-excitation conditions were also satisfied for the other modes, since the working current was increased to 120 A in order to provide chaotic generation. The absence of the components related to the excitation of the other modes in the spectrum and in the radiation pattern can be explained by the nonlinear competition of modes. Indeed, the results of simulation (see below and also [15]) show that the nonlinear competition of the modes at substantially differing coupling impedances may lead to one-mode self-modulation generation even if the self-modulation thresholds are overcome for each of the modes.

Let us demonstrate the suppression of the parasitic mode in the competition of two transverse modes: E_{01} and H_{11} . With neglect of the high-frequency field of the

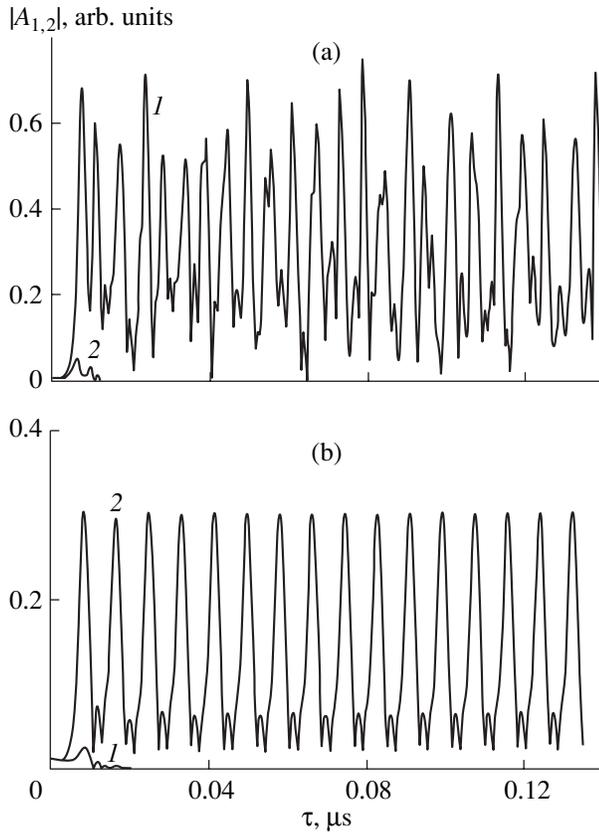


Fig. 4. Time dependences of the magnitudes of the (1) E_{01} and (2) H_{11} wave amplitudes under the conditions of mode competition at the current 120 A. (a) The coupling impedances of the waves are close to the real values: $|Z_1| = 0.5 \Omega$ and $|Z_2| = 0.15 \Omega$ ($p_1 = 0.75$, $p_2 = 0.09$, $v = 0.9$, and $J = 10^{-4}$). (b) The coupling impedance of the E_{01} wave is decreased ten times.

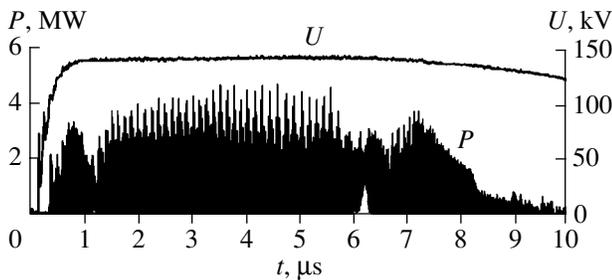


Fig. 5. Complete oscillogram of the microwave signal at the current 120 A.

space charge, mode interaction in a relativistic BWT is described by the system of equations

$$\frac{\partial A_1}{\partial \tau} - \frac{\partial A_1}{\partial \zeta} = -\frac{\bar{J}}{2\pi^2} \int_0^{2\pi} \int_0^{2\pi} e^{-i\vartheta_1} d\vartheta_{10} d\vartheta_{20},$$

$$p_1 \frac{\partial A_2}{\partial \tau} - \frac{\partial A_2}{\partial \zeta} = -\frac{\bar{J} p_2}{2\pi^2} \int_0^{2\pi} \int_0^{2\pi} e^{-i\vartheta_2} d\vartheta_{10} d\vartheta_{20},$$

$$\frac{d\vartheta_1}{d\zeta} = \frac{1}{\sqrt{1-\gamma^{-2}}} - \frac{1}{\sqrt{1-\gamma_0^{-2}}}, \quad (6)$$

$$\frac{\partial \vartheta_2}{\partial \zeta} = v \left(\frac{1}{\sqrt{1-\gamma^{-2}}} - \frac{1}{\sqrt{1-\gamma_0^{-2}}} \right),$$

$$\frac{\partial \gamma}{\partial \zeta} = \text{Re}(A_1 \exp(i\vartheta_1) + v A_2 \exp(i\vartheta_2)).$$

Here, we use the normalized variables: $\zeta = \omega_1 z/c$ is the longitudinal coordinate, $\tau = \omega_1(t - z/v_0)/(1/\beta_0 + 1/\beta_1)$ is time, $\vartheta_{1,2} = \omega_{1,2}t - h_{1,2}z$ are the phases of the electrons relative to each of the waves, $A_{1,2} = eE_{1,2}/(mc\omega_{1,2})$ are the amplitudes of the waves, $v_{1,2} = \beta_{1,2}c$ are the group velocities of the waves, $\bar{J} = eI|Z_1|/(2mc^2\beta_0^2)$, $z_{1,2}$ are the coupling impedances of the waves, $p_1 = (\beta_0^{-1} + \beta_2^{-1})/(\beta_0^{-1} + \beta_1^{-1})$, $p_2 = |Z_2|\omega_2/(|Z_1|\omega_1)$, and $v = \omega_2/\omega_1$. The boundary and initial conditions for system (6) are given by

$$A_{1,2}|_{\zeta=l} = 0, \quad A_{1,2}|_{\tau=0} = A_0(\zeta), \quad (7)$$

$$\vartheta_{1,2}|_{\zeta=0} = \vartheta_{1_0, 2_0} \in (0, 2\pi), \quad \gamma|_{\zeta=0} = \gamma_0.$$

We performed the numerical simulation of Eqs. (6) for a BWT with the parameters close to those used in the experiments. Figure 4 shows the results of simulation for the injection current 120 A. As follows from Fig. 4a, the H_{11} mode is suppressed by the nonlinear competition with the E_{01} mode, having the larger coupling impedance and the smaller starting current. With the given parameters, chaotic generation is observed for the E_{01} mode. A substantial decrease in the coupling impedance Z_1 of the working mode E_{01} leads to complicated self-modulation generation at the H_{11} mode if the E_{01} mode is not excited (Fig. 4b). Thus, the simulation confirms that one-mode self-modulation generation at the working mode E_{01} is a possibility if the parasitic mode is nonlinearly suppressed. Note that, generally speaking, the injection current for the latter mode also far exceeds the starting current.

2. EXPERIMENTAL STUDY OF SELF-MODULATION GENERATION

Electron beams with a pulse duration to 10 μs were produced by the Saturn pulsed accelerator based on a hot-cathode magnetron injector gun. Under the working conditions, the accelerating voltage was as high as 150 kV and the injection current was varied from 10 to 140 A. The hot cathode worked in the space-charge-limited-current mode. The current was controlled by redistributing the potentials among the three electrodes of the electron gun.

The electrodynamic system represented a slightly corrugated axisymmetric waveguide with the param-

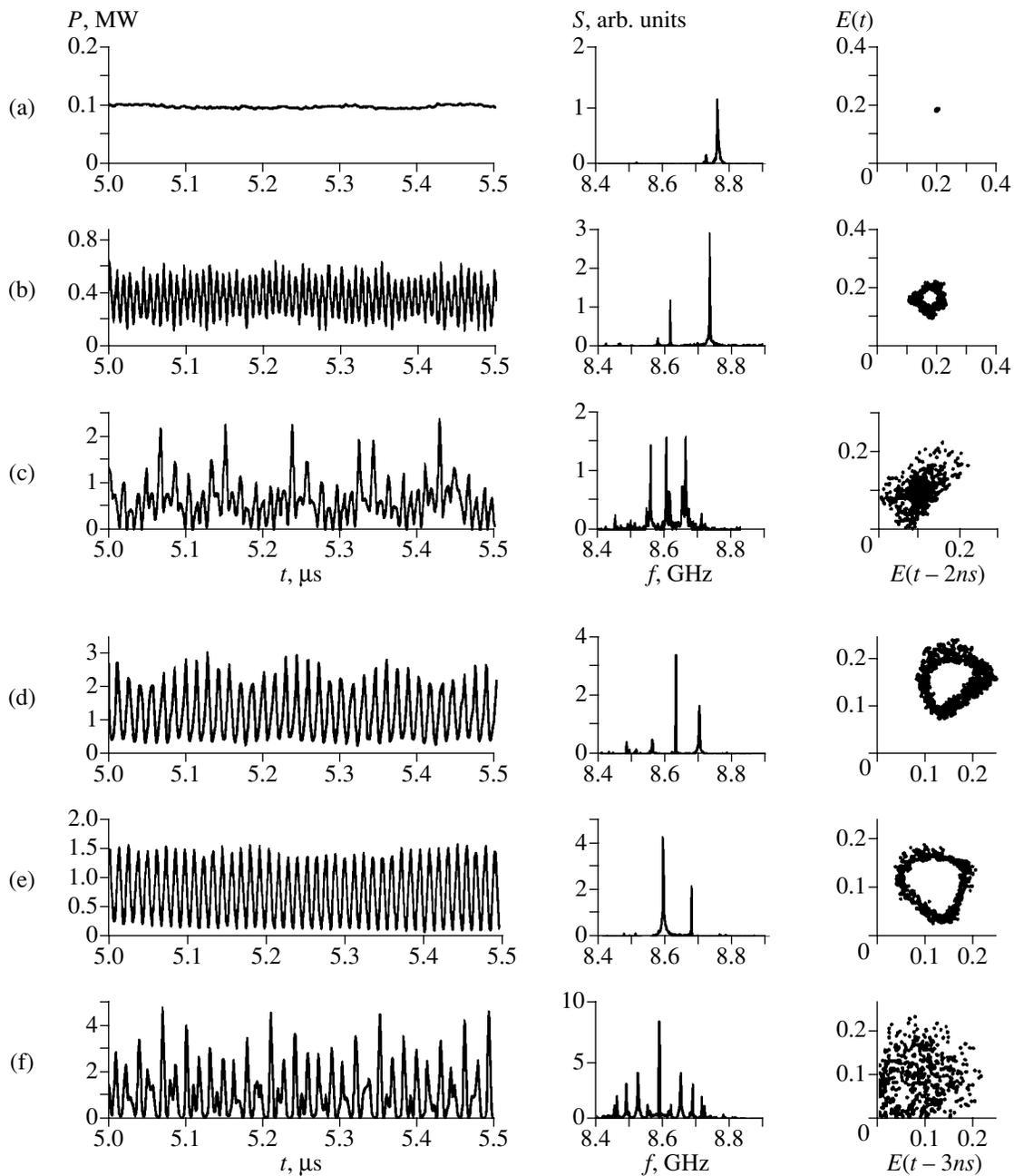


Fig. 6. Experimental oscillograms of the output power (left column), radiation spectra (middle column), and phase portraits (right column) at the injection currents (a) 10, (b) 30, (c) 55, (d) 70, (e) 90, and (f) 120 A.

ters mentioned in Section 1. To increase the breakdown strength [14], we maintained the temperature of the electrodynamic system at 500°C .

After passing the vacuum-tight window, the radiation was detected in an anechoic chamber. The receiving horn was placed at the maximum of the radiation pattern for the working mode E_{01} . The envelope of the output signal was recorded by a crystal detector with a response time of no more than 2 ns and by Hewlett-Packard and Tektronix digital oscilloscopes with a

bandwidth of 500 MHz and a digitization period of no more than 1 ns. Thus, the entire envelope of a high-frequency pulse could be memorized for the subsequent examination of its fragments. The heterodyne method was used to take the spectrum of the output radiation. The oscilloscope also recorded the difference signal during the microwave pulse, and then the spectrum of the signal was reconstructed. The output power was measured both from the calibrated damping in the channel and from the energy of the microwave pulse for more accurate measurements. In the latter case, a spe-

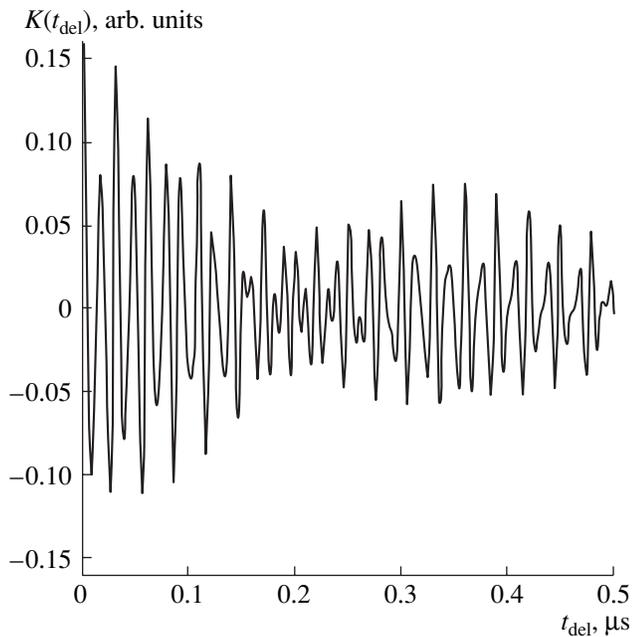


Fig. 7. Autocorrelation function corresponding to the oscillogram obtained at the current 120 A.

cial calorimeter [19] enabled the determination of the pulse energy and, hence, power with the accuracy $\pm 10\%$.

As mentioned in the previous section, we observed the excitation of only one working mode E_{01} with the center frequency 8.7 GHz. Figure 5 shows a typical oscillogram in the chaotic modulation mode at the injection current 120 A. The self-modulation is seen to cover virtually the entire microwave pulse with a duration of about 10 μs . However, the more detailed analysis of the envelope shows that the character of self-modulation varies during the pulse, which can be related to minor changes in the voltage. As a result, the coupling impedance slightly also varies within the pulse. In addition, the electrodynamic system inevitably exhibits parasitic reflections from the collector end, which, as known, lead to the strong frequency dependence of the output power, the dependence varying with voltage [6].

In connection with this, we analyzed the bifurcation values of the current, using the central fragments of the oscillograms with the fixed peak voltage (≈ 150 kV) of the ≈ 500 -ns-wide pulse. Figure 6 shows these fragments together with the corresponding spectra and the phase portraits on the $(|A(0, t)|, |A(0, t - t_{\text{del}})|)$ plane. It is seen that stationary generation, observed at currents smaller than 7 A (Fig. 6a), changes to self-modulation as the current increases (Figs. 6b–6f). At 30 A, the self-modulation is near-sinusoidal with the period 8 ns (Fig. 6b). Then, the self-modulation becomes more and more complicated and finally chaotic at 50 A (Fig. 6c). At currents exceeding 50 A, the self-modulation mode

takes a simpler form. Specifically, the process is quasi-harmonic with the period 14 ns for currents from 70 to 90 A (Figs. 6d, 6e). Finally, the self-modulation becomes chaotic again with a relative spectral width of about 4% when currents exceed 100 A (Fig. 6f). Note that the bifurcation values of the current and the characteristic periods of self-modulation are in good agreement with the results of simulation. It follows from the numerical simulation that the self-modulation modes alternate with increasing injection current because of the space-charge effect. Note also that the autocorrelation function (Fig. 7) constructed from the oscillograms shown in Fig. 6f is in good agreement with that based on the results of simulation at the injection current 120 A (Fig. 2).

The experimental results exhibit the good pulse-to-pulse reproducibility except for the trailing edge of the voltage pulse. The leading and trailing edges of the microwave pulse are not mirror-symmetric presumably because of microwave breakdowns that arise when the electrodynamic system is bombarded by positive ions of the beam. At accelerating voltages above 150 kV, the microwave pulse becomes much shorter. Therefore, we can admit that, at injection currents exceeding 100 A, the power of self-modulation generation is limited by microwave breakdown. At the injection current 50 A and the voltage 150 kV, the calorimetrically measured power of self-modulation generation was 1 MW; in other words, the electron efficiency reached 10%. The power of chaotic generation was about 2 MW at the current 120 A.

ACKNOWLEDGMENTS

We are grateful to M.I. Petelen, É.B. Abubakirov, N.F. Kovalev, and M.I. Yalandin for the helpful discussions.

This work was supported by the Russian Foundation for Basic Research (grant no. 00-02-16412a).

REFERENCES

1. N. S. Ginzburg, N. I. Zaïtsev, E. V. Ilyakov, *et al.*, Pis'ma Zh. Tekh. Fiz. **24** (20), 66 (1998) [Tech. Phys. Lett. **24**, 816 (1998)].
2. N. S. Ginzburg, N. I. Zaïtsev, E. V. Ilyakov, *et al.*, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelin. Din., No. 5, 60 (1999).
3. M. I. Yalandin, V. G. Shpak, S. A. Shunaïlov, and M. R. Ul'maskulov, Pis'ma Zh. Tekh. Fiz. **25** (10), 19 (1999) [Tech. Phys. Lett. **25**, 388 (1999)].
4. N. M. Ryskin and V. N. Titov, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelin. Din. **6** (1), 75 (1998).
5. I. V. Pegel', Izv. Vyssh. Uchebn. Zaved., Fiz. **39** (12), 62 (1996).
6. B. Levush, Th. M. Antonsen, A. Bromborsky, *et al.*, IEEE Trans. Plasma Sci. **20**, 263 (1992).
7. N. S. Ginzburg, S. P. Kuznetsov, and T. N. Fedoseeva, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **21**, 1037 (1978).

8. N. S. Ginzburg and S. P. Kuznetsov, in *Relativistic High-Frequency Electronics: Problems of Radiating Power and Frequency Rise* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1981), pp. 101–144.
9. B. P. Bezruchko, L. B. Bulgakova, S. P. Kuznetsov, and D. I. Trubetskov, in *Lectures on Microwave Electronics and Radiophysics: 5th Winter School–Workshop for Engineers* (Saratovs. Gos. Univ., Saratov, 1980), Vol. 5, pp. 25–77.
10. V. P. Bezruchko, S. P. Kuznetsov, and D. I. Trubetskov, in *Nonlinear Waves: Stochasticity and Turbulence* (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1980), pp. 29–36.
11. V. P. Bezruchko, S. P. Kuznetsov, and D. I. Trubetskov, *Pis'ma Zh. Éksp. Teor. Fiz.* **29**, 180 (1979) [JETP Lett. **29**, 162 (1979)].
12. N. I. Zaıtsey, E. V. Ilyakov, G. S. Korablev, *et al.*, *Prib. Tekh. Éksp.*, No. 3, 138 (1995).
13. N. F. Kovalev, M. I. Petelin, M. D. Raızer, *et al.*, *Pis'ma Zh. Éksp. Teor. Fiz.* **18**, 232 (1973) [JETP Lett. **18**, 138 (1973)].
14. E. V. Ilyakov, G. S. Korablyov, I. S. Kulagin, *et al.*, *IEEE Trans. Plasma Sci.* **26**, 332 (1998).
15. N. S. Ginzburg, N. I. Zaıtsey, E. V. Ilyakov, *et al.*, *Izv. Vyssh. Uchebn. Zaved., Radiofiz.* **41**, 1565 (1998).
16. N. F. Kovalev, *Élektron. Tekh.*, Ser. 1: *Élektron. SVCh*, No. 3, 102 (1978).
17. E. S. Venttsel', *Theory of Probability* (Nauka, Moscow, 1964).
18. N. S. Ginzburg, N. A. Zavol'skiı, V. E. Zapevalov, *et al.*, *Zh. Tekh. Fiz.* **70** (4), 99 (2000) [Tech. Phys. **45**, 480 (2000)].
19. N. I. Zaıtsey, E. V. Ilyakov, Yu. K. Kovneristyı, *et al.*, *Prib. Tekh. Éksp.*, No. 2, 153 (1992).

Translated by A. Chikishev

ELECTRON AND ION BEAMS,
ACCELERATORS

On the Propagation of a Paraxial Beam
of Charged Particles
in an External Electromagnetic Field

N. D. Naumov

Received November 17, 2000

Abstract—A small-angle approximation method as applied to a curvilinear beam of relativistic particles in mutually orthogonal uniform magnetic and nonuniform electric fields is elaborated. Analytical solutions to the paraxial equation for curvilinear beams of charged particles are obtained. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

The necessity for analyzing the motion of charged particle fluxes in external electromagnetic fields is dictated by practical problems of electron–ion optics, mass spectroscopy, etc. [1–3]. One method for theoretically treating the dynamics of charged particle beams is the paraxial theory. This theory allows one to find the geometry of beam-forming electrodes that provide the flux of required configuration, as well as to consider the direct formulation of the problem when the parameters of an external electromagnetic field in which the beam propagates are given.

In general the case, when the beam axis is a spatial curve, the equations of the paraxial theory are rather complicated [4]. As has been shown in [5] for a nonrelativistic beam, simpler expressions arise if the beam axis is a plane curve. In this paper, the approach of [5] is generalized to the relativistic case. A narrow beam with a small ratio of its transverse size to the radius of curvature is considered. In this case, analytical results accurate up to first-order terms in this small parameter are obtained.

FORMULATION OF THE PROBLEM

To be specific, let us assume that the beam propagates perpendicular to the z axis in mutually orthogonal nonuniform electric field $\mathbf{E}_{\text{ext}} = E_x \mathbf{e}_x + E_y \mathbf{e}_y$ and uniform magnetic field $\mathbf{B}_0 = B_0 \mathbf{e}_z$. The propagation of the beam is convenient to consider in a curvilinear coordinate system (x, q, ζ) :

$$\mathbf{x} = \mathbf{X}(s) + q\mathbf{n} + \zeta\mathbf{b}.$$

The curve $\mathbf{X}(s)$ is the beam axis in the plane $z = 0$; s is the path length for axial particles reckoned from the point of beam injection; \mathbf{t} , \mathbf{n} , and $\mathbf{b} = \pm \mathbf{e}_z$ are the vectors of the Frenet trihedral for the curve $\mathbf{X}(s)$. The direction

of the vector \mathbf{b} depends on the direction of the external electric field.

The beam axis position given by the trajectories of axial particles is defined by the solution of the relativistic equation of motion for a single particle in an external field. Substituting the expression for particle momentum $\mathbf{p} = mu\mathbf{t}$ into this equation of motion, we obtain

$$muu' = \gamma e E_{01}, \quad k = \frac{\gamma e E_{02}}{mu^2} - \frac{\Omega}{u}. \quad (1)$$

Here, E_{0i} are the components of the external electric field at the beam axis: $\mathbf{E}_{\text{ext}}(\mathbf{X}(s)) = E_{01}\mathbf{t} + E_{02}\mathbf{n}$, $\gamma = \sqrt{c^2 + u^2}/c$, and $\Omega = e\mathbf{b} \cdot \mathbf{B}_0/mc$. The prime denotes differentiation with respect to s . In order to determine the characteristics of the curvilinear beam of charged particles moving in an external electromagnetic field, one has to find the trajectory equation for a particle with injection conditions different from those for an axial one. The difference is that the particle is injected at some distance away from the axis and at an angle to it; in addition, the initial kinetic energy of the particle does not coincide with that of an axial particle. The kinetic equation in the small-angle approximation [6] can be used to take into account effects like thermal straggling of the particles and multiple elastic scattering by gas molecules.

TRAJECTORY EQUATION

A trajectory equation for a charged particle in an external electromagnetic field may be found by the Maupertius principle. In relativistic mechanics, this principle is written in the form [7]

$$\delta \int \left(m\Lambda dl + \frac{e}{c} \mathbf{A} \cdot d\mathbf{x} \right) = 0.$$

Here, $\Lambda = \sqrt{(E - e\Phi)^2/m^2c^2 - c^2}$, E is the total particle energy, Φ and \mathbf{A} are the potentials of the external field, and l is the path length. Calculations similar to those for a nonrelativistic particle in a potential field [8] allow one to find the trajectory equation

$$m\Lambda \frac{d}{dl} \left(\Lambda \frac{d\mathbf{x}}{dl} \right) = \frac{e}{mc^2} (E - e\Phi) \mathbf{E}_{\text{ext}} + \frac{e}{c} \Lambda \left[\frac{d\mathbf{x}}{dl} \times \mathbf{B}_0 \right]. \quad (2)$$

It is easy to check that the quantity $m\Lambda(\mathbf{x}(l))$ is the absolute value of the particle momentum. In particular, $\Lambda(\mathbf{X}(s)) = u$ for axial particles.

In order to find an approximate trajectory equation in the vicinity of the curve $\mathbf{X}(s)$, one should represent the particle momentum in the form

$$m\Lambda \frac{d\mathbf{x}}{dl} = mu[(1 + \lambda)\mathbf{t} + q'\mathbf{n}]. \quad (3)$$

For near-axial trajectories, the values of kq and λ are small. Substituting expression (3) into Eq. (2) and omitting second-order terms, we arrive at the set of ordinary differential equations

$$q'' + \frac{u'}{u}q' + k^2q + \left(k + \frac{\kappa}{\gamma^2} \right) \lambda = \frac{\gamma e g_2}{mu^2} q, \quad (4)$$

$$\lambda' + \Gamma \frac{u'}{u} \lambda = \kappa q' - kq \frac{u'}{u} + \frac{\gamma e g_1}{mu^2} q. \quad (5)$$

Here, $\kappa = k + \Omega/u$, $\Gamma = 1 + 1/\gamma^2$, and g_i are the components of the external electric field

$$\mathbf{E}_{\text{ext}} = (E_{01} + g_1q)\mathbf{t} + (E_{02} + g_2q)\mathbf{n}$$

in the vicinity of the curve $X(x)$.

PARAXIAL EQUATION

In order to obtain the paraxial equation in the closed form, it is necessary to find the explicit form of the function λ . As has been found recently [5], the functions g_i have the form

$$g_1 = E'_{02} + kE_{01}, \quad g_2 = kE_{02} - E'_{01}. \quad (6)$$

The substitution of expression (6) for g_1 into Eq. (5) makes it possible to represent the latter in the form

$$\lambda' + \Gamma \frac{u'}{u} \lambda = \frac{\gamma e}{mu^2} \frac{d}{dx} q E_{02},$$

which is more convenient for integration.

In view of relationships (1), we arrive at the expression for λ

$$\lambda = \kappa q + C \frac{\gamma}{u^2},$$

where $C = u_0^2(\lambda_0 - \kappa_0 q_0)/\gamma_0$ is a constant.

Substituting the obtained equation for λ and the expression (6) for g_2 into (4), we finally find, in view of (1),

$$q'' + \frac{u'}{u}q' + \left(k^2 + \frac{\kappa^2}{\gamma^2} + \frac{u''}{u} + \frac{u'^2}{\gamma^2 u^2} \right) q + \frac{C}{u^2} \left(\gamma k + \frac{\kappa}{\gamma} \right) = 0. \quad (7)$$

The curvature of the beam axis is $k = \sqrt{X''^2 + Y''^2}$. Thus, the trajectory equation for a particle injected into the external field under consideration in the vicinity of the curve $X(s)$ has been obtained up to first-order terms.

The angle $q'_0 = \alpha$ between the injection direction and the beam axis, as well as the distance q_0 of the injection point to the beam axis, are initial conditions for Eq. (7). The initial kinetic energy T of a particle defines the constant C appearing in (7):

$$C = \frac{D}{m} (\lambda_0 - \kappa_0 q_0), \quad \lambda_0 = \frac{1}{D} (T - W), \quad (8)$$

$$D = W - \frac{m^2 c^4}{W},$$

where $W = \gamma_0 mc^2$ is the initial kinetic energy of an axial particle.

Generally, numerical methods should be used to solve the differential equation obtained. Analytical results may easily be found for a beam with a constant curvature of the axis. Such beams are used in mass spectroscopy [3]: the beam axis configures into a circular arc in both a sector magnetic field and the electric field of a cylindrical capacitor.

In a sector magnetic field, the beam curvature is $k = 1/R$, where $R = mc u_0 / e B_0$ is the circle radius. Since $\kappa = 0$ and the particle velocity is constant in this case, Eq. (7) is substantially simplified:

$$q'' + k^2 q + k \lambda_0 = 0. \quad (9)$$

Then, the path of a charged particle in a sector magnetic field is given by

$$q = (q_0 + \lambda_0 R) \cos \psi + R(\alpha \sin \psi - \lambda_0),$$

where $\psi = s/R$.

For the beam in the nonuniform electric field of a cylindrical capacitor, $k = \kappa = 1/R$, where R is the radius of a circle described by axial particles with a kinetic energy $W = (A + \sqrt{A^2 + 4m^2 c^4})/2$. Here, the constant $A = eU/\ln(b/a)$, where U is the voltage across the capacitor and a and b are the inner and outer radii of the capacitor plates, respectively.

In this case, Eq. (7) is written in the form

$$q'' + k^2 \Gamma [k(q - q_0) + \lambda_0 R] = 0,$$

which is similar to (9).

The solution of this equation defines the path of a relativistic particle in the field of a cylindrical capacitor:

$$q = q_0 + \lambda_0 R (\cos \psi - 1) + \alpha R_0 \sin \psi,$$

where $\psi = s/R_0$ and $R_0 = R/\sqrt{2 - A/W}$.

KINETIC EQUATION

Now let us turn to the kinetic equation

$$\mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + e \left(\mathbf{E}_{\text{ext}} + \frac{1}{c} [\mathbf{v} \times \mathbf{B}_0] \right) \cdot \frac{\partial f}{\partial \mathbf{p}} = I_{\text{el}} \quad (10)$$

for a curvilinear beam of charged particles. Here, $f(\mathbf{x}, \mathbf{p})$ is the distribution function and I_{el} is the elastic collision integral.

We write Eq. (10) in the curvilinear coordinates, taking $\mathbf{p} = mu[(1 + \lambda)\mathbf{t} + \alpha\mathbf{n} + \beta\mathbf{b}]$. In the case of a narrow beam, the angles α and β , along with λ , are small. As a result, up to first-order terms, we find

$$Lf + \left[\kappa\alpha - (\lambda\Gamma + kq) \frac{u'}{u} + \frac{e\gamma g_1}{mu^2 q} \right] \frac{\partial f}{\partial \lambda} + \left[\frac{e\gamma g_2}{mu^2} q - k^2 q - \left(k + \frac{\kappa}{\gamma} \right) \lambda - \alpha \frac{u'}{u} \right] \frac{\partial f}{\partial \alpha} = 0, \quad (11)$$

where χ^2 is the mean square of the scattering angle per unit path length and

$$L = \frac{\partial}{\partial s} + \alpha \frac{\partial}{\partial q} + \beta \frac{\partial}{\partial \zeta} - \beta \frac{u'}{u} \frac{\partial}{\partial \beta} - \frac{\chi^2}{4} \left(\frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \beta^2} \right).$$

In order to simplify Eq. (11), one should turn from λ to the new variable $\mu = (\lambda - \kappa q)u^2/\gamma$. Finally, in view of (1) and (6), we find the kinetic equation for a curvilinear paraxial beam of relativistic particles:

$$Lf - \left[\alpha \frac{u'}{u} + \left(k^2 + \frac{\kappa^2}{\gamma^2} + \frac{u''}{u} + \frac{u'^2}{\gamma^2 u^2} \right) q + \frac{\mu}{u^2} \left(k\gamma + \frac{\kappa}{\gamma} \right) \right] \frac{\partial f}{\partial \alpha} = 0. \quad (12)$$

Analytical solutions of Eq. (12) can be found for beams with a constant curvature considered above. In particular, the method proposed in [5] can be employed to find the Green's function.

PARAXIAL REPRESENTATION

Let us return to the paraxial equation again. The proper time of an axial particle,

$$\tau = \int_0^\pi \frac{dx}{u(x)},$$

may be used as a longitudinal variable instead of s . This means the transition to the parametric representation of the particle path. Then, Eq. (7) is written in the form

$$\ddot{q} + \left[\left(2k^2 + \frac{\kappa^2}{\gamma^2} \right) u^2 + \frac{\mathbf{u} \cdot \dot{\mathbf{w}}}{u^2} - \left(\frac{\mathbf{u} \cdot \mathbf{w}}{\gamma uc} \right)^2 \right] q + C \left(\gamma k + \frac{\kappa}{\gamma} \right) = 0, \quad (13)$$

where the dot denotes differentiation with respect to τ , $\mathbf{u} = \dot{\mathbf{X}}$, and $\mathbf{w} = \dot{\mathbf{u}}$. In this notation, the path curvature takes the form $k^2 = |\mathbf{u} \times \mathbf{w}|^2/u^6$.

For a uniform electromagnetic field, $\mathbf{u} \cdot \dot{\mathbf{w}} - (\mathbf{u} \cdot \mathbf{w}/\gamma c)^2 = ku\Omega$ and Eq. (12) is reduced to

$$\ddot{q} + \left(k^2 + k\kappa + \frac{\kappa^2}{\gamma^2} \right) u^2 q + C \left(\gamma k + \frac{\kappa}{\gamma} \right) = 0. \quad (14)$$

Equations (13) and (14) have to be solved jointly with the equation of motion for an axial particle

$$m\mathbf{w} = e\gamma \mathbf{E}_{\text{ext}} + \frac{e}{c} [\mathbf{u} \times \mathbf{B}_0].$$

For a nonrelativistic particle, $\gamma = 1$, the variable u is equal to the axial particle velocity v and the parameter τ coincides with the particle motion time. Because of this, Eq. (13) takes the simpler form

$$\ddot{q} + \left[(2k^2 + \kappa^2)v^2 + \frac{\mathbf{v} \cdot \dot{\mathbf{v}}}{v^2} \right] q + C(k + \kappa) = 0, \quad (15)$$

where $\mathbf{v} = \dot{\mathbf{X}}$ (the dot means differentiation with respect to time).

The expression for path curvature changes correspondingly: $k^2 = |\mathbf{v} \times \dot{\mathbf{v}}|^2/v^6$.

For $T = mc^2 + K$ and $W = mc^2 + K_0$, expressions (8) in the nonrelativistic limit are recast as

$$C = \frac{2K_0}{m} (\lambda_0 - \kappa_0 q_0), \quad \lambda_0 = \frac{1}{2} \left(\frac{K}{K_0} - 1 \right).$$

Equation (15) has to be supplemented by the equation of motion

$$m\dot{\mathbf{v}} = e\mathbf{E}_{\text{ext}} + \frac{e}{c} [\mathbf{v} \times \mathbf{B}_0].$$

A BEAM IN AN ELECTRIC FIELD

The solution of Eq. (15) can be found for beams in uniform fields, where $\dot{\mathbf{v}} = 0$. In a uniform electric field $\mathbf{E}_0 = E_0 \mathbf{e}_y$, the beam axis in the cross section is a segment of parabola:

$$Y = wX^2/2v_0^2 \sin^2 \Phi - X \cot \phi.$$

Here, $X = v_0 t \sin \phi$, $w = eE_0/m$, ϕ is the angle between the vector of the electric field and the injection direction for axial particles with the initial velocity $v_0 = \sqrt{2K_0/m}$, and the longitudinal coordinate s is parametrically represented as

$$s = \frac{1}{2w} \left[v_0 (wt - v_0 \cos \phi) + v_0^2 \left(\cos \phi + \sin^2 \phi \ln \frac{wt - v_0 \cos \phi + v}{v_0(1 - \cos \phi)} \right) \right],$$

where $v = \sqrt{v_0^2 + wt(wt - 2v_0 \cos \phi)}$ is the velocity of axial particles.

In this case, the beam axis curvature is $k = \kappa = wv_0 \sin \phi / v^3$. Therefore, Eq. (15) takes the form

$$\ddot{q} + 3v_0^2 \sin^2 \phi \frac{w^2}{v^4} q = Q, \quad Q = -2Cw \sin \phi \frac{v_0}{v^3}. \quad (16)$$

The solution of Eq. (16) may be found with the method of variation of constants [9]. First, one has to find the functions q_1 and q_2 representing two independent solutions of the homogeneous equation corresponding to Eq. (16). It is easy to check that these functions satisfy the condition

$$q_1 \dot{q}_2 - \dot{q}_1 q_2 = C_0, \quad (17)$$

where C_0 is a constant.

Then, the solution of the inhomogeneous equation is calculated using the functions q_1 and q_2 :

$$q = q_1 \left(C_1 - \frac{1}{C_0} \int q_2 Q dt \right) + q_2 \left(C_2 + \frac{1}{C_0} \int q_1 Q dt \right).$$

The constants C_i are defined by the initial conditions.

One of the solutions of the homogeneous equation corresponding to (16) should be found in the form of a power law of v . This yields $q_1 = v - 2v_0^2 \sin^2 \phi / v$. Using condition (17), we find the second independent solution $q_2 = (wt - v_0 \cos \phi) / v$. Eventually, for a beam in an electric field, we have

$$q = \frac{1}{v} [(q_0 + \alpha v_0 t)(v_0 - wt \cos \phi) - \lambda_0 v_0 w t^2 \sin \phi].$$

A BEAM IN MUTUALLY ORTHOGONAL FIELDS

A trochoid,

$$X = Vt - \frac{1}{\omega} \sin \psi [V \cos \phi - v_0 \sin(\psi + \phi)],$$

$$Y = \frac{1}{\omega} [v_0 \cos(\psi + \phi) + V \sin \psi]$$

is the path of a nonrelativistic charged particle in mutually orthogonal uniform electric and magnetic fields. Here, $V = cE_0/B_0$, $\psi = \omega t$, and $\omega = eB_0/2mc$. As in the previous problem, ϕ is the angle between the electric field vector and the direction of particle injection. For brevity sake, we here omit the relation between s and t , which may be represented via the elliptic integral of the second kind.

This implies that the velocity of an axial particle is defined by the expression

$$v^2 = v_0^2 + 4V \sin \psi [V \sin \psi + v_0 \cos(\psi + \phi)].$$

The beam axis curvature is $k = \omega(1 + \rho)/v$, where $\rho = v_0^2(v_0^2 - 2V \sin \phi)/v^2$.

As a result, Eq. (15) for a beam in mutually orthogonal fields is written in the form

$$\ddot{q} + \omega^2(1 + 3\rho^2)q + 2\rho\omega \frac{v_0^2}{v} \left(\lambda_0 + 2q_0 \omega \frac{V}{v_0^2} \sin \phi \right) = 0. \quad (18)$$

Unfortunately, an analytical solution of Eq. (18) cannot be found for an arbitrary value of the initial velocity of axial particles. The exception is the specific case $v_0 = 2V \sin \phi$, when the beam axis is a segment of a cycloid

$$X = Vt + \frac{V}{2\omega} [\sin 2\phi - \sin 2(\psi + \phi)],$$

$$Y = \frac{V}{2\omega} [\cos 2\phi - \cos 2(\psi + \phi)].$$

In this case, a characteristic feature of the particle motion in the vicinity of the beam axis is that the particle path does not depend on the initial energy:

$$q = q_0 \cos \psi + 2\alpha \frac{V}{\omega} \sin \phi \sin \psi,$$

where $\psi = \arccos(\cos \phi - 2\omega s/V) - \phi$.

The practical importance of the method developed is in the possibility of estimating the parameters of a curvilinear beam of relativistic particles propagating in an external electromagnetic field. These estimates are valid when the beam is not yet significantly broadened

due to effects like thermal straggling of the particles and multiple elastic scattering by gas molecules.

REFERENCES

1. V. M. Kel'man and S. Ya. Yavor, *Electron Optics* (Nauka, Moscow, 1968).
2. S. I. Molokovskii and A. D. Sushkov, *High-Intensity Electron and Ion Beams* (Énergoatomizdat, Moscow, 1991).
3. R. Jayaram, *Mass Spectrometry: Theory and Applications* (New York, 1966; Mir, Moscow, 1969).
4. V. A. Syrovoi, *Radiotekh. Élektron. (Moscow)* **37**, 1692 (1993).
5. N. D. Naumov, *Zh. Tekh. Fiz.* **70** (10), 90 (2000) [*Tech. Phys.* **45**, 1320 (2000)].
6. T. P. Hughes and B. B. Godfrey, *Phys. Fluids* **27**, 1531 (1984).
7. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).
8. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics* (Nauka, Moscow, 1973; Pergamon, New York, 1988).
9. J. Mathews and R. L. Walker, *Mathematical Methods of Physics* (Benjamin, Menlo Park, 1970; Atomizdat, Moscow, 1972).

Translated by M. Fofanov

ELECTRON AND ION BEAMS, ACCELERATORS

A High-Frequency Undulator for Ion Beam Focusing and Acceleration

É. S. Masunov

Moscow State Engineering Physics Institute (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia

e-mail: masunov@edhem.mephi.msk.su

Received December 6, 2000

Abstract—The use of high-frequency field harmonics nonsynchronous with the beam for focusing and accelerating charged particles in periodic resonant structures is discussed. Conditions for effective acceleration of particles in the transverse and longitudinal fields of a high-frequency undulator are found. A specific implementation of particle focusing and acceleration, as well as the numerical calculation of deuterium ion dynamics in a new type of a linear accelerator, are reported. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

In [1, 2], the dynamics of nonrelativistic ion beams in a linear accelerator where the particles are accelerated in the field of a periodic electrostatic undulator and in an alternating rf field has been considered. It has been shown that the superposition of the fields of the electrostatic undulator and of the periodic rf resonator in which waves synchronous with the beam are absent (such a resonator will be referred to as an rf undulator for brevity) provides the simultaneous acceleration and focusing of the beam particles. An undulator accelerator can be made compact if the same electrodes are used to excite the rf field in the accelerating channel and the periodic electrostatic field, that is, if the rf undulator is structurally integrated with the electrostatic undulator. Tentative mock-up experiments [3] have shown that the implementation of these devices may face serious problems associated with rf breakdown when the rf and the electrostatic fields of large amplitude are generated simultaneously. In addition, the effective operation of an undulator accelerator requires the use of a high-power high-voltage dc source along with an rf generator. Our technique for focusing and accelerating the beam using nonsynchronous waves makes it possible to omit the electrostatic fields.

It is the aim of this work to formulate basic demands for an rf undulator to provide the effective acceleration and focusing of ion beams with low injection energies.

EQUATION OF MOTION

The field excited in a periodic resonator will be found as a periodic solution of the Maxwell equations with given boundary conditions. Under the assumption that the cross size of the resonator is much smaller than the wavelength of the rf field, the Fourier series coefficients for the field can be calculated with the electrostatic approximation to the problem. In this case, the

longitudinal and transverse components of the electric field can be represented as the sum over the spatial harmonics:

$$\mathbf{E}_{\perp} = \sum_n \mathbf{E}_{n,\perp}(x, y) \sin(h_n z + \alpha) \cos(\omega t), \quad (1)$$

$$E_z = \sum_n E_{n,z}(x, y) \sin(h_n z + \alpha) \cos(\omega t),$$

where $h_n = h_0 + 2\pi n/D$, $h_0 = \mu/D$, μ is the phase advance of the field per period, and D is the structure period.

Consider the equation of motion for a nonrelativistic ion beam in field (1), assuming that the particle velocity v differs from the phase velocity of all harmonics: $v_{\text{ph},n} = \omega/h_n$, $n = 0, 1, 2, \dots$

In general, the interaction of the particles with nonsynchronous harmonics of the rf field does not change the average energy of the beam but causes fast oscillations in the longitudinal and transverse directions. In further analysis, it is necessary to make a distinction between a longitudinal undulator, for which the transverse components of the field are absent [$\mathbf{E}_{n,\perp}(0, 0) = 0$], and a transverse undulator, for which the longitudinal component field at the axis vanishes [$E_{n,z}(0, 0) = 0$]. Accordingly, the parameter α in sum (1) should be set equal to zero in the former case and $\pi/2$ in the latter.

As has been shown in [2], even in the absence of synchronism between the beam particles and the spatial harmonics of the wave, effective beam–field interaction takes place if

$$v = \omega/k_z, \quad (2)$$

where $k_z = (h_n \pm h_p)/2$ ($k_z \neq h_n \neq h_p$; $n = 0, 1, 2$; $p = 0, 1, 2$).

The parameter k_z defines the wave number of the combined wave resulting when the fields of the n th and p th harmonics are added. In fact, introducing a slowly

varying phase $\psi = \int k_z dZ - \omega t$ and coordinate $\mathbf{R} = (X, Y, Z)$ and averaging over the fast longitudinal and transverse oscillation [2, 4], one comes to the equation of motion in the smooth approximation:

$$\frac{d^2 \mathbf{R}}{dt^2} = -\frac{\partial}{\partial \mathbf{R}} U_{\text{eff}}, \quad (3)$$

where the function U_{eff} depends on the amplitudes of the harmonics of the nonsynchronous waves.

Introducing the dimensionless amplitude of the field harmonics

$$\mathbf{e}_{n(\perp, z)} = \frac{e\lambda}{2\pi mc^2} \mathbf{E}_{n(\perp, z)},$$

reduced particle velocity β ; dimensionless coordinates $\xi = 2\pi Z/\lambda$, $\rho = 2\pi X/\lambda$, and $\eta = 2\pi Y/\lambda$; and dimensionless time $\tau = \omega t$, one can write the equation of motion in the form

$$\frac{d\beta}{d\tau} = -\nabla U_{\text{eff}}, \quad (4)$$

where

$$U_{\text{eff}} = U_1 + U_2 + U_3, \quad (5)$$

$$U_1 = \frac{1}{16} \sum_n (\mathbf{e}_n)^2 \left(\frac{1}{(\Delta_n^-)^2} + \frac{1}{(\Delta_n^+)^2} \right), \quad (5a)$$

$$U_2 = \frac{1}{16} \sum_{h_p + h_n = 2k} \left(\frac{e_{n,z} e_{p,z} - \mathbf{e}_{n,\perp} \mathbf{e}_{p,\perp}}{(\Delta_n^-)^2} \right) \times \cos(2\psi + 2\alpha), \quad (5b)$$

$$U_3 = \frac{1}{8} \sum_{|h_p - h_n| = 2k} \left(\frac{\mathbf{e}_n \mathbf{e}_p}{(\Delta_n^-)^2} \right) \cos(2\psi). \quad (5c)$$

Here, $\Delta_n^\pm = (h_n \pm k_z)/k_z$. In the expression for U_2 , the sum is taken only over the harmonics for which $h_n + h_p = 2k_z$; in the expression for U_3 , over the harmonics for which $|h_n - h_p| = 2k_z$. The function U_{eff} can be considered as the effective potential function that specifies the Hamiltonian of the beam-wave system and helps to study the 3D beam dynamics in the smooth approximation. From Eq. (4), it follows that the existence of an absolute minimum of U_{eff} is the necessary condition for beam focusing and acceleration. The longitudinal bunching and acceleration of the beam are possible if the particle velocity is close to the velocity $\beta_s = \omega/c k_z$ of a synchronous particle whose phase $\psi = \psi_s$ remains constant or slowly varies with the longitudinal coordinate. The variation of β_s can be written as

$$\frac{d\beta_s}{d\tau} = -\frac{\partial}{\partial \xi} U_{\text{eff}} \Big|_{\psi = \psi_s, \rho = 0, \eta = 0}. \quad (6)$$

From this expression, one can find the law of variation of the structure period $D(\xi)$ and the range for the phase ψ_s of the combined wave where the acceleration of the particles is the most efficient.

ANALYSIS OF BEAM DYNAMICS

By way of example, let us consider the problem of accelerating and focusing of a ribbon-shaped beam in the slot channel of a plane rf undulator. Figure 1 shows three possible electrode arrangements in this accelerating system for $\mu = \pi$. An alternating-sign field along the beam path can be generated by applying an rf potential $U_v = \pm U_0 \cos \omega t$ to a regularly (periodically) arranged transverse electrodes ("rods"). Depending on the difference between the phases of U_v applied to adjacent electrodes, a longitudinal (Fig. 1a) or a transverse (Fig. 1b) undulator can be implemented. Under the assumption that the field in the narrow slot channel depends only on the longitudinal coordinate Z and the transverse coordinate Y , the amplitudes of the electric field harmonics in the longitudinal undulator are given by

$$E_{n,z} = E_{n,0} \cosh(h_n Y), \quad E_{n,y} = E_{n,0} \sinh(h_n Y). \quad (7)$$

For the transverse undulator,

$$E_{n,z} = E_{n,0} \sinh(h_n Y), \quad E_{n,y} = E_{n,0} \cosh(h_n Y). \quad (8)$$

Now we find conditions for the focusing and accelerating of a ribbon beam, using expressions (7) and (8) for the field amplitudes. In analyzing the function U_{eff} it would suffice to leave the first two harmonics with $n = 0$ and 1 in sum (5) if the contribution from all other harmonics is assumed to be negligibly small. We start with the study of a periodic structure using the $\mu = 0$ mode. In this case, if the beam particle velocity is close to $\beta = 2D/\lambda$, the expression for U_{eff} can be written as

$$U_{\text{eff}} = \frac{1}{8} \left[e_0^2 + \frac{5}{9} e_1^2 \cosh\left(\frac{4\eta}{\beta_s}\right) + 2e_0 e_1 \cosh\left(\frac{2\eta}{\beta_s}\right) \cos(2\psi) \right]. \quad (9)$$

From (9), we come to the equation that describes the variation of the particle velocity:

$$\frac{d\beta}{d\tau} = \frac{e_0 e_1}{2\beta_s} \cosh\left(\frac{2\eta}{\beta_s}\right) \sin(2\psi). \quad (10)$$

The acceleration and self-focusing of the particles are possible if the phase of a synchronous particle ψ_s that is in the combined wave field lies in the intervals of $[\pi/4, \pi/2]$ and $[5\pi/4, 3\pi/2]$. In this case, according to (9), the focusing condition for all particles being accelerated will be met near the axis of the system ($\eta/\beta_s \ll 1$) if the amplitude of the first harmonic equals or exceeds that of the zero one ($e_1 \geq e_0$).

It is interesting to compare this result with another possibility of acceleration in the rf undulator where the

field is generated at the mode $\mu = \pi$. If the particle velocities in the beam are close to $\beta = D/\lambda$, we have

$$U_{\text{eff}} = \frac{1}{4} \left[e_0^2 \cosh\left(\frac{\eta}{\beta_s}\right) + e_1^2 \cosh\left(\frac{3\eta}{\beta_s}\right) + 3e_0e_1 \cosh\left(\frac{\eta}{\beta_s}\right) \cos(2\psi) \right] \quad (11)$$

instead of (9).

In (11), the second term in sum (5a), which is responsible for the interaction of the beam with the counterpropagating wave and whose contribution to the effective potential U_{eff} is negligible, is omitted for simplicity. Then, the equation for longitudinal motion can be written in the form

$$\frac{d\beta}{d\tau} = \frac{e_0e_1}{\beta_s} \cosh\left(\frac{\eta}{\beta_s}\right) \sin(2\psi), \quad (12)$$

that is, acceleration is possible at the same phases of the synchronous particle as before but the acceleration rate for a given structure period is four times higher. One more important advantage of the rf undulator operating at the $\mu = \pi$ mode is that the transverse focusing of the particles is possible for any ratio of the amplitudes of the zero and first field harmonics. This is easy to check by analyzing the form of potential function (11) as a function of the coordinates η and ψ .

Note that all the above results are valid for both types of the undulators. However, the transverse undulator is easier to design and implement. In fact, at low velocities of the ion beam, the spatial period D of the undulator rapidly increases with energy. For given values of the rf potential on the electrodes, it is difficult to provide the constancy of the field harmonic amplitudes at the axis of the system under such conditions. In the transverse undulator, the channel aperture can be made constant. This allows one to easily keep the maximal value of the accelerating field amplitude throughout the accelerator length for given electrode potentials.

In a plane undulator, the electrodes can be arranged in such a way (Fig. 1c) that the condition

$$E_{n,y}(y, 0) = E_{n,z}(y, 0) \quad (13)$$

for the mode $\mu = \pi$ is satisfied in any cross section of the channel. In this case, $U_2 = U_3 = 0$ in expression (5) and the effective potential U_{eff} is independent of the particle phase. This means that the acceleration equals zero but the undulator can be considered as a device providing the transverse focusing of a ribbon-shape beam throughout its length.

NUMERICAL SIMULATION OF THE ACCELERATOR

To exemplify the efficiency of acceleration in an rf undulator, we numerically simulated the dynamics of a ribbon deuterium ion beam. The energy of injection of

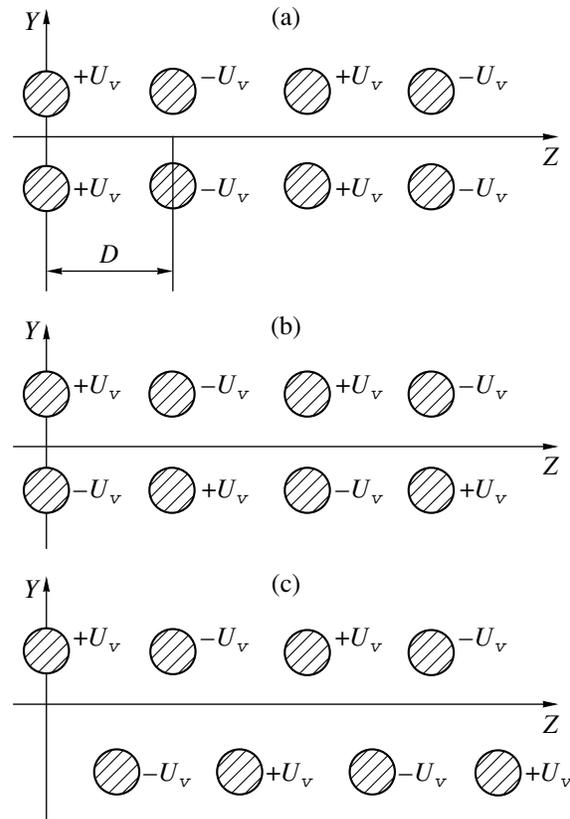


Fig. 1. Arrangement of electrodes in a planar rf undulator (the phase advance $\mu = \pi$).

the unmodulated ribbon beam is $W_{\text{in}} = 150$ keV; the operation mode in the transverse rf undulator, $\mu = \pi$; and the operating wavelength, 1.5 m. The structure period $D = \beta_s \lambda$, where β_s is found from Eq. (6). The accelerator comprises the buncher of length $L_b = 0.7$ m and the basic section of length $L = 1.8$ m. In the buncher, the field harmonic amplitude smoothly grows as $E(z) = E_{\text{max}} \sin(\pi z/2L_b)$ and the phase ψ_s linearly decreases from $\pi/2$ to $3\pi/8$. At this acceleration stage, the field amplitude remains unchanged, $E = E_{\text{max}}$, and the synchronous particle phase is $\psi_s = 3\pi/8$. The first-to-zero field harmonic amplitude ratio $\chi = e_1/e_0$ was selected such that the particle losses were as low as possible. The gain in the particle energy W and the coefficient K of current transmission depend on E_{max} . In our case, $W = 1.1$ MeV and $K = 64\%$ for $E_{\text{max}} = 150$ kV/cm and $\chi = 0.6$. As the amplitude of the fundamental harmonic increases to $E_{\text{max}} = 250$ kV/cm at $\chi = 0.3$, we find the final energy $W = 1.37$ MeV and $K = 63.5\%$. Obviously, the acceleration rate can be increased and K can be made larger by using special optimization techniques and by choosing more appropriate parameters of the buncher. However, even this example implies that, if the initial particle velocity is low, the system suggested compares well with conventional rf-focusing

accelerators and sometime outperforms them in efficiency of acceleration.

CONCLUSIONS

We have shown that an rf undulator can effectively bunch, accelerate, and focus ion beams. Acceleration is possible in both longitudinal and transverse rf field. The rf system suggested compares well with RFQ accelerators in rate of acceleration. By varying the field amplitude and the period of the rf undulator, one can provide the effective longitudinal bunching of the beam at the frequency equal to the doubled frequency of the rf field. In this case, a relatively high capture efficiency for the particles to be accelerated is attained.

A plane rf undulator of special geometry can be used for the focusing of ribbon-shaped unmodulated beams. Finally, an rf undulator offers wide possibilities for accelerating neutralized ion beams, that is, those consisting of oppositely charged particles (for example, D^+ and D^-). As follows from Eqs. (6), (10), and (12), the phase of the synchronous particle ψ_s does not depend on the charge sign. Therefore, positively and negatively charged ions can be accelerated in a single bunch [5]. This means that the use of neutralized beams tackles the

space-charge-related problem of limited beam intensity. As follows from the numerical simulation, all basic data obtained by analyzing the beam dynamics in the smooth approximation are close to those that are based on the exact calculation of the beam dynamics in the polyharmonic field of an rf undulator. For the ion energies considered, the difference in the output integral beam characteristics is within 5%.

REFERENCES

1. É. S. Masunov, USSR Inventor's Certificate No. 1 508 354 (1987).
2. É. S. Masunov, Zh. Tekh. Fiz. **60** (1), 152 (1990) [Sov. Phys. Tech. Phys. **35**, 962 (1990)].
3. N. V. Avrel'in, V. N. Leonov, É. S. Masunov, *et al.*, Preprint No. 041-90, MIFI (Moscow Engineering Physics Institute, Moscow, 1990).
4. E. S. Masunov, in *Proceedings of the 18th International Linac Conference* (CERN, Geneva, 1996), Vol. 2, p. 487.
5. E. S. Masunov and A. P. Novicow, in *Proceedings of the 4th European Particle Accelerator Conference, London, 1994*, Vol. 2, p. 1171.

Translated by V. Isaakyan

**SURFACES, ELECTRON AND ION
EMISSION**

Field Electron Emission in Graphite-Like Films

A. N. Obraztsov, I. Yu. Pavlovskii, and A. P. Volkov

Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

Received May 5, 1998; in final form, November 20, 2000

Abstract—Results of investigation of carbon films deposited with the use of gas-phase chemical reactions in the plasma of a dc discharge are presented. Films obtained at different parameters of the deposition process varied widely in their structure and phase composition, from polycrystalline diamond to graphite-like material. Comparative study of the structure and phase composition of the films using Raman spectroscopy, cathodoluminescence, electron microscopy, and diffractometry, as well as the obtained field electron emission characteristics, have shown that the threshold value of the electric field strength for electron emission decreases with a decrease in the size of diamond crystallites and growth of the fraction of non-diamond carbon. The lowest threshold fields (less than 1.5 V/ μm) are obtained for films consisting mainly of graphite-like material. A model based on the experimental data is proposed, which explains the mechanism of field electron emission in carbon materials. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

Various carbon materials are being considered now for use as efficient cathodes based on the phenomenon of field electron emission. This emission process needs no cathode heating and produces a beam of electrons with low initial velocities, which considerably simplifies the design of high-vacuum electronic devices. One of the most attractive applications of such “cold” cathodes is flat cathodoluminescent displays.

The possibility of using diamond materials as electron sources is based on their intrinsic property of negative electron affinity (NEA), which has been predicted theoretically and discovered experimentally both in single diamond crystals and in polycrystalline diamond films prepared by the method of gas-phase chemical deposition (GPCD) [1–5]. The occurrence of NEA depends mainly on the type of interatomic electronic bonds in the material and is typical not only of diamond but also of other wide band semiconductors [6–9]. In particular, emitters with a surface having NEA require considerably lower electric field strengths for initiating field electron emission, 1–10 V/ μm , compared with values of 10^3 – 10^4 V/ μm typical of the most metals and semiconductors.

It is evident that for obtaining stable electron emission the field emission cathode should be made of a material having high enough electron conduction. However, the problem of the synthesis of *n*-type semiconducting diamonds is still unresolved. On the other hand, in polycrystalline diamond films the electrical conductivity can be due to various structural defects, which introduce a set of additional levels into the forbidden gap of the diamond [5–11] or produce extended formations of non-diamond carbon [12, 13]. Many studies indicate that defects significantly improve the emissive properties of GPCD-grown diamond films (up

to concentrations at which formation of an amorphous material begins) retaining as its essential feature the diamond-type hybridization of the valence electron bonds of carbon atoms. In such emitters the threshold electric-field strength for electron emission is found in the range from 2 to 20 V/ μm [5, 12–15].

On the other hand, it is well known that the field electron emission is also observed in carbon having a graphite-like type of chemical bond between atoms. Crystalline graphite belongs to the semimetals, it has a high conductivity and a rather large electronic work function (about 4 eV) corresponding to positive electron affinity. However, in some cases the emission parameters of cathodes based on graphite-like materials are close to the parameters of diamond field-emission cathodes and can even considerably exceed them, exhibiting higher emission current density and, at the same time, lower threshold values of the electric field strength [16–18]. Moreover, recently, a large number of papers have been published reporting low-voltage field emission from carbon nanotubes [19–22], though, for materials consisting of carbon fibers having a similar structure, such observations had been made much earlier [23, 24].

So, the published data indicate that the possession of diamond structure is not absolutely necessary for efficient carbon field emitters. From the practical point of view, the production of nondiamond carbon materials appears to be a simpler task, but, in order to produce on their basis the “cold” cathodes, additional studies are necessary to elucidate the mechanisms of field emission. With this purpose in mind, in this work, a comparative study is carried out of the emissive, structural and other specific properties of carbon films grown by GPCD.

MATERIALS AND PROCEDURES

Studied in this work were films deposited on silicon substrates (*p*-Si(100), $\rho = 10 \text{ } \Omega \text{ cm}$, $20 \times 20 \text{ mm}$) by gas-phase chemical deposition using a mixture of hydrogen and methane activated by a dc arc discharge, in an installation described in detail elsewhere [25]. Prior to the deposition process all substrates were nucleated with diamond crystallites of nanometer size using ultrasonic treatment in a suspension of ultra-fine diamond in alcohol. The deposition process was carried out at a gas-mixture pressure of 8.9–9 kPa. The deposition time for all the samples studied in this work was 45 min. By varying the substrate temperature in the range from 850 to 1100°C and the methane concentration in the gas mixture from 0.5 to 10% carbon films containing different amounts of diamond and non-diamond phases were obtained.

The morphological and structural characteristics of the obtained films were studied using electron microscopy and atomic force microscopy, as well as electron diffraction and Raman scattering. Data on electronic properties of the films were obtained from cathodoluminescence (CL) studies.

Electron emission properties of the carbon films were studied using two procedures. In the first procedure the dependence of the emission current on the applied electric field strength was measured. For this purpose the sample was placed in a chamber pumped down to a vacuum of 10^{-6} Pa. The sample holder could be cooled to the liquid nitrogen temperature or heated to +350°C. The anode in the form of a tungsten rod 5 mm in diameter with a flat and polished end face could be displaced relative the cathode with a micrometer screw. The electric field strength in the gap between the anode and cathode was taken to be equal to $E = V/d$, where V is the potential difference between the anode and cathode separated with a vacuum gap of width d . The accuracy of positioning the anode relative to the cathode was 5 μm . Measurements of the current-voltage characteristics (IVC) were carried out automatically in a range of voltage between the anode and the cathode from 100 to 1500 V and with the maximum current not exceeding 2 mA.

The second procedure was used for studying the electron emission and provided data on the spatial distribution of emission centers. The sample was placed in a holder and displaced with a micrometer screw relative to the fixed anode having the form of a transparent electrically conductive film of indium oxide deposited on a glass plate. The conductive film was coated with an electron-excited phosphor a few micrometers thick. The separation between the anode and cathode being relatively small (from 50 to 500 μm), the area where the phosphor luminescence occurred due to electron bombardment could be considered coinciding with the regions of field electron emission. To prevent excessive ablation of the phosphor by incident electrons the measurements were carried out in the pulsed mode. The

voltage pulse duration was about 1 μs , the pulse repetition rate varied from 30 to 500 Hz, the peak voltage could be regulated between 200 and 2000 V. The obtained pattern of emission centers was registered using a video or photographic camera. The measurements using this procedure were carried out at room temperature.

RESULTS

1. The Phase Composition and Structural Properties of the Carbon Films

The films obtained under conditions described above were continuous and had an approximately equal thickness (about 1 μm). The surface morphology depended markedly on the deposition conditions. This dependence, as shown by electron and atomic force microscopy, was essentially the same as described in other studies (see, for example [26]). The phase composition of the films determined by Raman spectroscopy (using 488 nm line of a argon laser) also varied in accordance with known data [27–29]. Figure 1 shows typical Raman spectra of the studied GPCD-grown films. Spectrum 1 is that of a film deposited at a substrate temperature of $T_s = 950^\circ\text{C}$ and methane concentration (k) of about 1%. A feature of this film is an intensive narrow “diamond” line at 1330 cm^{-1} testifying to the predominantly diamond-like type of the constituent carbon. A wide line at about 1580 cm^{-1} is due to carbon in the form of amorphous graphite and the line at 1350 cm^{-1} to graphite in the form of nanocrystals of a size less than 10 nm [27]. Films having in their Raman spectra the 1330 cm^{-1} line consisted mainly of faceted crystals with a characteristic diamond cut. Spectrum 2 in Fig. 1 was obtained from a film deposited at $T_s = 850^\circ\text{C}$ and $k = 2\%$ and contained characteristic lines at 1140 and 1470 cm^{-1} , whose simultaneous presence in the Raman spectrum correlated with the decrease in the size of diamond crystallites in a GPCD film down to a few nanometers [28]. Such films, as a rule, had a mirror-like surface. Spectrum 3 in Fig. 1 was obtained from a film deposited at $T_s = 1050^\circ\text{C}$ and $k = 8\%$. This spectrum featured an intensive line at a frequency of 1575 cm^{-1} and a relatively weak line around 1350 cm^{-1} . The spectral shape practically coincided with Raman spectra of microcrystalline graphite with crystallites about 30 nm in size [29].

In Raman spectra of the films prepared at different parameters of the GPCD growth the relative intensities of the lines, their width, and, to some extent, spectral positions varied in accordance with variations of the phase composition, size of crystallites and the degree of amorphism of the constituent diamond and graphite. For the sake of convenience, in what follows, the films with Raman spectra containing a distinct “diamond” line around 1330 cm^{-1} are called “polycrystalline diamond” films; films having lines in their spectra at fre-

quencies 1140 and/or 1470 cm^{-1} are referred to as “nanocrystalline diamond” films; and, finally, films with spectra similar to spectrum 3 in Fig. 1 are referred to as “graphite-like” films.

The expediency of this classification has been proven in cathodoluminescent studies (electron energy in the beam of 25 keV). GPCD films, which had different sets of lines in their Raman spectra, also had differing CL spectra. Shown in Fig. 2 are the CL spectra for the same films as in Fig. 1. The spectra of polycrystalline diamond films (spectrum 1 in Fig. 2) contained well-known lines due to various structural defects and nitrogen and silicon impurities in the diamond (see, for example [30]). The relative line intensities of the CL spectra were different for different points at the surface of diamond crystallites, as described in [31], and varied with variation of the deposition conditions. For nanocrystalline films the CL intensity was typically lower than the sensitivity threshold of the measurement setup (spectrum 2 in Fig. 2). Such a low CL intensity can be due, in part, to a high rate of nonradiative recombination in diamond films with high concentration of structural defects. This is indicated by enhanced light absorption in the visible part of spectrum in nanocrystalline diamond films [32]. Finally, GPCD films, whose Raman spectra contained a narrow “graphite” line, typically had two CL lines at around 330 and 500 nm. In different samples of these graphite-like films variation were observed in the relative intensities and positions of these line over the ranges 310–340 and 490–530 nm, respectively. CL lines in about the same spectral ranges were observed earlier in diamond GPCD films [32, 33] but no interpretation of their origin has yet been suggested.

Conclusions concerning the phase composition of GPCD films made on the basis of the Raman and CL spectra have been corroborated by electron-diffraction studies and high-resolution electron microscopy data. Electron diffraction patterns of the polycrystalline and nanocrystalline films were typical of polycrystalline diamond. Electron diffraction patterns from graphite GPCD films had a number of specific features. Figure 3 displays an electron diffraction pattern of such film, clearly showing reflections corresponding to the crystal lattice of graphite. The arched shape of the line corresponding to electrons reflected from the basal plane of graphite indicates that the predominant orientation of the (0001) crystallographic planes of graphite is in a direction perpendicular to the substrate surface. Possible deviation from this direction does not exceed $\pm 20^\circ$.

This crystallographic orientation of the near-surface layer of graphite films was revealed with high-resolution electron microscopy. Clearly seen in the images of the film surface obtained by this technique are the parallel lines at separations equal to the interplanar spacing of graphite (~ 0.34 nm). Some crystallites in the film had a cross-sectional dimension less than 10 nm,

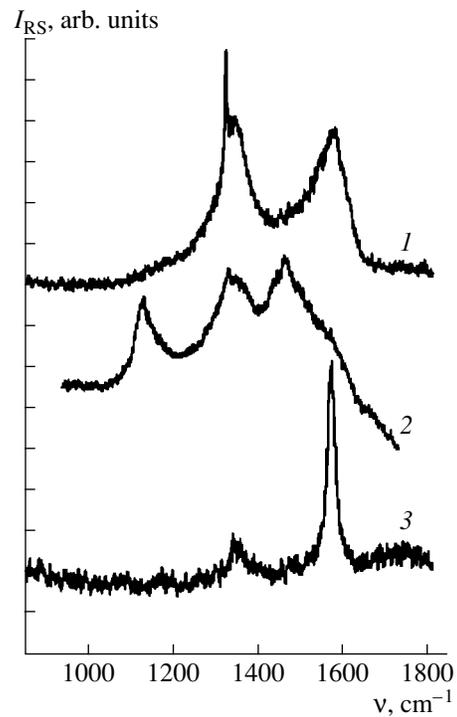


Fig. 1. Raman spectra, $I_{RS}(\nu)$, of (1) polycrystalline diamond, (2) nanocrystalline diamond, and (3) graphite-like films grown by GPCD.

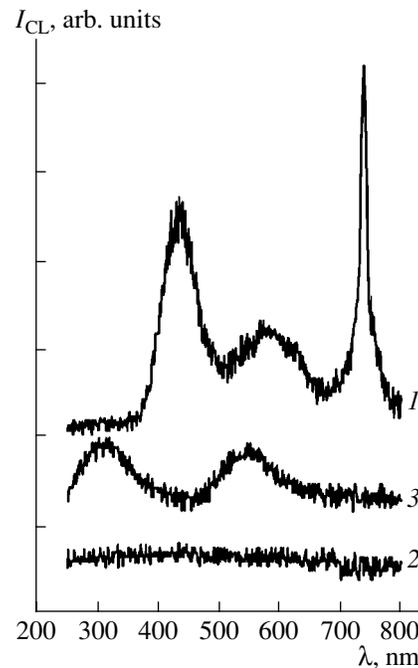


Fig. 2. Cathodoluminescence spectra, $I_{CL}(\lambda)$, of (1) polycrystalline diamond, (2) nanocrystalline diamond, and (3) graphite-like films grown by GPCD.

whereas their dimension along the basal plane could be in excess of $1\ \mu\text{m}$.

2. Field Emission Properties of Carbon Films

As demonstrated in this study, practically all carbon films grown by GPCD under the conditions specified above emitted electrons when a potential difference was applied between the film cathode and the anode. For polycrystalline diamond films, according to earlier experiments [1–15], the threshold value of the electric-field strength in the cathode–anode gap for the field electron emission to occur ($E_{\text{th}} = V/d$) was lower for films with greater amounts of the non-diamond carbon

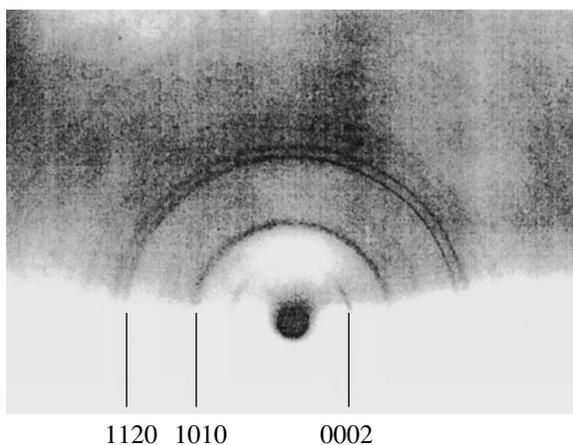


Fig. 3. Electron diffraction pattern of a graphite-like film grown by GPCD.

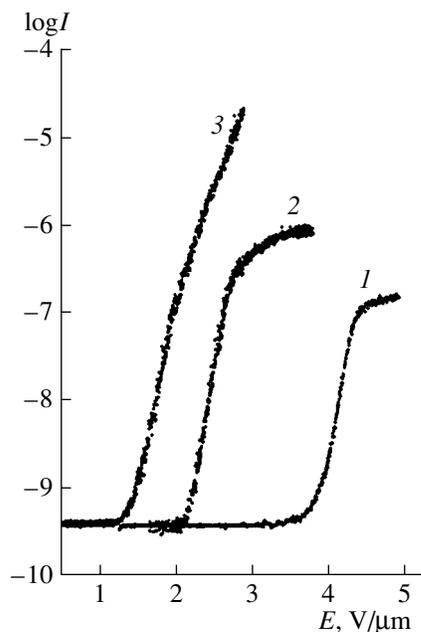


Fig. 4. Current–voltage characteristics of the field emission current from (1) polycrystalline diamond, (2) nanocrystalline diamond, and (3) graphite-like films grown by GPCD. The emission current density I is given in mA/cm^2 .

phase; simultaneously, the emission current density increased. The minimum value of E_{th} for such films was about $4\ \text{V}/\mu\text{m}$. For example, in Fig. 4 a current–voltage characteristic is given of the same polycrystalline diamond film (curve 1) whose characteristic are shown in Figs. 1 and 2 (curve 1). Figure 5 shows (1) the distribution of emission centers registered at $8\ \text{V}/\mu\text{m}$ for the same film. In this case, though the integrated emission current measured while registering the IVC was fairly stable, in observations with the use of a luminescent screen considerable instability was detected of the emission current from individual emission centers.

The tendency to lower threshold field values and higher emission currents continued with the decrease in the diamond crystallites in GPCD films down to sizes corresponding to nanocrystalline films. A current–voltage characteristic of one of such films, whose Raman and CL spectra are shown in Figs. 1 and 2, can be seen in Fig. 4 (curve 2). Nanocrystalline films had a much higher density of emission centers (up to $10^3\ \text{cm}^{-2}$) (Fig. 5b) and, at the same time, higher emission stability as regards both the average current density and the emission from individual centers.

Finally, in the case of graphite-like films the electron emission was observed at applied electric field strengths of less than $1.5\ \text{V}/\mu\text{m}$. The current–voltage characteristic in Fig. 4, curve 3, was obtained for a graphite film, the Raman and CL spectra of which are presented in Figs. 1 and 2, respectively. The emission current density in graphite films was $1\ \text{mA}/\text{cm}^2$ at electric field strengths as low as $4\ \text{V}/\mu\text{m}$. Simultaneously, a substantial increase in the emission stability and density of emission centers was observed. The latter was estimated as not lower than 10^6 – $10^7\ \text{cm}^{-2}$, taking into account that the average size of the phosphor grains was about $1.5\ \mu\text{m}$. A corresponding image of the luminescent screen at an electric-field strength of $8\ \text{V}/\mu\text{m}$ is shown in Fig. 5c.

Figure 6 shows current–voltage characteristics of the field emission current obtained for a graphite-like film at room temperature (1), at liquid nitrogen temperature (2), and at 300°C (3). Some decrease of the emission current density and increase of the threshold electric field strength at lower emitter temperatures can be due to residual gas molecules adsorbed at its surface [34]. More appreciable changes of the current density were observed at elevated temperatures. A possible reason for the decrease of the emission current density could be an intensive degassing of the highly porous carbon film. Simultaneously, a noticeable increase in the signal-to-noise ratio was observed, especially in the range of low currents. It should be noted, however, that these temperature variations of the current density refer only to a relatively small region of the IVC around the threshold emission values. Variation of the temperature of the carbon cathode over the entire temperature range studied had only a weak effect on the threshold electric-field strength, which remained below $1.5\ \text{V}/\mu\text{m}$.

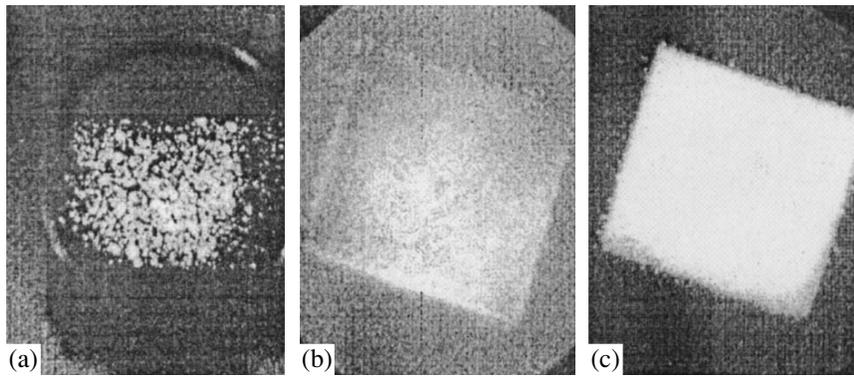


Fig. 5. Luminescence from a cathodoluminescent-phosphor screen under irradiation by electrons emitted from (a) polycrystalline diamond, (b) nanocrystalline diamond, and (c) graphite-like films grown by GPCD on 20×20 mm substrates.

Thus, the above experimental data indicate that the electron emission properties as well as the structural and morphological features of the obtained polycrystalline and nanocrystalline diamond films are in good agreement with the published data. Using the GPCD technique, films of graphite-like carbon material were grown whose structural and morphological properties have not been described earlier. Note that in studies by the same methods of the field electron emission from different samples of graphite and vitreous carbon practically no emission current at field strengths up to $20 \text{ V}/\mu\text{m}$ was detected.

DISCUSSION

The above comparative studies have established that among different carbon materials (from polycrystalline diamond to graphite) produced using the same GPCD technique the best field emission parameters (lowest threshold electric field strength, highest emission current and density of emission centers) have the films consisting mainly of plate-like crystallites, in which carbon atoms are arranged in a typically graphite structure in the form of a few parallel layers oriented predominantly at a normal to the substrate surface. This sort of field electron emission can be due to enhancement of the electric field at the tips on the surface, as it occurs in tip cathodes [7]. However, for the considered carbon films the required enhancement should be no less than 5000, which means that even with small transverse dimensions of the graphite-like tip structures (1–5 nm) their length should be at least a few microns. Direct observations of the film surface using electron microscopy and atomic force microscopy give no evidence of tips of this length.

In addition, with such small-size emitting tips the electric field strength near the surface, which depends on the surface distribution of the free charge carriers, should be appreciably affected by features of the atomic structure of the material. This is especially true of graphite, whose conductivity is determined by a small overlap of the electronic $p\pi$ -orbitals orthogonal

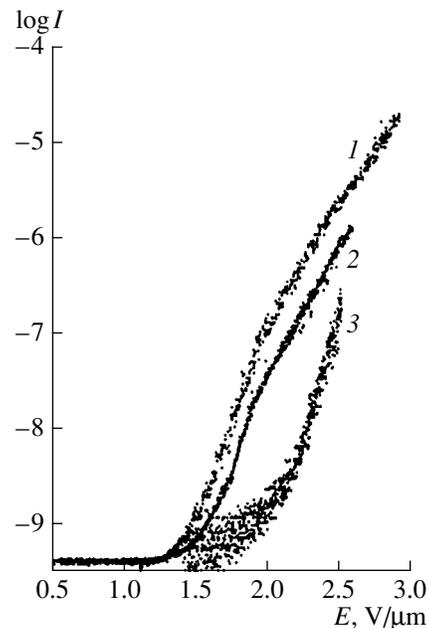


Fig. 6. Current–voltage characteristics of the field emission current from a graphite-like film. The emission current density I is given in mA/cm^2 .

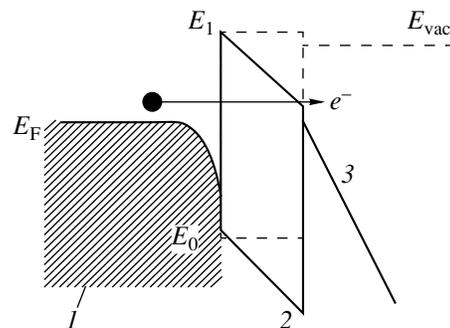


Fig. 7. Energy band diagram of the field electron emitter based on graphite-like film: (1) graphite-like carbon; (2) carbon with modified electronic configuration; (3) vacuum.

to the basal crystallographic planes of graphite. It can be supposed that for the carbon atoms located at the very edge of such planes the electric configuration will be different, resulting in other than sp^2 hybridization of the bonds. At the same time, instead of a continuous spectrum of electric states, as in graphite, a forbidden zone can arise in the electronic spectrum of such atoms. A similar effect is observed in different carbon clusters, with the forbidden zone having a width from 0 to 5 eV [6, 35–37].

The occurrence of the forbidden zone in the density of electric states can explain the features of the CL spectrum of graphite-like films (Fig. 2): the CL band around 3.6 eV could be due to the radiative recombination of the electron–hole pairs taking place in the near-surface region under the action of the electron beam. The long-wavelength CL band in this case can be due to transitions of electrons from the E_1 level to the Fermi level E_F in the graphite-like part of the film.

Whatever the nature of the above bands in the CL spectrum, their very presence indicates the principally important fact of modification of the electric properties of part of carbon atoms comprising graphite-like films. This modification causes, in particular, emergence of energy levels separated by a gap of about 4 eV. It is evident that such atoms are localized on the surface of GPCD films and form an extremely thin layer. In this case the energy band diagram of this film can be schematically represented by a region having the band structure of graphite with an adjacent thin film having a forbidden gap (E_1-E_0) of a width of about 4 eV (Fig. 7). In accordance with the characteristics common to all wide band semiconductors [6–9], it can be expected that the energy level corresponding to an electron in a vacuum, E_{vac} , would be lower than the bottom of the upper band E_2 ; i.e., the electron affinity would be negative. In this case the only reason for the threshold electric-field strength for electron emission to be non-zero would be the need to overcome a potential barrier at the interface between the modified near-surface region and the graphite clusters comprising the rest of the GPCD film. The relatively weak temperature dependence of the observed electron emission, especially at low temperatures (Fig. 6), confirms its field-related nature.

This mechanism can also explain the low-voltage electron emission in other carbon materials with graphite-type structure, such as carbon nanotubes [20, 21]. Indeed, these clusters also consist mainly of sp^2 -coordinated graphite-like carbon. The carbon atoms located at the end of such a tube should have a modified-electron configuration, which may lead to NEA or, at any rate, to a considerable reduction of the work function, thus favoring the electron emission. It should be noted that experimental data give evidence of electron emission only from “open” carbon tubes having at their ends atoms with dangling electronic bonds [20, 21].

An analogous emission mechanism can also be effective in the polycrystalline diamond films, which

always contain a fraction of graphite-like carbon at intercrystalline boundaries. According to tunnel microscopy data, the field electron emission centers in such films are localized precisely at these boundaries [12].

CONCLUSION

In this study of carbon films grown by GPCD an extremely efficient field electron emission has been observed at a threshold electric field strength of about 1.5 V/ μm , with high emission current density (1 mA/cm²) and high density of emission centers (10⁶–10⁷ cm⁻²). These films consist of plate-like crystallites of graphite, whose basal planes are oriented perpendicular to the substrate surface. The surface layer of atoms has a modified electronic configuration, which significantly reduces the electronic work function thus increasing the efficiency of the field electron emission.

ACKNOWLEDGMENTS

The authors are sincerely grateful to V. I. Petrov and E. V. Rakova for their help with this study.

REFERENCES

1. F. J. Himpsel, J. A. Knapp, J. A. van Vechten, and D. E. Eastman, *Phys. Rev. B* **20**, 624 (1979).
2. J. van der Weide, Z. Zhang, P. K. Baumann, *et al.*, *Phys. Rev. B* **50**, 5803 (1994).
3. S. P. Bozeman, P. K. Bauman, B. L. Ward, *et al.*, *Diamond Relat. Mater.* **5**, 802 (1996).
4. N. Eimori, Y. Mori, A. Hatta, *et al.*, *Diamond Relat. Mater.* **4**, 806 (1995).
5. W. Zhu, G. P. Kochanski, S. Jin, and L. Seibles, *J. Appl. Phys.* **78**, 2707 (1995).
6. J. Robertson, *Thin Solid Films* **296**, 61 (1997).
7. V. V. Zhirnov, G. J. Wojak, W. B. Choi, *et al.*, *J. Vac. Sci. Technol. A* **15**, 1733 (1997).
8. E. J. Chi, J. Y. Shim, H. K. Baik, and S. M. Lee, *Appl. Phys. Lett.* **71**, 324 (1997).
9. R. J. Nemanich, P. K. Baumann, M. C. Benjamin, *et al.*, *Diamond Relat. Mater.* **5**, 790 (1996).
10. Z.-H. Huang, P. H. Cutler, N. M. Miskovsky, and T. E. Sullivan, *Appl. Phys. Lett.* **65**, 2562 (1994).
11. L. N. Dworsky, J. E. Jaskie, and R. C. Kane, US Patent No. 5180951.
12. A. V. Karabutov, V. I. Konov, S. M. Pimenov, *et al.*, *J. Phys. IV* **6**, C5-113 (1996).
13. A. A. Talin, L. S. Pan, K. F. McCarty, *et al.*, *Appl. Phys. Lett.* **69**, 3842 (1996).
14. B. S. Satyanarayana, A. Hart, W. I. Milne, and J. Robertson, *Appl. Phys. Lett.* **71**, 1430 (1997).
15. B. L. Weiss, A. Badzian, L. Pillone, and T. Badzian, *Appl. Phys. Lett.* **71**, 794 (1997).
16. A. A. Dadykin, A. G. Naumovets, V. D. Andreev, *et al.*, *Diamond Relat. Mater.* **5**, 771 (1996).

17. A. Y. Tcherepanov, A. G. Chakovskoi, and V. B. Sharov, *J. Vac. Sci. Technol. B* **13**, 482 (1995).
18. A. L. Suvorov, E. P. Sheshin, V. V. Protasenko, *et al.*, *Vide, Couches Minces, Suppl.*, No. 271, 326 (1994).
19. Yu. V. Gulyaev, L. A. Chernozatonskii, Z. Ja. Kosakovskaja, *et al.*, *J. Vac. Sci. Technol. B* **13**, 435 (1995).
20. W. A. De Heer, A. Châtelain, and D. Ugarte, *Science* **270**, 1179 (1995).
21. Y. Saito, K. Hamaguchi, K. Hata, *et al.*, *Nature* **389**, 555 (1997).
22. W. A. De Heer, J.-M. Bonard, K. Fauth, and A. Châtelain, *Adv. Mater.* **9**, 87 (1997).
23. F. S. Baker, A. R. Osborn, and J. Williams, *J. Phys. D* **7**, 2105 (1974).
24. S. Yamamoto, S. Hosoki, S. Fukuhara, and M. Futamoto, *Surf. Sci.* **86**, 734 (1979).
25. I. Yu. Pavlovskii and A. N. Obraztsov, *Prib. Tekh. Éksp.*, No. 1, 152 (1998).
26. B. V. Spitsyn, in *Handbook of Crystal Growth*, Vol. 3: *Thin Films and Epitaxy* (Elsevier, Amsterdam, 1994), p. 402.
27. R. E. Shroder, R. J. Nemanich, and J. T. Glass, *Phys. Rev. B* **41**, 3738 (1990).
28. T. López-Ríos, E. Sandre, S. Leclercq, and E. Sauvain, *Phys. Rev. Lett.* **76**, 4935 (1996).
29. P. Lespade, R. Al-Jishi, and M. S. Dresselhaus, *Carbon* **20**, 427 (1982).
30. R. J. Graham, T. D. Moustakas, and M. M. Disco, *J. Appl. Phys.* **69**, 3212 (1991).
31. A. N. Obraztsov, G. V. Saparin, S. K. Obyden, and I. Yu. Pavlovsky, *Scanning* **19**, 455 (1997).
32. A. N. Obraztsov, H. Okushi, H. Watanabe, and I. Yu. Pavlovskii, *Fiz. Tverd. Tela (St. Petersburg)* **39**, 1787 (1997) [*Phys. Solid State* **39**, 1594 (1997)].
33. Y. Mori, N. Eimori, H. Kozuka, *et al.*, *Appl. Phys. Lett.* **60**, 47 (1992).
34. A. N. Obraztsov, I. Yu. Pavlovsky, and A. P. Volkov, *J. Vac. Sci. Technol. B* **17**, 674 (1999).
35. Yu. V. Gulyaev, N. I. Sinitsyn, and O. E. Glukhova, in *Proceedings of the International Vacuum Microelectronics Conference, IVMC'97, Kyongju, Korea, 1997*, p. 523.
36. J. Robertson and E. P. O'Reilly, *Phys. Rev. B* **35**, 2946 (1987).
37. J. Robertson, *Adv. Phys.* **35**, 317 (1986).

Translated by B. Kalinin

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Ion-Beam Modification of Track Membrane Surface

V. A. Pronin*, V. N. Gornov*, A. V. Lipin*, P. A. Loboda*,
B. V. Mchedlishvili**, A. N. Nechaev**, and A. V. Sergeev**

* Russian Federal Nuclear Center, Zababakhin All-Russia Research Institute of Technical Physics,
Snezhinsk, Chelyabinsk oblast, 456770 Russia

e-mail: p.a.loboda@vniitf.ru

** Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninskii pr. 59, Moscow, 117333 Russia

Received February 2, 2001

Abstract—An ion-beam method to modify the track membrane surface is suggested. An ion gun based on a magnetron sputterer is developed. This gun provides ion energies in the range of 5 eV–1 keV, ion current density up to 0.8 mA/cm², and an ion beam aperture of 90 mm. After the track membrane surface has been irradiated by argon ions with an energy of 50–100 eV, the angle of surface–water contact decreases from 65°–75° to 10°–25°. If the irradiating ion energy is 300–800 eV, the angle of contact increases from 65°–75° to 90°–100°. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

Many important service parameters of membranes used in the separation and purification of biopolymer solutions, such as selectivity, capacity, and possibility of regeneration, to a great extent depend on the wetting ability of the membrane surface [1]. However, the production of polymer-based track membranes (TMs) with a desired wetting ability is sometimes difficult. Therefore, the membrane surface is modified in different ways to make it more hydrophilic. Currently, emphasis is on surface modification by processing in a glow-discharge low-temperature plasma. For example, the TM surface was modified by gaseous monomer polymerization or by processing with argon or oxygen ions with an average energy of 4.5 eV in a high-frequency plasma [2–4].

Of great practical interest is TM modification by argon ions whose energy is controllably varied between 5 eV and 1 keV. Physical desorption from the surface has been observed at an ion energy up to 10 eV; and chemical desorption, from 25 to 50 eV. Thus, at energies below 50 eV, the polymer surface is purified and activated due to physical and chemical desorptions; accordingly, the angle of contact decreases to 30°–40° [5].

At higher ion energies, 0.1–1.0 keV, surface hardening or the formation of a graphitized layer may take place because of partial destruction, break, or cross-linking in polymer linear chains, as well as the generation of active radicals [6].

A topical problem is also TM surface modification by embedding particles or applying a metal layer 10 to 1000 Å thick. These processes are successfully accomplished by ion-beam or ion–plasma processing of materials [7]. Both types of ion processing have been considered in a variety of works, among which [5, 7, 8],

where methods of surface cleaning, ion etching, and ion sputtering for obtaining film coatings have been discussed, are of the greatest interest.

At the same time, ion processing as applied to TM surface modification is poorly studied and has not been considered in those papers. That is why this point calls for special theoretical and experimental (in terms of process equipment and control tools) investigation.

Briefly, the advantages of ion-beam surface modification are high processing uniformity over large areas, high reproducibility, and the possibility of precisely controlling the process. In addition, the method is easy to implement and environmentally safe.

In this work, we apply ion-beam methods and equipment to modify the surface of TMs and study hydrophilic–hydrophobic balance on them at ion energies up to 1 keV.

EXPERIMENTAL EQUIPMENT

Magnetron Sputterer

Ions to modify samples were generated with a magnetron sputterer, a planar disk-shaped magnetron with a target diameter of 100 mm. Depending on the operating mode, the samples were placed 50–150 mm away from the target to ensure the uniform irradiation of the membrane over a circular area of diameter 80 mm. The sample was biased negatively or positively according to which ions, argon or oxygen (or nitrogen), were used for irradiation. In this operating mode of the magnetron, the flux of the sputtered cathode material is negligibly small compared with the ion flux. If it is necessary to irradiate the sample by ions of a gas and a sputtered metal, the magnetron operated in the mode when

ion irradiation and metal deposition were carried out simultaneously.

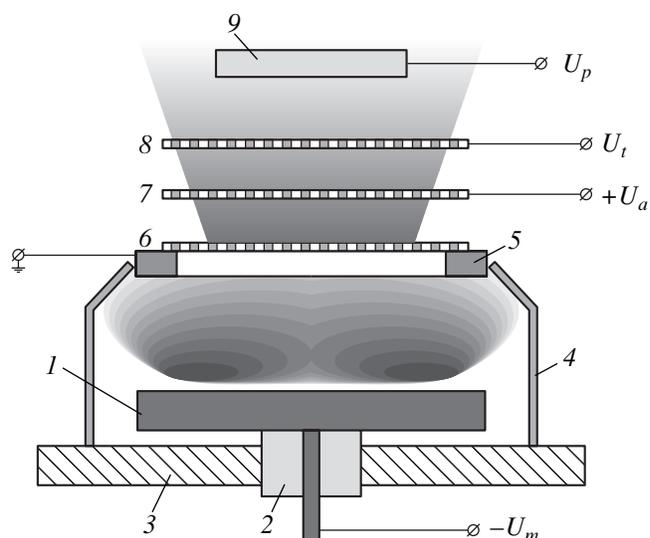
The design of the magnetron allows irradiation by argon, nitrogen, or oxygen ions with an average energy of 300 eV, as well as to apply metal layers of thickness from several hundredths of a millimeter to several millimeters in the ion-deposition mode.

The ion flux density q_i produced by the magnetron source can be varied between 10^{14} and $5 \times 10^{15} \text{ cm}^{-2} \text{ s}^{-1}$. Thus, with such a magnetron source, the surface can be modified by both the desorption of adatoms and the application of metallic coatings.

Ion Gun Built Around the Magnetron Sputterer

The magnetron was used as the basis for developing an ion gun that extracts ions from the near-cathode plasma cloud, accelerates them, collimates the beam, and uniformly irradiates the sample with a diameter of 50 to 80 mm. The ion beam cross section has a diameter of 90 mm. Structurally, the ion-flux-controlling system is a monolithic unit that has three insulated grids arranged in such a way that the ion density is maximal and uniformly distributed within the beam. The unit is mounted and adjusted above the magnetron sputterer.

The design of the magnetron-based ion source is schematically in the figure. After several accelerations



Magnetron-based ion gun; 1, magnetron sputterer; 2, insulator; 3, chamber base; 4, screen; 5, anode; 6, grounded grid; 7, extracting grid; 8, decelerating grid; and 9, sample.

and collisions, electrons knocked out from target 1 by incident ions are collected by grounded anode 5 and grid 6, connected to the anode. Since the radius of grid meshes is smaller than the Debye radius, a resulting plasma cloud is confined to the grid. If a negative poten-

Table 1. Measurements of the wetting angles on ion-beam-modified surfaces

Run no.	I , mA	U , V	U_1 , V	I_1 , mA	M_1	i	i_1	t , min	M_t	ϕ , deg	ϕ_1 , deg	Comments
1	100	280	-100	0.06	W	Ar ⁺	W	3	F-42 Teflon	75	13	W layer is seen
2	100	280	-100	0.06	W	Ar ⁺	W	3	PETP	71	16	"
3	100	280	-100	0.06	W	Ar ⁺	W	3	PETP-I	75	19	"
4	100	280	-100	0.06	W	Ar ⁺	W	3	PETP-TM	69	22	"
5	100	280	-100	0.06	W	Ar ⁺	W	10	F-42 Teflon	77	11	Semitransparent W layer
6	100	280	-100	0.06	W	Ar ⁺	W	10	PETP	68	11	"
7	100	280	-100	0.06	W	Ar ⁺	W	10	PETP-I	73	23	"
8	100	280	-100	0.06	W	Ar ⁺	W	10	PETP-TM	69	23	"
9	100	280	-25	0.4	W	Ar ⁺	W	1	F-42 Teflon	74	37	"
10	150	250	-25	0.5	Ti	Ar ⁺	Ti	3	F-42 Teflon	75	24	"
11	150	260	-25	0.5	Ti	Ar ⁺	Ti	6	F-42 Teflon	74	11	"
12	150	280	-100	1	Ti	Ar ⁺	Ti	20	PETP-I	72	20	"
13	150	280	-100	1	Ti	Ar ⁺	Ti	20	F-42 Teflon	75	10	"
14	150	280	-100	1	Ti	Ar ⁺	Ti	20	PETP	74	17	"
15	150	280	-100	1	Ti	Ar ⁺	Ti	20	PETP-TM	65	19	"
16	150	350	-10	0.5	Al ₂ O ₃	Ar ⁺	Al	1	F-42 Teflon	73	25	"
17	150	280	10	35	Ti	O ⁻ , N ⁻ , e ⁻	Ti	20	F-42 Teflon	74	16	"
18	150	280	10	35	Ti	O ⁻ , N ⁻ , e ⁻	Ti	20	PETP-I	73	23	"
19	150	280	10	35	Ti	O ⁻ , N ⁻ , e ⁻	Ti	20	PETP-TM	70	27	"

Note: I and V , magnetron current and voltage; I_1 , ion current; U_1 , sample potential; M_1 , cathode material; i , type of irradiating particles (ions or electrons); i_1 , attendant atoms and ions; t , irradiation time; M_t , target (sample) material; ϕ and ϕ_1 , wetting angles before and after irradiation, respectively; PETP-TM, PETP-based TM; PETP-I, PETP film preirradiated by silicon ions

Table 2. Wetting angles measured on the TM samples irradiated by the ion-gun-produced Ar ion beam

Sample type	E , eV	J , mA/cm ²	t , min	ϕ , deg	ϕ_1 , deg	ϕ_2 , deg
M 0.05	800	0.85	15	76	35	96
M 0.1				70	45	101
M 0.2				68	68	102
Ar				74	25	102
M 0.05	600	0.65	15	76	35	96
M 0.1				72	52	98
M 0.2				69	57	100
Ar				74	36	95
M 0.05	500	0.55	15	72	45	96
M 0.1				70	63	99
M 0.2				68	60	102
Ar				70	64	98
M 0.05	300	0.55	5	68	32	98
M 0.1				71	60	100
M 0.2				70	58	103
Ar				73	64	101
M 0.05	100	0.3	15	72	44	75
M 0.1				70	67	101
M 0.2				73	98	102
Ar				75	42	65
M 0.05	50	0.12	10	68	15	42
M 0.1				73	18	56
M 0.2				69	17	58
Ar				74	21	48
M 0.05	10	0.08	10	65	20	51
M 0.1				73	17	68
M 0.2				70	19	65
Ar				72	23	63
M 0.05	10	0.03	15	74	16	69
M 0.1				71	13	65
M 0.2				70	18	66
Ar				73	20	58

Note: M 0.05, M 0.1, and M 0.2 are PETP-based TMs with pore diameters of 0.05, 0.1, and 0.2 μm ; Ar, 10- μm -thick PETP-based film preirradiated by an Ar ion beam with a flux density $q_{\text{Ar}} = 1.3 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$; E is the Ar ion energy; J is the ion current density; t is the irradiation time; and ϕ , ϕ_1 , and ϕ_2 are the wetting angles before irradiation, 1 h after irradiation, and 2 days after irradiation.

tial U_a is applied to grid 7, placed behind grid 6, one can extract positive ions from the plasma. These can be noble gas (most often argon) ions or those of a metal sputtered. If grid 7 is biased positively, negative ions can be extracted and directed toward the sample. These can be ions of oxygen, nitrogen, or other gaseous

admixture introduced into argon and capable of generating negative ions. The ions accelerated in the first intergrid gap are decelerated by the field of grid 8, to which a potential U_i is applied. It can be positive or negative relative to the ion sign. By varying the potential applied to grid 8, one can control the ion energy at sample 9. The energy and sign of the ions can additionally be controlled by applying a control potential U_p to the sample.

EXPERIMENTAL STUDY OF MODIFICATION

In experiments, we studied the effect of irradiation by argon and metal ions on the angle of surface–water contact. The measurement procedure has been developed in the Shubnikov Institute of Crystallography, Russian Academy of Sciences [8, 9].

As samples, we used an F-42 teflon film, a poly(ethylene terephthalate) (PETP) film, and a PETP-based track membrane. When argon ion irradiation was combined with irradiation by metal ions, 0.05- to 0.15- μm -thick films of aluminum, titanium, tungsten, or molybdenum were deposited.

In experiments where the ion beam was formed by the magnetron source alone, we observed how the surface condition is affected by nonmonoenergetic ions with energies from 25 to 100 eV, i.e., energies that are sufficient for removing surface contaminants by physical and chemical desorptions.

These experiments showed the decrease in the wetting angle from the initial value 65° – 75° to 10° – 20° when the argon irradiation alone was used. However, the combined processing by argon and metal ions was shown to give the same effect. This means that the combined irradiation does not offer any advantages in terms of the wetting angle.

The processing of the sample surface by electrons and negative nitrogen and oxygen ions produced from residual air in the chamber or from nitrogen and oxygen impurities present in technical-grade argon also decreased the wetting angle to 15° – 25° .

In all the above cases, the attained values of the wetting angle (10° – 25°) did not decrease further upon increasing the irradiation time or the ion energy up to 300 eV. It should be noted that the wetting angles are the smallest for the surface-modified teflon films (10° – 15°), while for the PETP films and PETP-based TMs, the angles equal 15° – 25° . We measured the wetting angles on the samples modified vs. time of exposure to the atmosphere and found that for the teflon and PETP samples, the wetting angle is restored to the initial value 65° – 75° within 5 and 1–2 days, respectively.

The experimental data are listed in Table 1. From this table, it follows that the most optimal parameters for TM surface modification by irradiating by energetically inhomogeneous ions are average energy of about

100 eV, ion flux density $q_{Ar} \sim 10^{15} \text{ cm}^{-2} \text{ s}^{-1}$, and irradiation time of 10–15 min.

Next, we studied PETP-based track membranes irradiated by the argon beam generated in the ion gun and having pores with diameters of 0.05, 0.1, and 0.2 μm , as well as the 10- μm -thick PETP film preprocessed with an argon beam with $q_{Ar} = 1.3 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$. The ion energy was varied between 10 and 800 eV; and the current density, between 0.05 and 0.8 mA/cm^2 which corresponds to irradiation doses of 0.3–0.6 C/cm^2 . The samples were 40 and 60 mm distant from the control grid of the ion gun. The residual pressure in the chamber was kept at 0.1–0.5 Pa. The ion current density near the TMs was measured with a Faraday cup. The associated experimental data are summarized in Table 2.

From Table 2, it follows that when the TMs are processed by argon ions with an energy of 300–800 eV, the surface becomes more hydrophobic. The wetting angle increases from 65°–75° to 90°–100° presumably because of the graphitization of the membrane surface [6].

For ions with lower energies (50–100 eV), conversely, the surface acquires pronounced hydrophilic properties: the wetting angle decreases to 15°–30°. However, within 1 or 2 days of storage in air at room temperature, the wetting angles turn back to the initial value 65°–75° and the surface becomes hydrophobic again.

CONCLUSION

It is suggested to use ion-beam irradiation for the modification of the TM surface by varying its hydrophilic–hydrophobic balance. The magnetron-based ion gun developed generates ions with energies from 5 eV to 1 keV and a current density up to 0.8 mA/cm^2 . The ion beam aperture is 90 mm.

The TM surface processing by energetically inhomogeneous ions with their energy varying between 50

and 100 eV results in the pronounced hydrophilicity of the surface: the angle of surface–water contact decreases from 65°–75° to 10°–25°.

Conversely, the processing of the surface by monoenergetic argon ions with energies of 300–800 eV that are generated by the ion gun enhances the hydrophobicity of the surface: the wetting angle increases from 65°–75° to 90°–100°.

ACKNOWLEDGMENTS

This work was supported by the International Research Center (project no. 918).

REFERENCES

1. A. M. Cherkasov and V. A. Pasechnik, *Membranes and Sorbents in Biotechnology* (Khimiya, Leningrad, 1991).
2. F. Shue, G. Clarotti, J. Sleds, *et al.*, *Makromol. Chem., Macromol. Symp.* **73**, 217 (1993).
3. A. Mass, H. Jaaba, F. Shue, *et al.*, *J. Macromol. Sci., Pure Appl. Chem.* **A34** (1), 67 (1997).
4. F. Shue, G. Clarotti, A. Ait, *et al.*, *Macromol. Rep.*, **A31**, Suppl. 6, 7, 1161 (1994).
5. G. F. Ivanovskii and V. I. Petrov, *Ion-Plasma Materials Treatment* (Radio i Svyaz', Moscow, 1986).
6. Yu. I. Koval', T. B. Borzenko, and V. A. Kudryashev, *Poverkhnost'*, Nos. 10–11, 78 (1994).
7. B. S. Danilin and V. Yu. Kireev, *Low-Temperature Plasma Use for Materials Etching and Purification* (Énergoatomizdat, Moscow, 1987).
8. *Sputtering by Particle Bombardment*, Ed. by R. Behrisch (Springer-Verlag, New York, 1991; Radio i Svyaz', Moscow, 1998), Vol. III, pp. 509–511.
9. *Proceedings of the Russian Conference on Membranes and Membrane Technology "Membranes-95," Moscow, 1995.*

Translated by V. Isaakyan

**EXPERIMENTAL INSTRUMENTS
AND TECHNIQUES**

Parameter Optimization of a Stationary Deepwater Amplitude Cerenkov Muon Detector

V. S. Kinchakov

Computer Center, Far-East Division, Russian Academy of Sciences, Khabarovsk, 680063 Russia

e-mail: kinchakov@as.fe.ru

Received April 4, 2000; in final form, March 26, 2001

Abstract—A condition that should be met in the design of both amplitude and temporal detectors for the muon track azimuth to be uniquely recovered is analytically found. An optimal algorithm for muon track reconstruction with an amplitude Cerenkov detector is elaborated. The optimization involves not only the design of the detector but also statistical methods for estimating auxiliary muon track parameters. An original procedure for computing the confidence interval of the muon track parameters is suggested. This procedure is based on the algorithm used for solving the nonlinear programming problem. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

The role of stationary deepwater muon detectors in locating muon and neutrino sources is difficult to overestimate [1, 2]. For this purpose, a vertical string of four modules is used. However, such a string cannot uniquely reconstruct the muon track [1]. In this work, we show that the muon tracks can be uniquely reconstructed by using asymmetric detectors consisting of no less than six modules. However, the number of the modules per detector should be minimized to reduce the detector cost. Unlike a string, the detecting system suggested in this work consists of six modules arranged at the vertexes of two equilateral triangles. The planes of the triangles are parallel to each other, and their center line is normal to their planes. The triangles are H apart, and the lower one is rotated through an angle of 60° about the center line. The triangles are inscribed into a circle of diameter D .

We succeeded in revealing the general condition for the unique reconstruction of the muon track azimuth. This condition should be met in both the amplitude and temporal detectors. It turned out that the detector must contain no less than three modules with nonzero vertical coordinates without regard to the position of the origin of reference system. In addition, the projections of the modules onto the horizontal plane must not lie on a straight line with a slope of 1.

Along with the muon path reconstruction technique, this work also considers various methods for the estimation of statistical errors introduced into the muon path parameters calculated. The need for error estimation is due to the fact that the signal coming from a photoelectron multiplier (PEM) is of a statistical nature. It is shown that when the similarity method (SM) is combined with the least squares method (LSM), the muon track is reconstructed almost without displacement.

The effect of the detector size on the reconstruction accuracy is studied.

To date, the problem of confidence interval calculation as applied to both amplitude and temporal detectors has not been touched upon in the literature [1, 2]. In this work, we suggest an original procedure to determine the confidence interval for the muon path parameters by directly solving a set of inequalities obtained from the Poisson distribution of the PEM signals. Mathematically, this problem is reduced to a nonlinear programming problem that does not have the general solution algorithm. We elaborated a numerical algorithm to solve this problem and performed appropriate computations.

In our case, as for the case of a string [1], the problem of recovering the coordinates of a relativistic muon from its Cerenkov radiation is solved under the following assumptions: the muon path is a straight line in the recording region and the detector modules are mathematical points.

MUON TRACK RECONSTRUCTION ALGORITHM

The muon path reconstruction method adopted in this work is based on the solution of a set of nonlinear equations derived by comparing the amplitudes of responses from various PEMs of the Cerenkov detector. The amplitude A_i of the response from the i th PEM is given by

$$A_i = n_i \cos \beta_i, \quad (1)$$

where β_i is the angle between the normal to the window of the i th PEM and a Cerenkov photon path and n_i is the flux of Cerenkov photons traveling the distance R_i from

the muon track. The flux n_i is expressed as [2]

$$n_i = n_0 \exp(-R_i/R_0)/(2\pi R_i). \quad (2)$$

Here, R_0 is the distance within which the flux is attenuated e times and n_0 is the number of generated photons per unit muon track length. Let us place the origin at the center of the upper triangle of the modules and direct the z axis vertically. The muon straight path can be described by four independent parameters: Θ and ϕ , zenith and azimuth angles of the direction vector \mathbf{a} of an arbitrary trajectory, respectively; and x_1 and y_1 , coordinates of the point of intersection of the muon path with the plane $z = 0$. Then, the shortest distance spacing between the path and the i th detector is

$$l_i = [(z_i^0)^2(a_x^2 + a_y^2) + (x_1 - x_i^0)^2(a_y^2 + a_z^2) + (y_1 - y_i^0)^2(a_x^2 + a_z^2) + 2(y_1 - y_i^0)a_y(a_z z_i^0 - a_x(x_1 - x_i^0)) + 2a_x a_z(x_1 - x_i^0)z_i^0]^{1/2}. \quad (3)$$

Here, x_i^0 , y_i^0 , z_i^0 are the Cartesian coordinates of the radius vector \mathbf{r}_i^0 that defines the location of the i th PEM. Note that the desired parameters Θ , ϕ , x_1 , and y_1 enter into expression (3) nonlinearly. It should also be stressed that for modules that have the coordinate $z_i^0 = 0$, the replacement $\phi \rightarrow \pi/2 - \phi$ remains the associated l_i unchanged. For modules with $z_i^0 \neq 0$, this statement is valid if

$$y_i^0 = x_i^0 + y_1 - x_1. \quad (4)$$

This means the fundamental indistinguishability of these paths not only by amplitude but also by temporal detectors satisfying Eq. (4), since the reconstruction algorithm for the latter is also based on expression (3). Therefore, any detector that uniquely recovers the parameter ϕ of any muon path must have no less than three modules for which $z_i^0 = 0$ and the projections of these modules onto the plane $z_i^0 \neq 0$ must not satisfy straight-line equation (4).

The symmetry $\phi \rightarrow \pi - \phi$ in the solution of the muon path equation, which has been found for a string [1], fails for the detector under study.

For R_i , we have

$$R_i = [(x_B - x_i^0)^2 + (y_B - y_i^0)^2 + (z_B - z_i^0)^2]^{1/2}, \quad (5)$$

where the Cartesian coordinates of the point emitting a photon that enters a PEM are given by

$$x_B = (a_z x_1/a_x - a_z x_i^0/a_x' + z_i^0 - z_1)/(a_z/a_x - a_z'/a_x'), \quad (6)$$

$$y_B = (a_z y_1/a_y - a_z y_i^0/a_y' + z_i^0 - z_1)/(a_z/a_y - a_z'/a_y'), \quad (7)$$

$$z_B = (a_x z_1/a_z - a_x z_i^0/a_z' + x_i^0 - x_1)/(a_x/a_z - a_x'/a_z'). \quad (8)$$

Here, a_x' , a_y' , and a_z' are the Cartesian coordinates of the direction vector \mathbf{a} for a Cerenkov photon.

Formulas (6)–(8) follows from the equations for the muon and Cerenkov photon trajectories.

The spherical angles Θ' and ϕ' of the photon trajectory are related to the spherical angles Θ and ϕ of the muon trajectory as

$$\cos \Theta' = \cos \Theta \cos \alpha - \sin \Theta \sin \alpha \cos \psi_i^0, \quad (9)$$

$$\sin(\phi' - \phi) = \sin \psi_i^0 \sin \alpha / \sin \Theta', \quad (10)$$

$$\cos(\phi' - \phi) = (\cos \alpha - \cos \Theta \cos \Theta') / (\sin \Theta \sin \Theta'), \quad (11)$$

$$\cos \alpha = 1/n, \quad (12)$$

where n is the refractive index of the medium and ψ_i^0 is the angle between the plane passing through the muon path and the z axis and the plane containing both muon and photon trajectories.

If $\sin \Theta = 0$,

$$\cos(\phi' - \phi) = \cos \psi_i^0. \quad (13)$$

It is easy to check that

$$\cos \psi_i^0 = (-a_y A_2 + a_x B_2) / [(a_x^2 + a_y^2)(A_2^2 + B_2^2 + C_2^2)]^{1/2}, \quad (14)$$

$$A_2 = z_i^0 a_y / a_z + y_1 - y_i^0, \quad (15)$$

$$B_2 = -z_i^0 a_x / a_z - x_1 + x_i^0, \quad (16)$$

$$C_2 = (y_i^0 - y_1) a_x / a_z + (x_1 - x_i^0) a_y / a_z. \quad (17)$$

The sign of $\sin \psi_i^0$ depends on the sign of the mixed product

$$(\mathbf{a}[\mathbf{n}_1 \mathbf{n}_2]) = C_2(a_x^2 + a_y^2)/a_z - a_y B_2 - a_x A_2, \quad (18)$$

(bracketed vectors mean the vector product, and the vector parenthesis means the scalar product), where the vectors \mathbf{n}_1 and \mathbf{n}_2 specify the lines of intersection of the above-mentioned planes with the plane normal to the muon path (\mathbf{n}_1 lies in the plane that contains the muon path and the z axis).

The amplitudes of the responses from the i th and j th PEMs are related as

$$A_i/A_j = R_j \cos \beta_i \exp((R_j - R_i)/R_0)/(R_i \cos \beta_j). \quad (19)$$

The parameters Θ , ϕ , x_1 , and y_1 of the muon path are found by solving the set of nonlinear equations (19). The four necessary equations are obtained as follows. For each muon path, the modules were numbered in order of decreasing number of photons recorded. Taking the ratios of the signals from the first four modules to the signal from the fifth module (in the new numbering), we come to the desired set of equations. It should

be emphasized that, in general, a set of nonlinear equations may have different number of roots: from zero to infinity. Set (19) was solved numerically by the Newton method. As the first approximation, we used a straight line connecting the module of the upper triangle and the module of the lower triangle that recorded the greatest number of photons for a given muon path. With this initial approximation, one may expect that the starting point of the iterative process is near the proper solution. If the solution thus found satisfies new set (19) where the fourth module is replaced by the sixth one, the false solution is excluded. We performed direct test calculations, which showed that this algorithm quickly and uniquely reconstructs the true path. It appears that the quickness of path reconstruction is associated with the proper selection of the initial approximation.

STATISTICAL ESTIMATION OF THE ACCURACY OF MUON TRACK RECONSTRUCTION BY AMPLITUDE DETECTOR

To statistically estimate the accuracy of muon track reconstruction, we used the expression [2] for the average amplitude at the entrance of the i th PEM:

$$\bar{A}_i = C_A \exp(-l_i/(R_0 \sin \alpha))/l_i, \quad (20)$$

where $C_A = 8.57$ (m photoelectron).

This amplitude served as the parameter of the Poisson distribution of the signal A_i at the exit from the i th PEM. Attempts to find the solution of set (19) from the amplitudes A_i recorded sometimes failed. In this case, the estimation was performed either with the SM,

$$\sum_{j=1}^6 (\bar{A}_j - A_j)(1/l_j - 1/(R_0 \sin \alpha)) \frac{\partial l_j}{\partial x_k} = 0, \quad (21)$$

where $x_k = \Theta, \phi, x_1, \text{ or } y_1$ and the summation is over all sixth PEMs (since just this number of PEMs are necessary to uniquely determine the muon track parameters in this design of the detector), or with the modified SM (MSM) where two sets like set (21) are used and the summation is over two fives of PEMs needed for the unique reconstruction of the trajectory. In the latter situation, the trajectory parameters are estimated from the half-sum of the close solutions of these two sets or with the LSM [2] by minimizing the functional

$$S' = \sum_{j=1}^6 (A_{i\text{exp}}^2 - A_{i\text{theor}}^2)^2 / D(A_i^2), \quad (22)$$

$$D(A_i^2) = 4A_i^3 + 6A_i^2 + A_i. \quad (23)$$

The initial approximations for the trajectory parameters to solve the similarity equations were found by random drawing [3] and satisfy the inequalities

$$A_i^{\min} < A_i(\Theta^H, \phi^H, x_1^H, y_1^H) < A_i^{\max}, \quad (24)$$

where the confidence interval limits A_i^{\min} and A_i^{\max} were taken with a confidence coefficient of 0.99 [4].

The confidence coefficient was selected in such a way that it covered the PEM signals with the maximal deviation from the average level. Then, for a detector having less than six modules, the initial approximation may be far from the proper solution and close to other roots of set (19); in this case, we will obtain biased statistical estimates.

To conclude this section, we put forth our method for finding the confidence interval of the muon path parameters. It is based on the solution of the related nonlinear programming problem. Given the confidence coefficient α' and average amplitude \bar{A}_i (20) at the entrance of the i th PEM, one can find [4] the confidence intervals (A_i^{\min}, A_i^{\max}) and, hence, the corresponding intervals (l_i^{\max}, l_i^{\min}) in (20). To find the confidence intervals for the muon path parameters, it is suggested to iteratively look for the maximum of the objective function

$$F = \sum_{i=1}^N [l_i^{\max}(x_j^{\max}, x_j^{\min}) - l_i^{\min}(x_k^{\max}, x_k^{\min})] / [N(l_i^{\max} - l_i^{\min})], \quad (25)$$

where N is the total number of the detector modules, provided that

$$l_i^{\min} \leq l_i^{\max(\min)}(x_j^{\min}, x_j^{\max}) \leq l_i^{\max}, \quad (26)$$

$$j = 1, 2, 3, 4.$$

The relationship between the parameters l_i [see (3)] and the parameters $x_j = \Theta, \phi, x_1, \text{ or } y_1$ of the muon path is nonlinear and nonmonotonic, which greatly complicates the procedure of finding the confidence interval. The nonmonotonicity of the l_i vs. x_j relationship makes us consider all possible sets of the variables x_j^{\max} and x_j^{\min} in inequalities (26) and take those maximizing the objective function F at each of the iteration steps. Basically, such a problem is a problem of nonlinear programming that does not have the general solution algorithm. One can expect that generally an increase in the number N of the detector modules would shrink the confidence interval.

Since the unique reconstruction of the muon track requires the operation of all detector modules, it was assumed that all of them respond to muon transit. Also, each of the modules was assumed to have only one PEM.

RESULTS OF CALCULATION

Unless otherwise stated, the calculations discussed below were performed for the detector of diameter $D = 5$ m

Table 1. Confidence intervals vs. confidence coefficient α' for the muon trajectory parameters for two trajectories

Trajectory no.	α'	F	Θ_{\min} , rad	Θ_{\max} , rad	ϕ_{\min} , rad	ϕ_{\max} , rad	x_1^{\min} , m	x_1^{\max} , m	y_1^{\min} , m	y_1^{\max} , m
1	0.9	0.88	0.064	1.17	0.16	4.83	10^{-5}	13.1	10^{-5}	11.9
	0.99	0.90	0.126	1.57	0.05	1.13	0.73	30.4	0.13	36.4
2	0.9	0.94	10^{-5}	0.86	0.064	4.78	0.28	18.3	10^{-5}	2.80
	0.99	0.82	0.020	1.38	10^{-5}	6.07	10^{-5}	31.6	10^{-5}	24.2

and height $H = 15$ m. We consider first the computed confidence intervals for the trajectory parameters. A program intended for finding the maximum of objective function (25) (provided that inequalities (26) are met) was the complex combination of the gradient method and Monte Carlo method. To verify the program, we specified the parameters x_j^{\min} and x_j^{\max} , from which l_i^{\min} and l_i^{\max} was calculated. Then, the parameters specified were recovered with the method suggested. Table 1 lists the confidence interval limits for two sets of the muon trajectory parameters: (1) $\Theta = 20^\circ$, $\phi = 45^\circ$, $x_1 = 3.0$ m, and $y_1 = 3.0$ m and (2) $\Theta = 10^\circ$, $\phi = 45^\circ$, $x_1 = 1.0$ m, and $y_1 = 2.0$ m. The confidence intervals are relatively wide presumably because of the low amplitudes \bar{A}_i at the entrance to each of the PEMs (hence, A_i^{\min} and A_i^{\max} significantly differ from \bar{A}_i) and the small number of modules in the detector considered.

Next, let us consider the computation of the effective area over which the bundle of trajectories with the parameters Θ and ϕ is reconstructed:

$$S_{\text{eff}} = \int_S P_{(6)}(x_1, y_1, \Theta, \phi) dS, \quad (27)$$

where S is the area normal to the bundle; $P_{(6)}$ is the trajectory detection probability of the detector,

$$P_{(6)} = \prod_{i=1}^6 P_i, \quad (28)$$

and

$$P_i = 1 - \exp(-\bar{A}_i) \quad (29)$$

is the probability of detecting at least one photoelectron by a module.

The effective areas S_{eff} over which the bundle of trajectories with various angles Θ and ϕ are listed in Table 2, and Table 3 gives the track detection probabilities of the detector depending on the parameters Θ and $x_1 = y_1$. The detection probability of the detector is seen to considerably decrease with increasing distance to the detector center, becoming negligibly small for $x_1 = y_1 = 4$ m, and slightly drop with increasing Θ . The effective area S_{eff} of reconstruction of a planar trajectory bundle decreases with an increase in Θ and is almost indepen-

dent of ϕ . The latter result is quite predictable, since the given design of the detector is fairly symmetric with respect to this angle.

Table 4 summarizes the statistical estimates (obtained by the SM, MSM, and LSM) for the parameters of the two real trajectories. First, it should be noted that the reliability of statistical estimates (necessary for qualitative analysis to be valid) is achieved when the sample volume N_u is no less than 100 events (see the third and fourth rows in Table 4). One can argue that the LSM, compared with the SM and the MSM, gives, in general, more biased estimates of all the muon track parameters, excluding ϕ , which is adequately restored by this method for both trajectories. The two other methods restore ϕ much worse. The SM gives less biased estimates of the track parameters compared with the MSM. It was found that detectors of this configuration may provide nearly unbiased estimates for some of

Table 2. Effective area S_{eff} of reconstruction of the planar bundle of muon trajectories vs. Θ and ϕ ($H = 5$ m)

Θ , deg	ϕ , deg	S_{eff} , m ²
10	45	12.1
30	45	10.6
50	45	6.2
10	145	12.1
30	145	10.7
50	145	6.78

Table 3. Probability $P_{(6)}$ of muon track detection vs. parameters Θ and $x_1 = y_1$ of the muon trajectory

Θ , deg	$x_1 = y_1$, m	$P_{(6)}$
10	0.5	0.421
30	0.5	0.293
10	1.0	0.339
30	1.0	0.245
10	2.0	0.152
30	2.0	0.107
10	4.0	0.0132
30	4.0	0.00628

Table 4. Statistical estimates for the parameters Θ , ϕ , x_1 , and y_1 of the muon trajectory and the associated rms deviations σ_Θ , σ_ϕ , σ_{x_1} , and σ_{y_1} for two trajectories reconstructed by the SM, MSM, and LSM

N_u	Θ , deg	ϕ , deg	x_1 , m	y_1 , m	σ_Θ , deg	σ_ϕ , deg	σ_{x_1} , m	σ_{y_1} , m	Note
–	10	45	1.0	2.0	–	–	–	–	Real track parameters
100	7.32	71.9	0.825	2.02	0.96	3.45	0.0806	0.177	SM $H = 5$ m
900	10.5	79.3	1.1	2.9	0.32	0.91	0.021	0.05	MSM $H = 5$ m
100	9.16	78.6	0.997	2.74	1.00	2.90	0.0662	0.153	MSM $H = 5$ m
400	30.4	54.5	0.848	2.36	2.86	5.20	0.135	0.206	LSM $H = 5$ m
–	20	70	3.0	4.0	–	–	–	–	Real track parameters
100	19.2	101.0	1.53	2.21	1.65	5.16	0.114	0.141	SM $H = 5$ m
100	11.9	91.1	1.56	2.20	0.589	2.84	0.986	0.138	MSM $H = 5$ m
200	6.51	92.9	2.81	2.41	0.622	3.51	0.187	0.235	MSM $D = 10$ m
400	39.2	60.4	0.89	1.7	3.18	7.15	0.137	0.207	LSM $H = 5$ m

the real track parameters when the SM and LSM algorithms are combined (cf. the first, second, and fifth rows in Table 4). Basically, one can expect that the bias of the estimates will decrease when increasing number of modules.

For the second muon trajectory (more distant from the detector center), the detection probability (for the detector with $D = 5$ m) is relatively low (Table 3). Here, much larger detectors (the ninth row of Table 3) provide the much less biased estimate of the parameter x_1 and the much smaller rms deviation σ_{x_1} (cf. the eighth and ninth rows). However, the estimate of Θ becomes noticeably worse, although the initial estimate obtained by the SM is nearly unbiased (cf. the sixth and seventh rows). When the detector dimensions grow, the estimate of y_1 improves and that of ϕ remains the same.

CONCLUSION

We considered the asymmetric (nonstandard) configuration of an amplitude detector with the least possible number of modules necessary for the unique reconstruction of the muon track. It was established analytically that the unique recovery of the muon track azimuth is possible if the detector has no less than three modules with the coordinates $z_i^0 \neq 0$ and the coordinates x_i^0 and y_i^0 of the modules do not lie on the straight line given by (4). This statement holds for both amplitude and temporal detectors. The effective simple algorithm for the unique reconstruction of the muon track was elaborated.

Both the design (dimensions) of the detector and the procedure for statistically estimating the muon track parameters recovered were optimized. Direct computations showed that the SM is preferable for statistically estimating the parameters Θ , x_1 , and y_1 . The parameter ϕ of the muon path is best estimated with the LSM. The

detector with $D = H = 5$ m seems to be quite adequate for muon track reconstruction provided that the detection probability for a given track is no less than 0.1 when the SM and the LSM are combined. For trajectories far away from the detector, the bias of the statistical estimates for the parameters x_1 and y_1 can be decreased by increasing the detector dimensions. Basically, the estimates can also be improved by increasing the number of the modules.

The original procedure for finding the confidence interval for the muon track parameters was developed. It is based on the algorithm used to numerically solve the related nonlinear programming problem. Within this approach, the objective function was constructed and the effective algorithm for its maximization (provided that the set of inequalities following from the Poisson distribution of signals applied to PEMs) was found. With this method, the confidence intervals for the muon track parameters were found for two trajectories.

ACKNOWLEDGMENTS

The author thanks V.V. Kobylyanskiĭ for the support and encouragement.

REFERENCES

1. V. P. Pustovetov, Preprint No. 146, FIAN (Lebedev Inst. of Physics, USSR Academy of Sciences, Moscow, 1986).
2. V. I. Aĭnutdinov, I. A. Danil'chenko, and P. Ya. Yashin, in *Proceedings of the 1st All-Union Conference "Investigation of Muon and Neutrino in Large Aqueous Volumes," Alma-Ata, 1983*, p. 195.
3. G. E. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations* (Prentice-Hall, Englewood Cliffs, 1977; Mir, Moscow, 1980).
4. N. L. Bol'shev and N. V. Smirnov, *Tables of Mathematical Statistics* (Nauka, Moscow, 1968).

Translated by V. Isaakyan

EXPERIMENTAL INSTRUMENTS
AND TECHNIQUES

The Thermal State of a Small Spherical Body in an Air Plasma

A. P. Kuryshev and V. V. Sakhin

Baltic State Technical University, St. Petersburg, 198005 Russia

Received September 21, 2000

Abstract—Heat exchange between an air plasma and a low-heat-capacity spherical body made from a refractory (tungsten) or fusible (aluminum) material is studied under conditions when the body exhibits the properties typical of either a thermally thin or thermally thick body. The problem is solved for an unperturbed plasma temperature of 5000–20 000 K, a pressure of up to 1 atm, and a body (particle) size of 10 μm to 1 cm. The model developed accounts for the possible phase transitions at melting and boiling. It is shown that, under proper physical conditions, a refractory particle can relax to a steady thermal state (in various aggregate states). This property can be used in plasma sputtering of refractory coatings, which requires the preparation of a spatially homogeneous thermal state of the sputtered macroparticles that underwent a metal–liquid phase transition.
© 2001 MAIK “Nauka/Interperiodica”.

In [1], heat exchange between an argon plasma and a metallic or dielectric particle emitting thermoelectrons was studied. The efficiency of heat exchange was shown to be affected by electron and ion charge transfer and particle charging. Being of a quantum nature, the heat exchange mechanisms are related to the absorption of plasma electrons by a spherical particle, neutralization of the plasma ions arriving at the particle surface (Auger neutralization), and thermionic emission from the surface. The main objective of that paper was to determine the possibility of increasing the rate of the plasma processing of refractory particles heated up to high temperatures. The particle size was assumed to be consistent with a thermally thin body model. In this case, the thermal state of a particle has no impact on the thermal state of the surrounding plasma. Such a particle represents a body with a low heat capacity.

In this paper, following the physical formulation of the problem similar to that of [1], we study heat exchange between an air plasma and a low-heat-capacity spherical body made from a refractory (tungsten) or fusible (aluminum) material under conditions when the body exhibits the properties typical of either a thermally thin or thermally thick body. In this case, unlike [1], it is necessary to solve the boundary value problem for nonsteady heat conduction. The problem is solved for an unperturbed plasma with a temperature of 5000–20 000 K, a pressure of up to 1 atm, and a particle size of 10 μm .

It is shown that, under proper physical conditions, a refractory particle can relax to a steady thermal state (in various aggregate states). This is due to the switching-on of the two basic cooling mechanisms that limit the temperature growth, namely, radiation and thermionic emission. With an initial density of the total energy flux

onto the body ranging up to 10 MW/m², the major factor limiting the temperature is radiation; at higher energy flux densities, thermionic emission prevails. This fact can be used, in particular, in plasma sputtering of refractory coatings, which requires the preparation of a spatially homogeneous state of the sputtered macroparticles that underwent a metal–liquid phase transition. Conditions under which a fusible particle can relax to a steady state are determined by radiative cooling. Thermionic emission does not contribute significantly until the particle undergoes a liquid–vapor phase transition.

The thermal interaction is governed by electron, ion, neutral, and radiative heat-exchange mechanisms, which depend on the electric field in the double layer region [1]. The interaction between the plasma particles and the body is of a quantum nature. The electrons absorbed by the body give up an energy on the order of the work function, and the ions (in the process of surface Auger neutralization [2]) give up an energy on the order of the difference between the ionization energy and the work function. In both cases, the transferred energy is on the order of several electronvolts. The model under consideration accounts for the possibility of phase transitions at melting and boiling.

The problem is formulated as follows. The thermal state of a spherical body is determined by the solution to the well-known boundary value problem for the time-dependent heat conduction equation

$$\frac{\partial T}{\partial t} = a \left(\frac{\partial^2 T}{\partial r^2} + \frac{2}{r} \frac{\partial T}{\partial r} \right), \quad (1)$$

$$T(0, R) = T_0, \quad (2)$$

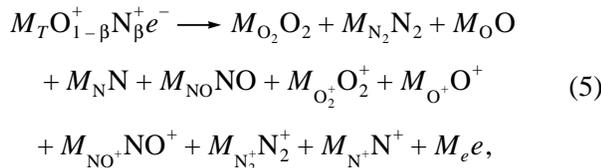
$$T(t, 0) < \infty, \quad (3)$$

$$\lambda \left. \frac{\partial T}{\partial r} \right|_{r=R} = q = \sum q_k. \quad (4)$$

The key point in problem (1)–(4) is boundary condition (4), which determines the energy exchange between the body and the plasma.

The total thermal flux from the plasma onto the body is affected by the charge state of the conducting body receiving energy from the plasma. The thermal and charge states of the plasma and the body are related to each other. Therefore, in the general case, we should solve a self-consistent problem. However, in the model in which a Maxwellian plasma is assumed to be in equilibrium at infinity and under conditions such that the electric field is screened by a thin boundary layer with a size on the order of the Debye length, which is much less than the plasma particle mean free path, the thermal and electrodynamic problems can be separated. When calculating the energy fluxes by the kinetic theory, one can employ the well-known solution to the electrodynamic problem for the distribution functions of electrons, ions, and neutral particles. From the physical standpoint, it is convenient to represent the expressions for the energy fluxes [1] in a clearer form—as the product of the fluxes of charged particles by their mean energy that is transferred to or taken away from the body.

The energy fluxes q_k depend on the plasma composition. The composition of the air plasma is determined by the equation of the model chemical reaction



where M_T is the mole number of the model air “molecules” $O_{1-\beta}^+ N_\beta^+ e^-$ and M_k is the mole number of the k th plasma component.

The equilibrium composition of the unperturbed isothermal plasma at given external plasma parameters (pressure p , temperature T_e , and parameter β) is calculated by solving a system of nonlinear algebraic equations for the molar fractions of the plasma components. The system consists of the equations for conservation of matter (the numbers of O^+ and N^+ ions and electrons), equations for the constants of chemical equilibrium, and the Dalton equation.

For the known plasma composition, the densities of the energy fluxes q_k onto the body can be determined by the following formulas, which are similar to those of [1].

1. For the electrons, we have

$$\begin{aligned} q_e &= J_e^-(2kT_e + e\Phi_b + e\Phi^-) - J_e^+(2kT_s + e\Phi_b), \\ J_e^- &= \frac{1}{4} n_e v_{Te} \exp\left(-\frac{e\Phi^+}{kT_e}\right), \\ J_e^+ &= \frac{B}{e} T_s^2 \exp\left(-\frac{e\Phi_b}{kT_s}\right), \\ v_{Te} &= \sqrt{\frac{8kT_e}{\pi m_e}}, \\ \Phi^\mp &= \pm \varphi_s \eta(\pm \varphi_s), \end{aligned} \quad (6)$$

where J_e^- is the flux of the plasma electrons onto the body, J_e^+ is the flux of the thermoemission electrons specified by the Richardson formula, v_{Te} is the thermal velocity of the plasma electrons, the quantities Φ^\mp accounts for the possible change in the electric field polarity, φ_s is the floating potential of the body, Φ_b is the work function, $e > 0$ is the absolute value of the electron charge, and $\eta(x)$ is the Heaviside step function.

2. For the ions of the k th species, we have

$$\begin{aligned} q_k &= J_k^-(2kT_e + e\Phi^+ + e(I_k - \Phi_b) - 2kT_s), \\ J_k^- &= \frac{1}{4} n_k v_{Tk} \exp\left(-\frac{e\Phi^-}{kT_e}\right), \\ v_{Tk} &= \sqrt{\frac{8kT_e}{\pi m_k}}, \end{aligned} \quad (7)$$

where J_k^- is the flux of the ions of the k th species onto the body, v_{Tk} is the thermal ion velocity, and I_k is the ionization potential of the corresponding plasma component.

3. For the neutral particles, we have

$$q_k = J_k^-(2kT_e - 2kT_s), \quad J_k^- = \frac{1}{4} n_k v_{Tk}, \quad (8)$$

where J_k^- is the flux of the neutral particles of the k th species onto the body and v_{Tk} is the thermal velocity of neutral particles.

Formulas (6)–(8) are derived under the assumption that all the electrons incident on the body are absorbed. The possible thermionic emission and the neutralization of the ions incident on the body (Auger ionization) [2] are also taken into account. Neutral particles are reflected diffusely from the body surface.

The floating potential of the body φ_s is determined from the condition that the ion charge flux is equal to the charge flux of the plasma electrons and thermoemission electrons. The energy flux density of thermal

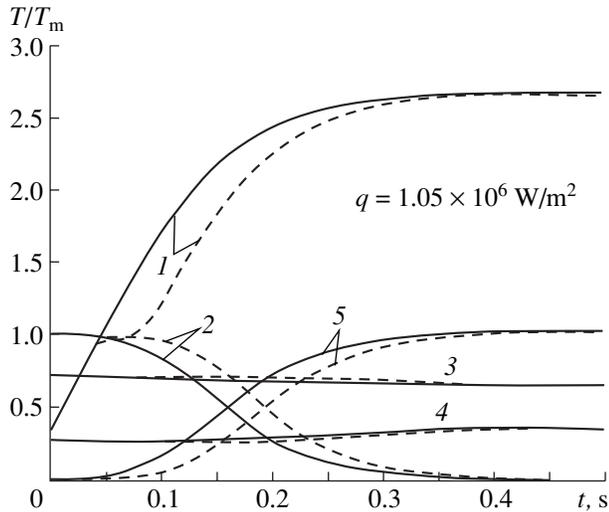


Fig. 1.

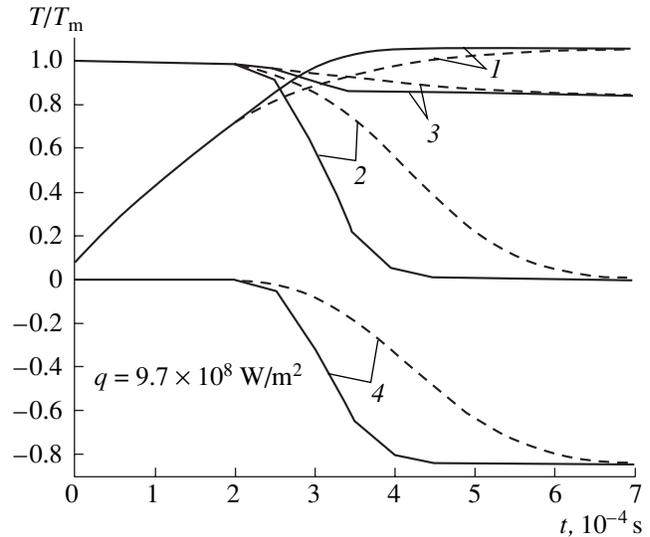


Fig. 2.

radiation from the hot body is given by the Stefan–Boltzmann formula. Phase transitions are incorporated in the mathematical model by introducing an artificial heat capacity, which allows one to account for the phase transition energies.

The boundary value problem (1)–(8) can be solved numerically. However, for a spherical body at $q, \lambda, a = \text{const}$, under physical conditions corresponding to sufficiently large thermal loads, one can also use an analytical solution to problem (1)–(4). This solution has the form

$$T\left(t_1 = \frac{a}{R^2}t, x = \frac{r}{R}\right) = \frac{q}{2\lambda R} \left(-\frac{3}{5} - 4S(x, t_1) + x^2 + 6t_1\right),$$

$$S(x, t_1) = \sum_{k=1}^{\infty} \frac{1}{\lambda_k^2} \frac{\sin(\lambda_k x)}{\sin(\lambda_k)} e^{-\lambda_k^2 t_1}. \tag{9}$$

The spectrum of the problem satisfies the transcendental equation $\tan(\lambda_k) = \lambda_k$, which has the following approximate solution: $\lambda_1 = 4.49340946$, $\lambda_2 = 7.725351837$, $\lambda_3 = 10.90412166$, $\lambda_4 = 14.0661939$,

$$\lambda_k \equiv \alpha_k = \frac{1}{a_k}, \quad \alpha_k = \frac{\pi}{2} + \pi k, \quad k = 5, 6, 7, \dots$$

Using solution (9), we determine the increase in the body temperature ΔT_n over a sufficiently small time step t_{step} at which we can assume $q, \lambda, a = \text{const}$. Then, for the temperature profile we obtain

$$\Delta T_n = T(t_n + t_{\text{step}}, x, q_n) - T(t_n, x, q_n),$$

$$q_n = q(t_n), \quad T(t_{i+1}, x) = T_0 + \sum_{n=1}^i \Delta T_n. \tag{10}$$

Semianalytic solution (9)–(10) has obvious advantages over numerical solutions. This solution allows one to control the validity of the model of a thermally thin or thermally thick body for any point inside the body and substantially reduce the computation time without knowing the exact temperature profile. A comparison between the semianalytic and numerical solutions shows that they agree well under thermal loading. Only when passing to a steady thermal state, in which the total energy flux onto the body vanishes (i.e., upon switching off thermal loading), approximate solution (9)–(10) fails to describe the flattening of the temperature profile over the bulk of a thermally thick body, which is evident from the physical meaning of this solution.

As an example, Figs. 1 and 2 show the calculated time evolution of the energy parameters specifying the character of the plasma–solid energy exchange for a spherical aluminum body with a radius of 0.1 mm at a plasma pressure of 10^{-3} atm and plasma temperature of 10^4 K (Fig. 1) and a tungsten body with the same radius at a plasma pressure of 1 atm and temperature of 2×10^4 K (Fig. 2). Curves 1 show the temperature of the body surface with (dashed) and without (solid) taking into account the solid–liquid phase transition; curves 2 show the total energy fluxes onto the body surface; and curves 3–5 show the contributions from ions, electrons, and radiation, respectively. The contribution of the neutrals to the total energy flux is negligible under the given conditions. The body temperature is normalized to the melting temperature, and the energy fluxes are normalized to the total initial fluxes. As follows from Figs. 1 and 2, the relaxation to the steady thermal state is determined by the switching-on of radiative cooling (for aluminum) or thermionic emission (for tungsten). To distinguish between the curves in Fig. 2, the total energy flux carried by the plasma electrons and thermoemission electrons is assigned the minus sign. Gen-

erally speaking, the phase transition should manifest itself as the straight portion on the curve. The absence of such a portion in Figs. 1 and 2 is related to the smooth switching-on of the artificial heat capacity accounting for the phase transition energy. The phase transition is more pronounced in Fig. 1. The almost-zero initial contribution of the electron component in Fig. 2 is due to the negative floating potential of the body (about -8.5 V), which suppresses electron heat conduction. Under the given conditions, an aluminum body is described by the model of a thermally thin body. For a tungsten body, the temperature difference between the surface and the center is about 200 K.

ACKNOWLEDGMENTS

This work was supported by the Integration Program.

REFERENCES

1. A. G. Gnedovets and A. A. Uglov, *Teplofiz. Vys. Temp.* **29**, 1184 (1991).
2. Y.-H. Ohtsuki, *Charged Beam Interaction with Solids* (Taylor and Francis, London, 1983; Mir, Moscow, 1985).

Translated by N. Ustinovskii

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Control of 3D Structure Growth by the Example of a Biocell

S. L. Grigor'ev

Received February 5, 2001

Abstract—This work, which relates to the area of biotechnology, touches upon the problem of decreasing nanoelectronic and nanooptic components and creating 3D controllable nanomechanic structures. The role of radiation chemical processes in controlling material transport, as well as the growth of membranes and fibers, is demonstrated with a biocell. The fibers are viewed as waveguides that direct control pulses. It is assumed that such waveguides make it possible to produce 3D neural networks. © 2001 MAIK “Nauka/Interperiodica”.

Self-assembling clusters are attracting more and more interest in the field of nanotechnology [1, 2]. Let us consider the effect of wave processes in nanoelectronic technologies by invoking the fundamental principle of wave–particle dualism. It is known that atomic oscillation in a substance may play an essential role in the creation and operation of a nanoelectronic structure. In inorganic crystals under steady-state conditions, thermal atomic oscillations generate standing waves, which do not transfer energy. During light-assisted crystal growth (photon) energy is transferred by wave processes, which significantly affect mass transfer. The energy of electromagnetic waves is used for the synthesis of compound substances in radiation chemistry. While in the nature, a photon flux stimulates photosynthesis, radiation chemistry uses the action of photons for the conformation and polymerization of hydrocarbons [3]. It appears that the life activity and growth of cells in living organisms are also accompanied by wave phenomena [4]. If so, this is the only example of a self-organized cluster system in the nature.

We will consider a living cell as a self-controllable growing structure. The cell (Fig. 1) consists of a membrane, a nucleus, and a cytoplasm that contains dissolved organic compounds and a number of organelles (mitochondria, ribosomes, etc., which are not considered here). The recent advances in genetics have shown that polyatomic molecules of the nucleus, molecules of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), bear hereditary information and control all stages of cell division [5]. Two basic stages in the cell life are distinguished: division and growth. The nuclear structures at these stages differ: before the cell is divided, the DNA molecules twist to form complexes (chromosomes). Cell division begins with the division of chromosomes. During the growth, the DNA complexes untwist and the DNA molecules are uniformly distributed in the nucleus. The chemical mechanism

behind genetic information transfer during the cell division and the mechanism of protein synthesis with the participation of DNA and RNA have been studied in detail [5, 6]. These mechanisms are based on a contact (chemical adsorption) reaction. One can also read

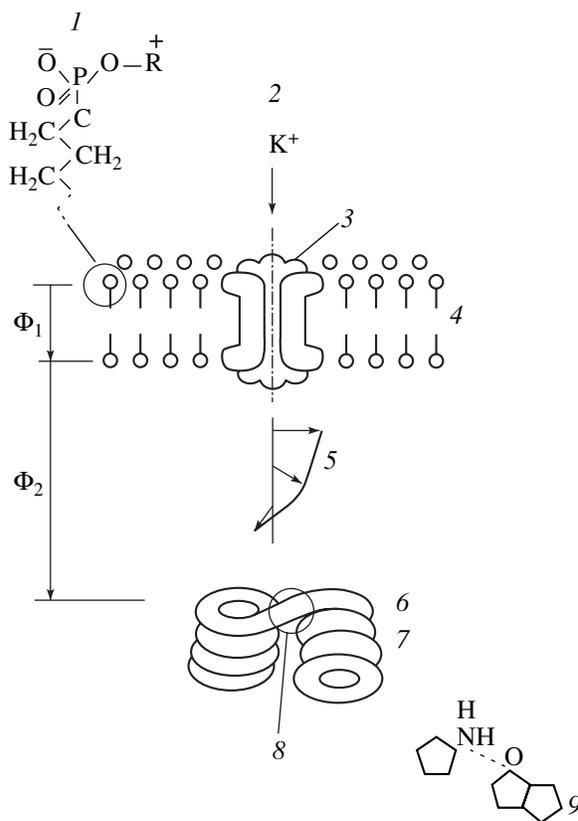


Fig. 1. Schematic representation of the cell: 1, phospholipid; 2, intercellular liquid; 3, membrane protein; 4, lipid layer of membrane; 5, circularly polarized wave (pulse); 6, DNA acid; 7, cell nucleus; 8, cytosine; and 9, guanine.

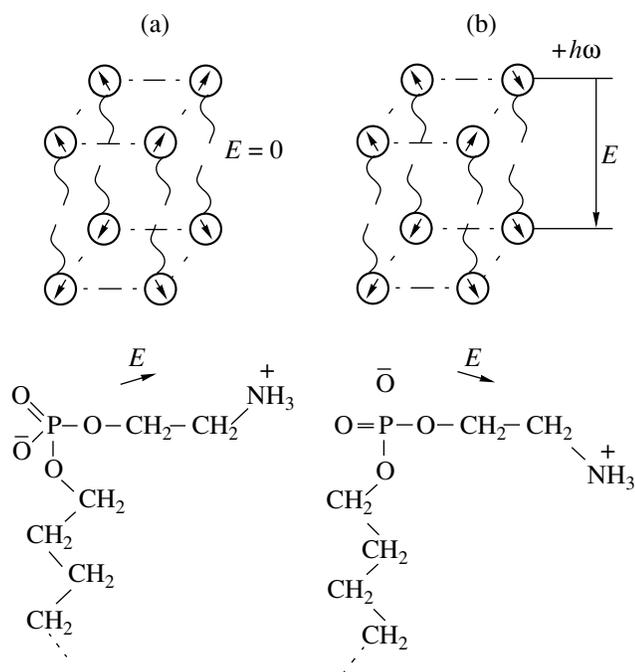


Fig. 2. Model of electromagnetic-pulse-induced (a) conformation and (b) polarization of the phospholipid membrane.

information from the nucleic acids using radiation techniques (X-ray diffraction, laser spectroscopy, etc.), but the interpretation of associated data is very cumbersome. However, within the cell as a self-controllable system, the radiation read-out technique can operate in real time and provide on-line control of the cell activity. Let us consider the control efficiency of this mechanism as applied to material transfer and membrane growth.

We will start with the oscillation properties of DNA. A DNA molecule comprises a chain of atomic groups that twisted into a helix and are bonded by relatively weak hydrogen bonds (Fig. 1). These molecules feature steady-state torsional oscillation with an amplitude up to 1 Å and a frequency in the range of 50–200 cm^{-1} (submillimeter waves) [7]. Moreover, phosphatic groups, nitrogen bases of DNA, and adsorbed proteins possess electric dipole moments [6]. The rotational moments of these dipoles may cause electromagnetic waves. Under steady-state conditions, the diffraction of these waves by a periodic structure generates a standing wave that does not transfer energy [8]. Under unsteady conditions, electromagnetic pulses (solitons) that propagate along the molecular chain may arise [9]. The life activity and growth of the cell are just an unsteady process. Reasons for such unsteadiness may be the evolutionary untwisting of the DNA chains (see the cell life stages described above) or a change in the membrane potential because of a change in the ionic composition of the intra- or intercellular liquid. Consider the latter case (Fig. 1). A change in the potential (ϕ) affects the polar molecular groups of DNA with the resulting

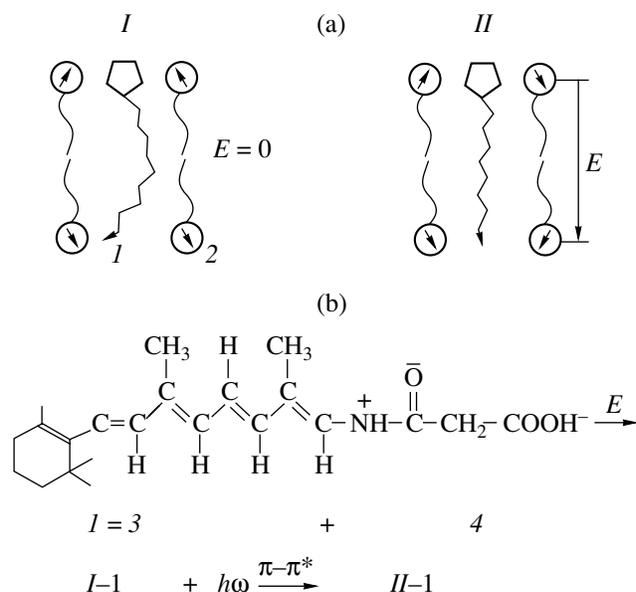


Fig. 3. Electromagnetic-pulse (photon)-induced polarization of the photoreceptor membrane of vertebrates (planar model). (a) Membrane: (I) *cis*- (1, rhodopsin; 2, lipid) and (II) *trans*-configurations. (b) Rhodopsin conformation: (I) rhodopsin, (3) retinol, and (4) glycoprotein; (I) *cis*- and (II) *trans*-configurations.

change in the frequency of the DNA-induced torsional oscillation (the quantum transition known as the Stark effect takes place). The membrane contains globular protein, ion channel (Fig. 1), which also has polar groups with a certain oscillation frequency. Eventually, frequency-modulated coupling between the protein and DNA occurs. This coupling will change the function of the membrane protein and may, for example, restore the normal potential by blocking the channel for potassium ions. If the unsteadiness is due to the DNA structure evolution, the variation of the coupling frequency will cause a change in the steady-state value of the potential. This may encourage material transfer to the cell and, thus, its growth.

The radiation chemical process of lipid membrane growth at the stage of cell growth can be represented as follows. It has been assumed that this process involves special transporting proteins [10]. Also, the dynamic conformational theory of membranes has been put forward [4, 6]. It seems likely that DNA radiation quanta make the groups of polar molecules rotate (the conformation of the molecules takes place), and the membrane becomes locally polarized. Because of this, its electrochemical properties are slightly modified. In particular, the strength of attraction of transporting protein (albumin) cations from the intercellular liquid may alter. Albumin transports molecules of lipids, peptides, and other building blocks of the cell. The model under consideration is consistent with the today's cell physiology concepts [10] and complements them.

At the current stage of our investigation, the behavior of individual molecules in the electromagnetic field generated in the cell is impossible to describe in detail. We can only suggest general rules and elaborate simple models.

The oscillations of the molecules are not arbitrary. They oscillate with a strictly defined frequency, and adjacent molecules oscillate synchronously with each other. For example, a photon pulse from the DNA guanine group (Figs. 1, 2) is in a phase correspondence with the oscillations of the upper, lower, and side phospholipid molecules of the membrane. An elementary electroneutral phospholipid cell consists of eight molecules with the various instantaneous directions and phases of oscillations (Fig. 2). The energy of this pulse will be transferred to the molecule whose oscillation phase differs from that of the pulse (phonon or soliton) to the greatest extent at the instant the pulse meets the cell. Therefore, only one of the molecules will change its configuration and the system will no longer be electroneutral. For example, the phospholipid polar centers, the negative one of the phosphatic group and the positive one of the amino group, will interchange relative to the outer surface of the membrane (Fig. 2b). This causes the local polarization of the membrane. The polarized state is retained until the excitation relaxes via a chemical reaction or heat evolution.

This model resembles the conformation of rhodopsin molecules in the eye photoreceptor [10, 11]. The conformation (bend) of a rhodopsin molecule (Fig. 3) under the action of a photon takes place in the nonpolar retinol part and relaxes with the formation of a polar glycoprotein "tail." Then, the resulting pulse (soliton) propagates along a nerve fiber. Thus, our model of membrane polarization by a radiation quantum applies to various molecules in a wide range of quantum energy (from 0.01 to 1 eV).

The conformation induced by electromagnetic field results in size effects (Fig. 2a). Consequently, it may affect the diffusion through the membrane (through both the lipid layer and the protein ion channels). Protein deformation due to solitons has been studied by Davydov [9].

It should also be noted that compound molecules may relax by passing from one energy level to another; i.e., they operate as frequency converters in electronic devices (for example, rhodopsin converts 1-eV photons to 0.01-eV solitons)

Once the mechanism of local control in the DNA molecule-membrane lipid system with surrounding lipid molecules has been described, we turn to the cell as a whole, which comprises several millions of molecules. For epithelial cells, which have a nearly spherical shape (symmetry), our mechanism applies to any section of the membrane. Other cells, however, have an extended form with a length-to-diameter ratio of 1000 or more (muscular and nerve fibers). Here, a section per nucleus may far exceed the size of the nucleus (the

fibers may have several nuclei). Therefore, the case in point is the radiation mechanism for controlling the growth of 3D structures. The fibers grow largely in the longitudinal direction, acting as waveguides that transfer control actions concentrated at their ends. Thus, a growing cellular structure is capable of transferring control actions at large distances. This fact favors the development of a radiation chemical process for the synthesis of artificial neural networks.

CONCLUSIONS (HYPOTHESES)

(1) DNA molecules of the nucleus and membrane molecules of the cell are coupled by millimetric electromagnetic waves that correspond to the torsional oscillations of the molecules.

(2) Control of the cell life activity is akin to the operation of a feedback generator where the membrane potential is the power supply and a DNA molecule is an element specifying the oscillation frequency.

(3) Electromagnetic waves, photons with an energy of 0.01 eV, polarize molecules of polar dielectrics, which are proteins, lipids, and polypeptides of the cell membrane. The polarization stimulates the electrochemical and radiation chemical synthesis of the cell materials and favors cell growth.

(4) Radiation chemical control allows the growth of 3D fiber structures, such as neural networks, where the fibers guide control actions.

REFERENCES

1. E. S. Soldatov *et al.*, *Usp. Fiz. Nauk* **168**, 217 (1998) [*Phys. Usp.* **41**, 202 (1998)].
2. E. G. Rapis, *Zh. Tekh. Fiz.* **70** (1), 122 (2000) [*Tech. Phys.* **45**, 121 (2000)].
3. L. T. Bugaenko *et al.*, *High-Energy Chemistry* (Khimiya, Moscow, 1988).
4. V. F. Antonov *et al.*, *Biophysics* (Arktos-Vika-Press, Moscow, 1996).
5. V. A. Orekhova, *Medical Genetics* (Vysheishaya Shkola, Minsk, 1997).
6. M. V. Vol'kenshtein, *Biophysics* (Nauka, Moscow, 1981).
7. K. D. Trueblood, in *Accurate Molecular Structures*, Ed. by A. Domenicano and I. Hargittai (Oxford Univ. Press, Oxford, 1992; Mir, Moscow, 1997).
8. V. A. Belyakov and A. S. Sonin, *Optics of Cholesteric Liquid Crystals* (Nauka, Moscow, 1982).
9. A. S. Davydov, *Biology and Quantum Mechanics* (Naukova Dumka, Kiev, 1978).
10. *Human Physiology*, Ed. by R. F. Schmidt and G. Thews (Springer-Verlag, Berlin, 1983; Mir, Moscow, 1996).
11. P. W. Atkins, *Physical Chemistry* (Oxford Univ. Press, Oxford, 1995).

Translated by V. Isaakyan

BRIEF
COMMUNICATIONS

Absorption of Electromagnetic Radiation by a Cylindrical Metal Particle

É. V. Zavitaev, A. A. Yushkanov, and Yu. I. Yalamov

Moscow Pedagogical Institute, ul. Radio 10a, Moscow, 107005 Russia

Received January 25, 2001

Abstract—The cross section of the absorption of electromagnetic radiation by a cylindrical metal particle is calculated. The calculation is performed in the low-frequency limit, in which the contribution of eddy currents to the absorption dominates, and for comparatively small particles (with radii of ≈ 10 nm), which allows us to neglect the skin effect. The case when the mean free path of electrons in the metal bulk substantially exceeds the radius of the cylindrical particle is considered in detail. The specific absorption cross sections for spherical and cylindrical particles are compared. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

The electromagnetic properties of small metal particles may essentially differ from those of massive metallic samples [1]. If the linear size R of a metallic sample is on the order of the electron mean free path Λ or smaller than Λ , the interaction of electrons with the boundary of the metallic sample substantially influences the behavior of the electrons in an external electromagnetic field, hence, the specific optical properties of the sample (metal particle). Therefore, one of the basic optical characteristics, namely, the absorption cross section, shows a nontrivial dependence on the ratio R/Λ for $R < \Lambda$. At room temperature, in metals with a high conductivity (aluminum, copper, silver, etc.), the electron free path Λ lies in the range 10–100 nm. The sizes of particles studied experimentally range within several nanometers; i.e., the case $R < \Lambda$ takes place.

The behavior of electrons in an external electromagnetic field in view of the interaction between the electrons and the boundary of the sample can be described in terms of the classical kinetic theory of conduction electrons in metals [2]. In this case, no constraints are imposed on the relation between the electron free path and the size of the sample.

The equations of macroscopic electrodynamics are valid only in the case of massive samples ($R \gg \Lambda$). Therefore, the well-known Mie theory, explaining the interaction of an electromagnetic wave with metal bodies in the framework of macroscopic electrodynamics, fails to describe the aforementioned size effect.

In [3, 4], the theory of interaction between an electromagnetic radiation and a spherical particle has been developed. Similar results in the limiting case $R \ll \Lambda$ at low frequencies (far-IR region) were found earlier [5, 6]. In the papers cited, the solution to the Boltzmann kinetic equation for conduction electrons in a metal was used. Another approach to the problem has been pro-

posed and is being developed in [7, 8]. Recent interest in the problem of interaction between an electromagnetic radiation and nonspherical particles [9] should be noted. In early articles [10–12], an attempt was made to take into account quantum-mechanical effects in the problem, which is of particular importance at low temperatures.

In this paper, the distribution function describing the linear response of conduction electrons in a metal cylinder to the variable magnetic field of a plane electromagnetic wave is calculated by the kinetic method. Special attention is given to the case when the electron free path substantially exceeds the radius of the cylinder, $R \ll \Lambda$ (free-electron regime). With the distribution function obtained, the dependences of the absorption cross section on the radius and frequency are found and compared with the absorption cross section for a spherical particle.

MATHEMATICAL MODEL AND CALCULATION

Consider a cylindrical particle of a nonmagnetic metal with a radius R and length L ($L \gg R$) in the field of an electromagnetic wave with a frequency ω bounded above by the near-IR range ($\omega < 2 \times 10^{15} \text{ s}^{-1}$). We assume that the direction of the magnetic field in the electromagnetic wave coincides with the cylinder axis. The nonuniformity of the external (wave) field and the skin effect are neglected (we assume that $R < \sigma$, where σ is the skin depth). In the frequency range considered, the contribution of electric dipole polarization currents is small compared to that of eddy currents induced by the external magnetic field [3]. Therefore, the effect of the external electric field is disregarded.

The common physical assumptions are also used: conduction electrons are considered as a degenerate Fermi gas and their behavior in an external variable

magnetic field is described by the Boltzmann equation in the relaxation time approximation. The boundary conditions imply the diffuse reflection of electrons from the inner cylinder surface (i.e., an electron may equiprobably be reflected at any angle between 0° and 90°).

With these assumptions, the absorption of the electromagnetic wave energy can be described as follows. A uniform periodic magnetic field $\mathbf{H} = \mathbf{H}_0 \exp(-i\omega t)$ induces an eddy electric field in the particle. Because of the symmetry of the problem, this field is defined by the Maxwell equation

$$\nabla \times \mathbf{E} = -\left(\frac{1}{c}\right) \frac{\partial \mathbf{H}}{\partial t} \quad (1)$$

and can be represented in the form

$$\mathbf{E} = \frac{1}{2c} \mathbf{r} \times \frac{\partial \mathbf{H}}{\partial t} = \frac{\omega}{2ic} \mathbf{r} \times \mathbf{H}_0 \exp(-i\omega t), \quad (2)$$

where \mathbf{r} is the radius vector (the origin of coordinates is at the axis of the particle) and c is the velocity of light.

The eddy electric field acts on conduction electrons of the particle and causes a deviation f_1 of their distribution function f from the equilibrium Fermi distribution function f_0 :

$$f(\mathbf{r}, \mathbf{v}) = f_0(\varepsilon) + f_1(\mathbf{r}, \mathbf{v}), \quad \varepsilon = \frac{m\mathbf{v}^2}{2}, \quad (3)$$

where \mathbf{v} and m are the velocity and mass of an electron, respectively.

This results in the generation of the eddy current

$$\mathbf{j} = e \int \mathbf{v} f \frac{2d^3(m\mathbf{v})}{h^3} = 2e \left(\frac{m}{h}\right)^3 \int \mathbf{v} f_1 d^3v \quad (4)$$

(h is the Planck constant and e is the electron charge) and also in energy dissipation in the bulk of the particle.

The energy \bar{Q} dissipating per unit time is given by [13]

$$\bar{Q} = e \int \overline{\mathbf{Re} \mathbf{E} \cdot (\mathbf{Re} \mathbf{j})} d^3r = \frac{1}{2} \text{Re} \int \mathbf{j} \cdot \mathbf{E}^* d^3r. \quad (5)$$

Here, the bar denotes time averaging and the asterisk, the complex conjugation. In (4), the standard normalization of the distribution function f , which implies that the density of electron states equals $2/h^3$, is used. Further, the step function [14]

$$f_0(\varepsilon) = \theta(\varepsilon_f - \varepsilon) = \begin{cases} 1, & 0 \leq \varepsilon \leq \varepsilon_f \\ 0, & \varepsilon_f < \varepsilon, \end{cases} \quad (6)$$

where $\varepsilon_f = (m v_f^2)/2$ is the Fermi energy, is employed to approximate the equilibrium function $f_0(\varepsilon)$.

The problem is reduced to finding the deviation f_1 of the distribution function from the equilibrium function f_0 due to the action of eddy field (2). In the linear approx-

imation with respect to the external field, the function f_1 satisfies the kinetic equation [2, 14]

$$-i\omega f_1 + \mathbf{v} \cdot \frac{\partial f_1}{\partial \mathbf{r}} + e\mathbf{v} \cdot \mathbf{E} \frac{\partial f_0}{\partial \varepsilon} = -\frac{f_1}{\tau}. \quad (7)$$

Here, the time dependence [$f_1 \sim \exp(-i\omega t)$] is assumed to be stationary and the collision integral is taken in the relaxation time approximation

$$(df_1/dt)_{\text{col}} = -\frac{f_1}{\tau}, \quad (8)$$

where τ is the relaxation time. The vectors in parentheses mean the scalar product.

In order to uniquely determine the function f_1 , it is necessary to specify its boundary value on the cylindrical surface of the particle. For the boundary condition, we take the diffuse reflection of electrons from the surface [2]:

$$f(\mathbf{r}, \mathbf{v}) = 0 \quad \text{at} \quad \begin{cases} |r_\perp| = R \\ \mathbf{r}_\perp \cdot \mathbf{v}_\perp < 0, \end{cases} \quad (9)$$

where \mathbf{r}_\perp and \mathbf{v}_\perp are the projections of the electron radius vector \mathbf{r} and velocity \mathbf{v} on a plane perpendicular to the cylinder axis.

Solving Eq. (7) by the method of characteristics [15], we obtain

$$f_1 = A(\exp(-vt') - 1)/v, \quad t' \geq 0, \quad (10)$$

where

$$\begin{aligned} \mathbf{v} &= 1/\tau - i\omega, \\ A &= e(\mathbf{v} \cdot \mathbf{E}) \frac{\partial f_0}{\partial \varepsilon} = \frac{e\omega}{2ic} \left(\frac{\partial f_0}{\partial \varepsilon} \right) [\mathbf{v} \times \mathbf{r}] \cdot \mathbf{H}_0 \exp(-i\omega t). \end{aligned} \quad (11)$$

Note that v and A are constant along the path (characteristic). The parameter t' is the time of electron motion along the path from the boundary to a point \mathbf{r} with a velocity \mathbf{v} and is given by

$$t' = \{ \mathbf{r}_\perp \cdot \mathbf{v}_\perp + [(\mathbf{r}_\perp \cdot \mathbf{v}_\perp)^2 + (R^2 - r_\perp^2)v_\perp^2]^{1/2} \} / v_\perp^2. \quad (12)$$

The meaning of t' follows from the following geometrical considerations. Projecting the obvious vector equality $\mathbf{r} = \mathbf{r}_0 + \mathbf{v}t'$, where \mathbf{r}_0 is the radius vector of an electron at the instant of its reflection from the surface of the particle, onto the plane perpendicular to the cylinder axis, we obtain $\mathbf{r}_\perp = \mathbf{r}_{0\perp} + \mathbf{v}_\perp t'$, where the vectors \mathbf{r}_\perp , $\mathbf{r}_{0\perp}$, and \mathbf{v}_\perp are the components of the initial vectors in the projection plane. Having squared both sides of the last equation and solved the resulting equation for t' , we arrive at formula (12).

Relationships (10)–(12) uniquely define the solution f_1 of Eq. (7) with boundary condition (9). This makes it possible to calculate current (4) and dissipated power (5).

Integrals (4) and (5) are convenient to evaluate in cylindrical coordinates both in the coordinate space (r_\perp , α , and r_z ; the Z axis is the polar axis; the vector \mathbf{H}_0 is parallel to the Z axis) and in the velocity space (v_\perp , φ , and v_z ; the v_\perp axis is the polar axis). The cylinder axis coincides with the Z axis. Field (2) in the cylindrical coordinates has only the φ component:

$$\mathbf{E} = E_\varphi \mathbf{e}_\varphi; \quad E_\varphi = \frac{i\omega}{2c} r_\perp H_0 \exp(-i\omega t). \quad (13)$$

Correspondingly, current (4) also has the φ component alone (the lines of the current are closed circles in planes perpendicular to the Z axis with the centers on the Z axis):

$$\begin{aligned} j_\varphi &= 2e \left(\frac{m}{h}\right)^3 \frac{1}{v} \int v_\varphi e(\mathbf{v} \cdot \mathbf{E}) \frac{\partial f_0}{\partial \varepsilon} (\exp(-vt) - 1) d^3 v \\ &= E_\varphi 2e^2 \left(\frac{m}{h}\right)^3 \frac{1}{v} \int v_\varphi^2 \delta(\varepsilon - \varepsilon_f) (1 - \exp(-vt)) d^3 v \\ &= E_\varphi 2e^2 \left(\frac{m}{h}\right)^3 \frac{1}{v} \left(\frac{2}{m}\right) \int_0^{v_f} \int_0^{2\pi} v_\perp^2 \sin^2 \varphi \delta \\ &\quad \times (v_z - \sqrt{v_f^2 - v_\perp^2}) \frac{1}{\sqrt{v_f^2 - v_\perp^2}} (1 - \exp(-vt)) v_\perp dv_\perp d\varphi dv_z \\ &= E_\varphi 2e^2 \left(\frac{m}{h}\right)^3 \frac{1}{v} \left(\frac{2}{m}\right) \int_0^{v_f} \int_0^{2\pi} \frac{v_\perp^3}{\sqrt{v_f^2 - v_\perp^2}} \\ &\quad \times (1 - \exp(-vt)) \sin^2 \varphi dv_\perp d\varphi. \end{aligned} \quad (14)$$

Indeed, using the properties of the δ function, we have

$$\begin{aligned} \delta(\varepsilon - \varepsilon_f) &= \frac{2}{m} \delta(v_z^2 + v_\perp^2 - v_f^2) \\ &= \frac{2}{m} \delta[v_z^2 - (v_f^2 - v_\perp^2)] \\ &= \frac{2}{m} \delta[(v_z - \sqrt{v_f^2 - v_\perp^2})(v_z + \sqrt{v_f^2 - v_\perp^2})] \\ &= \frac{1}{m\sqrt{v_f^2 - v_\perp^2}} [\delta(v_z - \sqrt{v_f^2 - v_\perp^2}) \\ &\quad + \delta(v_z + \sqrt{v_f^2 - v_\perp^2})]. \end{aligned}$$

In view of the symmetry of the problem, integration over the whole range of the velocities v_z is replaced by integration over the positive half-range and the result is doubled. Let us rearrange (12) by introducing the notation $\xi = r_\perp/R$ and using the relation $\mathbf{r}_\perp \cdot \mathbf{v}_\perp = r_\perp v_\perp \cos \varphi$. We find

$$t' = R(\xi \cos \varphi + \sqrt{1 - \xi^2 \sin^2 \varphi})/v_\perp.$$

In terms of the new variables

$$\eta = \xi \cos \varphi + \sqrt{1 - \xi^2 \sin^2 \varphi} \quad \text{and} \quad z = \frac{R}{v_f} \left(\frac{1}{\tau} - i\omega\right), \quad (15)$$

expression (14) takes the form

$$\begin{aligned} j_\varphi &= 2 \left(\frac{m}{h}\right)^3 e^2 E_\varphi \frac{2}{mv} \int_0^{v_f} \int_0^{2\pi} \frac{v_\perp^3}{\sqrt{v_f^2 - v_\perp^2}} \\ &\quad \times \left(1 - \exp\left(-\frac{\eta z v_f}{v_\perp}\right)\right) \sin^2 \varphi dv_\perp d\varphi. \end{aligned} \quad (16)$$

Further, we will consider in detail the case when the field frequency ω and the frequency of electron collisions in the bulk of a metal $1/\tau$ are small compared with the frequency of collisions of electrons with the particle surface. In other words, consider the case $|z| \ll 1$. Then, the expression for the current j_φ can be represented in the form

$$\begin{aligned} j_\varphi &= 2 \left(\frac{m}{h}\right)^3 e^2 E_\varphi \frac{2}{mv} \int_0^{v_f} \int_0^{2\pi} \frac{v_\perp^3}{\sqrt{v_f^2 - v_\perp^2}} \\ &\quad \times \frac{\eta z v_f}{v_\perp} \sin^2 \varphi dv_\perp d\varphi. \end{aligned} \quad (17)$$

Finding the mean dissipated power \bar{Q} by formula (5) and dividing it by the mean energy flux $cH_0^2/8\pi$ in the wave, we obtain the absorption cross section σ :

$$\begin{aligned} \sigma &= \frac{1}{2} \left(\frac{8\pi}{cH_0^2}\right) \int \text{Re}(j_\varphi E_\varphi^*) d^3 r = \frac{1}{2} \left(\frac{8\pi}{cH_0^2}\right) 2 \left(\frac{m}{h}\right)^3 \\ &\quad \times e^2 \left(\frac{i\omega}{2c}\right) H_0 \exp(-i\omega t) \frac{2}{mv} z v_f \left(-\frac{i\omega}{2c}\right) H_0 \exp(-i\omega t) \\ &\quad \times \left[\int_0^{R} \int_0^{2\pi} \int_0^L r_\perp^3 dr_\perp d\alpha dr_z \int_0^{v_f} \int_0^{2\pi} \frac{v_\perp^2}{\sqrt{v_f^2 - v_\perp^2}} \eta \sin^2 \varphi dv_\perp d\varphi \right] \\ &= \frac{2\pi^3 m^2 e^2 \omega^2 L v_f^2 R^5}{h^3 c^3} \int_0^1 \xi^3 d\xi \int_0^{2\pi} \eta \sin^2 \varphi d\varphi. \end{aligned} \quad (18)$$

Integral (18) is calculated in the explicit form. To do this, we make the substitution of the integration variable $\varphi \rightarrow \eta$ in formula (15) and obtain

$$\cos \varphi = \frac{\xi^2 + \eta^2 - 1}{2\eta\xi}, \quad d(\cos \varphi) = \frac{\eta^2 - \xi^2 + 1}{2\eta^2\xi} d\eta. \quad (19)$$

Eventually, the cross section σ is calculated in the

following way:

$$\begin{aligned}
 \sigma &= \frac{2\pi^3 m^2 e^2 \omega^2 L v_f^2 R^5}{h^3 c^3} \int_0^1 \xi^3 d\xi \int_0^{2\pi} (\xi \cos \varphi \\
 &+ \sqrt{1 - \xi^2 \sin^2 \varphi}) \sin^2 \varphi d\varphi = \frac{2\pi^3 m^2 e^2 \omega^2 L v_f^2 R^5}{h^3 c^3} \\
 &\times \int_0^1 \xi^3 d\xi \left[\int_0^\pi (\xi \cos \varphi + \sqrt{1 - \xi^2 \sin^2 \varphi}) \sin^2 \varphi d\varphi \right. \\
 &\left. + \int_\pi^{2\pi} (\xi \cos \varphi + \sqrt{1 - \xi^2 \sin^2 \varphi}) \sin^2 \varphi d\varphi \right] \quad (20) \\
 &= \frac{2\pi^3 m^2 e^2 \omega^2 L v_f^2 R^5}{h^3 c^3} \int_0^1 \xi^3 d\xi \\
 &\times \left[\int_{1-\xi}^{1+\xi} \eta \sqrt{1 - \frac{(\xi^2 + \eta^2 - 1)^2}{4\eta^2 \xi^2}} \left(\frac{\eta^2 - \xi^2 + 1}{2\eta^2 \xi} \right) d\eta \right. \\
 &\left. + \int_{1+\xi}^{1-\xi} \eta \sqrt{1 - \frac{(\xi^2 + \eta^2 - 1)^2}{4\eta^2 \xi^2}} \left(\frac{\eta^2 - \xi^2 + 1}{2\eta^2 \xi} \right) d\eta \right].
 \end{aligned}$$

Then, we change the order of integration:

$$\begin{aligned}
 &\int_0^1 d\xi \left[\int_{1-\xi}^{1+\xi} d\eta (\dots) + \int_{1+\xi}^{1-\xi} d\eta (\dots) \right] \\
 &= 2 \left[\int_0^1 d\eta \int_{1-\eta}^1 d\xi \xi^3 (\dots) + \int_1^2 d\eta \int_{\eta-1}^1 d\xi \xi^3 (\dots) \right] \quad (21) \\
 &= 2 \int_0^2 d\eta \int_{\eta-1}^1 d\xi \xi^3 (\dots).
 \end{aligned}$$

Since

$$\int_{\eta-1}^1 d\xi \xi^3 (\dots) = \int_{1-\eta}^1 d\xi \xi^3 (\dots) = \frac{\eta(4-\eta^2)^{3/2}}{24},$$

we have

$$\begin{aligned}
 \sigma &= \frac{4\pi^3 m^2 e^2 \omega^2 L v_f^2 R^5}{h^3 c^3} \int_0^2 \frac{(4-\eta^2)^{3/2}}{24} \eta \\
 &= \frac{16\pi^3 m^2 e^2 \omega^2 L v_f^2 R^5}{15h^3 c^3}.
 \end{aligned}$$

The electron concentration in a metal is given by

$$n = 2 \left(\frac{m}{h} \right)^3 \frac{4\pi v_f^3}{3};$$

hence,

$$\sigma = \frac{2\pi^2 e^2 L R^5 n \omega^2}{5c^3 m v_f}. \quad (22)$$

DISCUSSION

The classical cross section for a cylinder is calculated in the following way:

$$\begin{aligned}
 \sigma_{11} &= \frac{1}{2} \frac{8\pi}{c H_0^2} \int \text{Re}(j_\varphi E_\varphi^*) d^3 r \\
 &= \frac{1}{2} \frac{8\pi}{c H_0^2} \int (\Sigma_0 E_\varphi)(E_\varphi^*) d^3 r \quad (23) \\
 &= \frac{1}{2} \frac{8\pi}{c H_0^2} \Sigma_0 \frac{i\omega}{2c} H_0 \exp(-i\omega t) \frac{(-i\omega)}{2c} H_0 \exp(-i\omega t)
 \end{aligned}$$

$$\times \int_0^{R/2} \int_0^{R/2} \int_0^{2\pi} r_\perp^3 dr_\perp d\alpha dr_z = \frac{\pi \Sigma_0 \omega^2 R_1^4 2\pi L}{4mc^3} = \frac{\pi^2 e^2 n R_1^4 \tau L \omega^2}{2mc^3},$$

where $\Sigma_0 = (e^2 n \tau)/m$ is the static metal conductivity.

Let us compare the absorption cross sections for a cylinder calculated by the classical way and by the kinetic method. We have

$$\frac{\sigma_1}{\sigma_{11}} = \frac{4R_1}{5\tau v_f}. \quad (24)$$

By analogy, the classical absorption cross section for a sphere is

$$\sigma_{12} = \frac{8\pi^2 e^2 n \tau R_2^5 \omega^2}{15mc^3}. \quad (25)$$

The ratio of the specific (per unit volume) classical absorption cross sections for a cylinder and a sphere of the same radii is equal to

$$\frac{\sigma_{I,II,1}}{\sigma_{I,II,2}} = 1.25. \quad (26)$$

Compare the absorption by spherical and cylindrical particles in the low-temperature case. The estimates for a spherical particle can be obtained from [3]:

$$\sigma_{II1} = \frac{2\pi n e^2 R_1^3 \omega^2}{5mc^3 v_f}, \quad \sigma_{II2} = \frac{\pi n e^2 R_2^3 \omega^2}{5mc^3 v_f}. \quad (27)$$

In this case, the ratio of the specific absorption cross

sections for a cylinder and a sphere of the same radii is

$$\frac{\sigma_{III}}{\sigma_{II2}} = 1.6. \quad (28)$$

Thus, the exact kinetic calculation yields a substantial correction to the classical result.

REFERENCES

1. I. D. Morokhov, V. I. Petinov, L. I. Trusov, *et al.*, *Usp. Fiz. Nauk* **133**, 653 (1981) [*Sov. Phys. Usp.* **24**, 295 (1981)].
2. J. M. Ziman, *Electrons and Phonons* (Clarendon, Oxford, 1960; Inostrannaya Literatura, Moscow, 1962), Chap. 11.
3. A. G. Lesskis, V. E. Pasternak, and A. A. Yushkanov, *Zh. Éksp. Teor. Fiz.* **83**, 310 (1982) [*Sov. Phys. JETP* **56**, 170 (1982)].
4. A. G. Lesskis, A. A. Yushkanov, and Yu. I. Yalamov, *Poverkhnost*, No. 11, 115 (1987).
5. H. J. Trodahl, *Phys. Rev. B* **19**, 1316 (1979).
6. H. J. Trodahl, *J. Phys. C* **15**, 7245 (1982).
7. E. A. Bondar', *Opt. Spektrosk.* **75**, 837 (1983) [*Opt. Spectrosc.* **75**, 503 (1983)].
8. E. A. Bondar', *Opt. Spektrosk.* **80**, 89 (1996) [*Opt. Spectrosc.* **80**, 78 (1996)].
9. P. M. Tomchuk and B. P. Tomchuk, *Zh. Éksp. Teor. Fiz.* **112**, 661 (1997) [*JETP* **85**, 360 (1997)].
10. R. J. Kubo, *J. Phys. Soc. Jpn.* **17**, 975 (1962).
11. L. P. Gor'kov and G. M. Éliashberg, *Zh. Éksp. Teor. Fiz.* **48**, 1407 (1965) [*Sov. Phys. JETP* **21**, 940 (1965)].
12. A. A. Lushnikov, V. V. Maksimenko, and A. Ya. Simonov, *Dispersed Metal Films* (Akad. Nauk USSR, Kiev, 1976), p. 72.
13. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1992; Pergamon, New York, 1984).
14. W. A. Harrison, *Solid State Theory* (McGraw-Hill, New York, 1970; Mir, Moscow, 1972).
15. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2: *Partial Differential Equations* (Interscience, New York, 1962; Mir, Moscow, 1964).

Translated by M. Fofanov

**BRIEF
COMMUNICATIONS**

The Simulation of Two-Beam High-Dose Ion Implantation into Solid Targets

A. F. Komarov

Sevchenko Research Institute of Applied Physical Problems, Minsk, 220064 Belarus

e-mail: kff@rfe.bsu.unibel.by

Received July 27, 2000; in final form, January 18, 2001

Abstract—A physicomathematical model and a BEAM2HD program for the dynamic simulation of one- and two-beam high-dose ion implantation into multilayer and multicomponent targets are developed. The number of target layers is no more than three, and the number of sorts of atoms in each of the layers is no more than seven. The simulation is performed by the Monte Carlo method. Numerical results for the formation of $C_{x \rightarrow 3}N_{y \rightarrow 4}$ superhard layers by two-beam high-dose implantation of nitrogen ions into the $Si_3N_4/C/Si_3N_4/Si$ system are presented. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

Currently, ion range in solids is simulated by two basic methods. One is the Monte Carlo approach [1–4], which considers the motion of each of the ions separately; the other is based on a solution to the integro-differential equations for particle transfer [5–8]. The Monte Carlo method is flexible and allows exact calculations for multilayer and multicomponent targets, including intricately shaped targets. This makes it possible to simulate, for example, advanced VLSI processes or the interaction of a plasma with the vacuum chamber surface. However, most conventional methods for implantation simulation (including those mentioned [1–8] above) apply when the implantation dose is low. Hence, a number of high-dose-related effects, such as implant accumulation, cascade mixing, target sputtering, and target swelling, are ignored. Meanwhile, high-dose implantation has gained widespread acceptance for synthesizing compound materials used in microelectronics and other branches of the industry. Recently, dynamic simulation programs that are an extension of the well-known TRIM [4] and MARLOWE [1] statistical programs have been developed for taking into consideration some of the high-dose effects. A disadvantage of these programs is the impossibility of simulating simultaneous multibeam implantation. It might be expected, however, that the structure rearrangement and the formation of new phases during simultaneous and successive implantation processes radically differ.

In this work, we describe the Monte Carlo-based BEAM2HD program, which lets us simulate ion deceleration in amorphous solids. Specifically, it makes it possible to simulate one- and two-beam high-dose ion implantation into multilayer and multicomponent targets. The number of target layers must not exceed three, and the number of sorts of atoms in each of the layers

must be no more than seven. The potentialities of this program for predicting sputtering yield and implant profiles are discussed, and the effect of high doses on the implant profiles is studied. Sputtering yields are determined with the BEAM2HD program for energies from 1 to 300 keV at the normal incidence of argon ions onto the silicon surface. The 150-keV nitrogen ion concentration across the target is calculated for doses from 4.5×10^{17} to 1×10^{18} cm⁻².

We suggest a model for the formation of $C_{x \rightarrow 3}N_{y \rightarrow 4}$ superhard layers by high-dose two-beam implantation of nitrogen ions into the $Si_3N_4/C/Si_3N_4/Si$ system. Using the BEAM2HD program, we calculate the implant distribution in this multilayer system, determine the erosion depth for each of the layers, and find nitrogen ion energies and doses optimal for the formation of a near-stoichiometric (C_3N_4) layer.

THE MODEL

Scattering and Deceleration

When simulating the deceleration and sputtering of incident ions and recoil atoms, the BEAM2HD program considers the nuclear and electron decelerations separately. For the nuclear sputtering, the binary-collision approximation is used, while the electron deceleration is viewed as a continuous process. The angles of sputtering by nuclei are estimated by the approximate formula [4], and electron energy losses are calculated within the Lindhard–Scharff–Schott theory [13]. The approach suggested includes the high-dose effects [14, 15], namely, target surface sputtering, implant scattering by previously embedded atoms, and target swelling.

Our model gives a chance to describe one- or two-beam implantation. Either beam is characterized by the initial energy, ion weight and charge, dose, and the

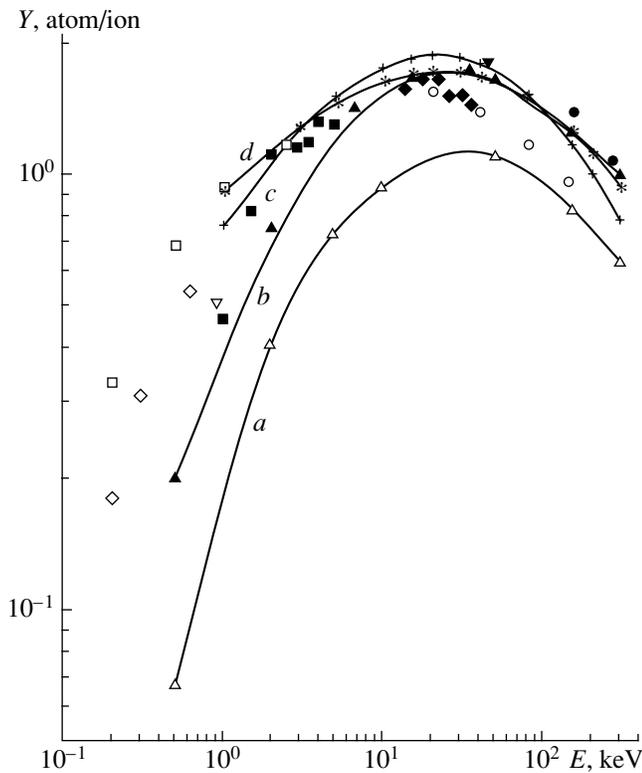


Fig. 1. Calculated and measured sputtering yields for silicon irradiated by normally incident argon atoms. (a, b) Calculated values [12] for $U_s = 7.8$ and 3.9 eV, respectively. (c, d) Results of this work: (c) calculation according to [15, 16] ($U_s = 4.63$ eV); (d) calculation according to [17] ($U_s = 7.8$ eV). Data points: \square [18], \blacksquare [19], \circ [20], \bullet [21], ∇ [22], \blacktriangledown [23], \diamond [24], and \blacklozenge [25].

number of trajectories to be simulated. Each of the trajectories is assigned a “pseudoparticle” that is some integral of the implanted dose by analogy with the TRIDYN [11] and HIDOS [12] programs. The trajectories of either beam are simulated in turn. The target thickness (the region being simulated) is initially divided into 100 equal dynamic layers, and during implantation, a change in the target composition is considered in the frame of approaches elaborated in [11, 12]. At the initial stage, each of the target layers (dynamic, not physical) may contain up to seven different components.

Target Swelling and Sputtering

Target swelling and sputtering during high-dose ion implantation are described in detail in our previous works [14, 15]. If a pseudoparticle (a trajectory being simulated) stops in some deep-seated (dynamic) layer, this layer swells and its thickness is expressed as

$$\Delta(\Delta l) = \frac{\Delta D_{\text{imp}} \times 10^{-16}}{N_{\text{imp}}} (\text{\AA}), \quad (1)$$

where $N_{\text{imp}} (\text{\AA}^{-3})$ is the atomic density of the implant.

If the total dose of this implant is D_{imp} and N_H trajectories are simulated, then the dose $\Delta D_{\text{imp}} = D_{\text{imp}}/N_H$ is assigned to each of the trajectories.

The draw of ion–recoil atom collisions and the stoichiometry change in each of the dynamic layers have been comprehensively described in [15].

The thickness of a layer sputtered from the target surface per pseudoparticle is given by

$$d_{\text{imp}} = \frac{Y(E_0)}{N_{\text{target}}} \Delta D_{\text{imp}} 10^{-16} (\text{\AA}), \quad (2)$$

where $N_{\text{target}} (\text{\AA}^{-3})$ is the atomic density of the target and $Y(E_0)$ is the sputtering yield.

The sputtering yield was estimated from the empirical formula [16, 17], which applies for all the available experimental data in a wide energy range for the normal incidence of the beam. Below, however, we will show that under the assumption of the normal irradiation of the target and a linear cascade, the Sigmund formula for $Y(E_0)$ [18] would be more appropriate for the specific energy interval and ion–atom combinations.

RESULTS AND DISCUSSION

Sputtering of Si by Ar Ions

Figure 1 compares the experimental sputtering yields obtained in [19–26] (see also [12]) with the calculated data obtained with the HIDOS [12] and BEAM2HD programs. In [12], the surface binding energy for silicon was set equal to $U_s = 7.8$ eV. We used the same value for the calculation according to [18] and a value $U_s = 4.63$ eV [17] for the calculation according to [16, 17]. In both cases, the agreement between the calculation and the experiment is the best for high energies. Throughout the energy range, good agreement between the values of Y calculated with the BEAM2HD program and the experimentally found values is observed when the empirical formula from [16, 17] is used. However, when the energy of silicon-sputtering argon ions lies in the interval $5 < E < 50$ keV, Y calculated according to [18] better agrees with the experimental data. As follows from Fig. 1, the HIDOS program [12], which includes complete collision cascades due to a primary ion under the condition $T \geq U_s$ (T is the energy imparted to the target atoms) gives a conservative estimate of the sputtering yield in the low-energy range. At the same time, the empirical formula [16, 17] provides fairly accurate values of Y in the energy range 0.1–300 keV. One can assume that the HIDOS program [12] underestimates the nuclear stopping cross section at low beam energies ($E < 10$ keV), since a decrease in the surface binding energy for silicon from 7.8 to 3.9 eV does not result in satisfactory agreement with the experiment in this energy range (curves a and b in Fig. 1).

High-Dose Implantation of N into Si

The formation of Si_3N_4 buried layers in SOI technology by ion implantation requires high doses of nitrogen [27–34]. In this case, the thickness of the top silicon layer is of crucial importance. If the high-dose effects considerably affect the concentration profiles or the ion range, the ion energy should be adjusted so as to attain a required implant depth.

Figure 2 shows the distributions of nitrogen ions (the energy 150 keV, the doses between 4.5×10^{17} and $1 \times 10^{18} \text{ cm}^{-2}$) in silicon that were obtained with the BEAM2HD program, HIDOS program [12], and Pearson-IV calculation [34]. The experimental data [33] are also presented for comparison. The shape and width of the profiles simulated are seen to be nearly independent of the ion dose. The projected ion range, however, decreases almost linearly with dose in the range of 5×10^{17} – $1 \times 10^{18} \text{ cm}^{-2}$, which we are interested in. At a dose of $1 \times 10^{18} \text{ cm}^{-2}$ (the nitrogen content 65 at. %), the peak of the profile shifts by $\approx 5\%$ (24 nm). The basic reason for such a shift is sputtering. The sputtering yield determined from formula (4) in our earlier work [14] is $Y(E_0) = 0.119$ atom/ion. The peaks of the calculated concentration profiles of nitrogen are roughly 8% higher than those of the HIDOS profiles. The projected ion range obtained in our work ($R_p = 321$ nm for a dose of $1 \times 10^{18} \text{ cm}^{-2}$) slightly differs from ≈ 318 nm reported in [12]. This discrepancy may be due to the different values of the nuclear stopping cross sections, not to the different thicknesses sputtered, because the sputtering yields in [12] and in our work are nearly the same: $Y = 0.12$ and 0.119 , respectively.

High-dose nitrogen implantation changes the stopping power and causes target swelling. It is expected that the projected range and the profile width must decrease with depth because of the progressively increasing atomic density. However, both effects are to a great extent shaded by target swelling (34 nm after irradiation with a dose of $1 \times 10^{18} \text{ cm}^{-2}$).

Figure 2 also compares the calculated (curve 1) and experimental [33] (curve 8) profiles of 150-keV nitrogen ions (for a dose of $4.5 \times 10^{17} \text{ cm}^{-2}$). It is seen that the peaks of the calculated and experimental profiles of nitrogen implanted into silicon are, as a whole, in good agreement. However, the calculated concentration profile of the nitrogen is much narrower and its peak is 25% higher. This discrepancy to some extent can be explained by the too high density of Si_xN_y precipitates taken in the calculations (3.19 g/cm^3 [35]). However, the basic reason for the discrepancy is the fact that none of the programs BEAM2HD, TRIDYN [11], and HIDOS [12] takes into account the radiation-stimulated diffusion of nitrogen implanted into silicon. At so high doses, this diffusion considerably broadens the peak and, accordingly, decreases the volume concentration near the peak. The difference between the analytical profile (curve 9) based on the Pearson-IV distribution

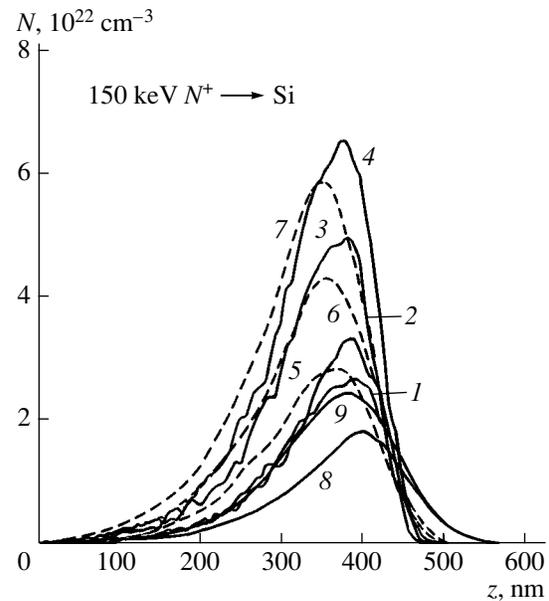


Fig. 2. Nitrogen profiles across the silicon target calculated for various implant doses. Continuous curves are the results of this work for doses (1) 4.5×10^{17} , (2) 5×10^{17} , (3) 7.5×10^{17} , and (4) $1 \times 10^{18} \text{ cm}^{-2}$; dashed curves are the results of [12] for doses (5) 5×10^{17} , (6) 7.5×10^{17} , and (7) $1 \times 10^{18} \text{ cm}^{-2}$; (8) experiment [33] for a dose of $4.5 \times 10^{17} \text{ cm}^{-2}$; and (9) Pearson-IV [34] for a dose of $4.5 \times 10^{17} \text{ cm}^{-2}$.

[34] and the experimental profile (curve 8) is also accounted for by the fact that the simulation ignores radiation-stimulated diffusion.

Thus, the changes in the ion range and in the profile width are small even for a dose of $1 \times 10^{18} \text{ cm}^{-2}$. For this dose, the layer-by-layer sputtering causes surface erosion to a depth of 26 nm. Basically, the results obtained are consistent with those of the theoretical [12] and experimental [33] works. For the more adequate description of the interaction between ion beams and solids during high-dose ion implantation, radiation-stimulated diffusion and the formation of new phases should be taken into account along with target sputtering, target swelling, and a change in the target density.

Simulation of the Formation of C_3N_4 Superhard Layers

We put forward a model for the formation of C_xN_y superhard layers by simultaneous high-dose implantation of nitrogen into the $\text{Si}_3\text{N}_4/\text{C}/\text{Si}_3\text{N}_4/\text{Si}$ multilayer system. The Si_3N_4 layers below and above the carbon film serve as barriers for nitrogen diffusion from this film. In addition, they act as seeds for the growth of the C_3N_4 phase, which has the same crystal lattice as Si_3N_4 . Figure 3 shows the results of simulation for simultaneous high-dose implantation of nitrogen into the $\text{Si}_3\text{N}_4/\text{C}/\text{Si}_3\text{N}_4/\text{Si}$ system. The density of

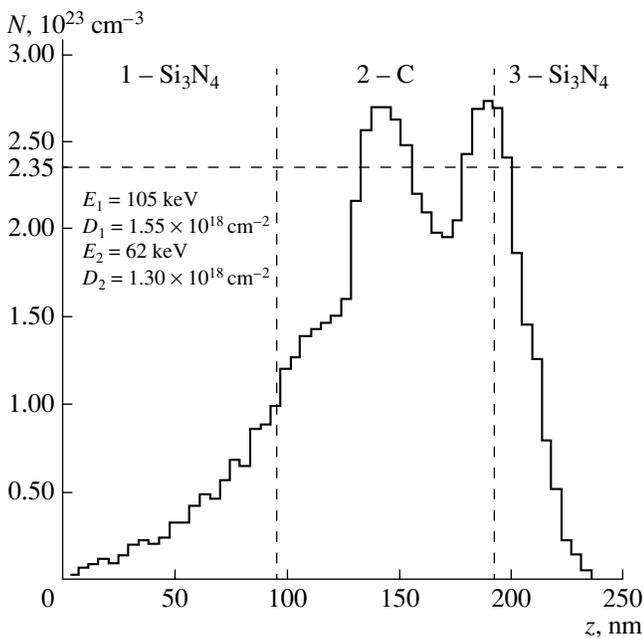


Fig. 3. Depth distribution of nitrogen implanted into the $\text{Si}_3\text{N}_4/\text{C}/\text{Si}_3\text{N}_4/\text{Si}$ multilayer system (two-beam implantation).

Si_3N_4 was taken to be equal to 3.19 g/cm^3 [35]; and that of the carbon film (the diamond structure), to 3.516 g/cm^3 . The thicknesses of the nitride and carbon layers were selected so as to fit the capabilities of our experimental equipment. By simulating the formation of the uniform concentration profile of nitrogen in the carbon film with a near-stoichiometric (4/3) ratio, we solved two problems.

(1) From the solution of the inverse problem [36] (polyenergetic ion implantation), we obtained the approximate values of the energies and doses needed for producing the uniform nitrogen profile with a volume concentration of nitrogen in the carbon film of $2.35 \times 10^{23} \text{ cm}^{-3}$. Two-beam implantation conditions were considered.

(2) With the BEAM2HD program (Fig. 3), we refined the energy and dose for either beam, as well as calculated the net concentration profile of the nitrogen in the carbon film. The energy values were determined from the condition of the complete stopping of the nitrogen ions in the carbon. The integral doses were derived from the condition that the net nitrogen profile in the carbon is uniform and that the nitrogen concentration corresponds to a near-stoichiometric (4/3) component ratio.

As follows from Fig. 3, the total nitrogen concentrations in the peaks slightly exceed $2.35 \times 10^{23} \text{ cm}^{-3}$ (for the given parameters of either beam). One can assume that the peaks will be smeared and decreased if radiation-stimulated diffusion is included. The sputtering of the layers causes surface erosion to a depth of 45 nm at

a net (double-beam) implant dose $D_{\text{imp}} = 2.85 \times 10^{18} \text{ cm}^{-2}$. To take into account carbon deposition from the residual gases in the chamber [37], and, hence, the resulting decrease in the nitrogen ion range, the energy of both beams was taken to be excessively too high. In Fig. 3, the net concentration profile of the nitrogen in the carbon film extends to $\approx 50 \text{ nm}$ to allow for the above factor.

To summarize, we note that the BEAM2HD program used in this work allows the prediction of the sputtering yield and the depth of sputtering-related erosion for each of the layers, as well as the implant profile for one- and two-beam high-dose ion implantation into multilayer and multicomponent targets.

CONCLUSION

A physicomathematical model and a BEAM2HD program for the simulation of one- and two-beam high-dose ion implantation into solid multilayer and multicomponent targets are developed. The program makes it possible to take into account the scattering of implanted ions by previously embedded impurity atoms, target sputtering due to various ion beams, and target swelling. The output data of the program can be represented in the form of tables and graphs that illustrate the distribution of implanted ions in the target, ion ranges, distribution of implantation-induced vacancies, and distributions of the implantation energy losses due to collisions with electrons and nuclei. With the BEAM2HD program, the profiles of the nitrogen atoms in the $\text{Si}_3\text{N}_4/\text{C}/\text{Si}_3\text{N}_4/\text{Si}$ multilayer system are found, the erosion depth for each of the layers is determined, and the nitrogen atom energies and doses optimal for the formation of the inhomogeneous layer close to C_3N_4 in composition are suggested.

ACKNOWLEDGMENTS

This work was partially supported by the Belarusian State Foundation for Basic Research and by the Mianowsky International Foundation (Foundation for Science Support, Warsaw).

REFERENCES

1. M. T. Robinson and M. I. Torrens, *Phys. Rev. B* **9**, 5008 (1974).
2. K. Guttner, H. Ewald, and H. Schmidt, *Radiat. Eff.* **13**, 111 (1972).
3. T. Ishitani, R. Shimizu, and K. Murata, *Jpn. J. Appl. Phys.* **11**, 125 (1972).
4. J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Range of Ions in Solids* (Pergamon, New York, 1985).
5. J. F. Gibbons, *Nucl. Instrum. Methods Phys. Res. B* **22**, 83 (1987).
6. Yu. D. Lizunov and A. I. Ryazanov, *Radiat. Eff.* **60**, 95 (1982).

7. A. F. Burenkov, F. F. Komarov, M. A. Kumakhov, and M. M. Temkin, *Tables of Ion Implantation Spatial Distributions* (Gordon and Breach, New York, 1986).
8. A. F. Burenkov, F. F. Komarov, and M. M. Temkin, *Povorkhnost'*, No. 4, 15 (1989).
9. D. S. Karpuzov, *Phys. Status Solidi A* **94**, 365 (1986).
10. A. M. Mazzone, *Appl. Phys. A: Solids Surf.* **A42**, 193 (1987).
11. W. Möller, W. Eckstein, and J. P. Biersack, *Comput. Phys. Commun.* **51**, 355 (1988).
12. A. Schönborn, N. Hecking, and E. H. Kaat, *Nucl. Instrum. Methods Phys. Res. B* **43**, 170 (1989).
13. J. Lindhard, M. Scharff, and H. E. Schiott, *K. Dan. Vidensk. Selsk. Mat. Fys. Medd.* **33** (14), 1 (1963).
14. A. F. Komarov, F. F. Komarov, P. Zukowski, *et al.*, *Nukleonika* **44**, 363 (1999).
15. A. F. Komarov, F. F. Komarov, A. L. Shukan, *et al.*, *Izv. Akad. Nauk Belarusi, Ser. Fiz.-Tekh. Nauk*, No. 3, 19 (1999).
16. Y. Yamamura, N. Matsunami, and N. Itoh, *Radiat. Eff.* **71**, 65 (1983).
17. N. Matsunami *et al.*, *At. Data Nucl. Data Tables* **31** (1), 2 (1984).
18. P. Sigmund, *Phys. Rev.* **184**, 383 (1969).
19. P. C. Zalm, *J. Appl. Phys.* **54**, 2660 (1983).
20. A. L. Southern, W. R. Willis, and M. T. Robinson, *J. Appl. Phys.* **34**, 153 (1963).
21. P. Blank and K. Wittmaack, *J. Appl. Phys.* **50**, 1519 (1979).
22. E. P. E. Nisse, *J. Appl. Phys.* **42**, 480 (1971).
23. J. M. Poate, W. L. Brown, R. Homer, *et al.*, *Nucl. Instrum. Methods* **132**, 345 (1976).
24. H. H. Andersen and H. L. Bay, *J. Appl. Phys.* **46**, 1919 (1975).
25. N. Laegreid and G. K. Wehnert, *J. Appl. Phys.* **32**, 365 (1961).
26. H. Sommerfeldt, E. S. Mashkova, and V. A. Molchanov, *Phys. Lett. A* **38A**, 237 (1972).
27. P. V. Pavlov, T. A. Kruze, D. I. Tetelbaum, *et al.*, *Phys. Status Solidi A* **36**, 81 (1976).
28. F. F. Komarov, V. S. Solov'yev, V. S. Tishkov, *et al.*, *Radiat. Eff.* **69**, 179 (1983).
29. K. J. Reeson, *Nucl. Instrum. Methods Phys. Res. B* **19/20**, 269 (1987).
30. G. A. Kachurin, V. D. Akhmetov, I. E. Tyschenko, and A. E. Plotnikov, *Nucl. Instrum. Methods Phys. Res. B* **74**, 399 (1993).
31. A. Markwitz, M. Arps, H. Bauman, *et al.*, *Nucl. Instrum. Methods Phys. Res. B* **124**, 505 (1997).
32. S. A. Krivelevich, M. I. Makoviichuk, and E. O. Parshin, *Mikroelektronika* **28**, 363 (1999).
33. M. Yu. Barabanenkov, Yu. A. Agafonov, V. N. Mordkovich, *et al.*, *Nucl. Instrum. Methods Phys. Res. B* **171**, 301 (2000).
34. A. F. Burenkov, F. F. Komarov, M. A. Kumakhov, and M. M. Temkin, *Spatial Distribution of Energy Released in Atomic Collision Cascade in Solids* (Énergoatomizdat, Moscow, 1985).
35. *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology, New Series*, Ed. by K. H. Hellwege (Springer-Verlag, Heidelberg, 1978), Vol. III/7cl.
36. A. F. Burenkov and F. F. Komarov, *Zh. Tekh. Fiz.* **58** (3), 559 (1988) [*Sov. Phys. Tech. Phys.* **33**, 334 (1988)].
37. F. F. Komarov, A. F. Komarov, V. V. Pilko, *et al.*, in *Proceedings of the International Symposium "Ion Implantation in Science and Technology," Lublin, Poland, 1997*, p. 38.

Translated by V. Isaakyan

**BRIEF
COMMUNICATIONS**

Small-Size Systems for Separation of Primary Ion Beams in Microprobe Equipment

S. A. Kuzema and S. N. Mordik

*Institute of Applied Physics, National Academy of Sciences of Ukraine,
ul. Petropavlovskaya 58, Sumy, 40030 Ukraine*

Received September 27, 2000

Abstract—The ion-optical properties and other characteristics of various versions of a double-focusing mass separator are studied. In these designs, the ion transit path is shortened by appropriately selecting the geometry of the ion-optical system. The resolving power of the separating systems is determined with allowance for second-order aberrations. © 2001 MAIK “Nauka/Interperiodica”.

In microprobe equipment where a beam of doubly charged helium ions is used as a primary beam, there appears the need for its purification by removing ion impurities of hydrogen and other gases. The resolving power of an associated mass separator must be about 200. Such a high value is difficult to achieve in conventional achromatic-focusing separators because of a high ion energy spread in the beam obtained from gas-discharge ion sources [1].

In a number of works [2–4], systems with direction and velocity focusing have been suggested. In these double-focusing separators, the optical arms of the energy analyzer are made shorter owing to the use of electrostatic lenses. Additional focusing devices can however be eliminated if ion-optical systems generating a parallel ion beam between the electrical and magnetic stages (see figure) are applied. By appropriately selecting the geometric and physical parameters of the energy and mass analyzers, three out of four optical arms of the separation system can be eliminated.

Let us confirm the aforesaid by considering an ion-optical system comprising a toroidal capacitor (energy analyzer) and a mass analyzer producing a permanent magnetic field.

RADIAL FOCUSING OF THE BEAM

To determine the focusing of an ion beam in the radial plane, we must find the deflection of an ion with arbitrary initial conditions from the central trajectory near the aperture diaphragm of the beam-forming system.

In the linear approximation, the expression for ion beam broadening in the image plane of a double-focusing system has the form

$$y_k = a_{11}y_0 + a_{12}y'_0 + a_{13}e_0 + a_{14}\mu_0, \quad (1)$$

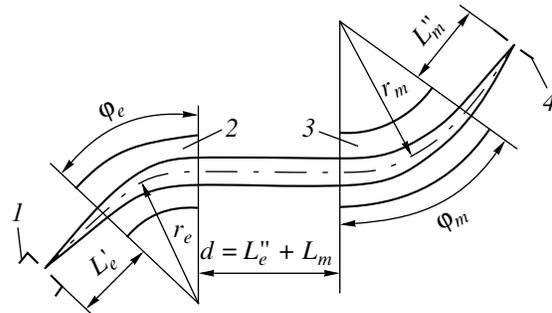
where y_k is the ion displacement in the image plane in units of the central trajectory radius in the mass ana-

lyzer and y_0 , y'_0 , e_0 , and μ_0 are the initial (at the exit from the ion source) displacement of the ion in units of the central trajectory radius, ion direction, and relative changes in its energy and momentum, respectively.

The coefficients a_{ij} entering into expression (1) can be found from the results of [5], where the focusing and dispersion properties of double-focusing systems without intermediate focusing between the electrical and magnetic stages have been estimated for the beam entering and leaving the system normally to the planes of both analyzers.

Since we use a sector nonuniform magnetic field as a magnetic stage, the results of [5] are applicable if the field index n in the expressions for a_{ij} is set equal to zero. Then, we have

$$\begin{aligned} a_{11} = & \pm \frac{r_e}{r_m} (\cos \sqrt{\eta_0} \varphi_e - l_e'' \sqrt{\eta_0} \sin \sqrt{\eta_0} \varphi) \\ & \times (\cos \varphi_m - l_m'' \varphi_m) \pm (-\sqrt{\eta_0} \sin \sqrt{\eta_0} \varphi_e) \\ & \times [\sin \varphi_m + (l_m' + l_m'') \cos \varphi_m - l_m' l_m'' \sin \varphi_m], \end{aligned} \quad (2)$$



Ion-optical system of the mass separator: 1, ion source; 2, energy analyzer; 3, mass analyzer; and 4, aperture diaphragm of the beam-forming system.

$$a_{12} = \pm \frac{r_e}{r_m} \left[\frac{\sin \sqrt{\eta_0} \varphi_e}{\sqrt{\eta_0}} + (l'_e + l''_e) + \cos \sqrt{\eta_0} \varphi_e - l'_e l''_e \sqrt{\eta_0} \sin \sqrt{\eta_0} \varphi_e \right] (\cos \varphi_m - l''_m \sin \varphi_m) \quad (3)$$

$$\pm (\cos \sqrt{\eta_0} \varphi_e - l'_e \sqrt{\eta_0} \sin \sqrt{\eta_0} \varphi_e) \times [\sin \varphi_m + (l'_m + l''_m) \cos \varphi_m - l'_m l''_m \sin \varphi_m],$$

$$a_{13} = \pm \frac{r_e}{r_m} \left(\frac{1 - \cos \sqrt{\eta_0} \varphi_e}{\eta_0} + l''_e \frac{\sin \sqrt{\eta_0} \varphi_e}{\sqrt{\eta_0}} \right) \times (\cos \varphi_m - l''_m \sin \varphi_m) \pm \frac{\sin \sqrt{\varphi_e}}{\sqrt{\eta_0}} \quad (4)$$

$$\times [\sin \varphi_m + (l'_m + l''_m) \cos \varphi_m - l'_m l''_m \sin \varphi_m],$$

$$a_{14} = 1 - \cos \varphi_m + l''_m \sin \varphi_m. \quad (5)$$

Here, $\eta_0 = 2 - (r_e/R_e)$; r_e and R_e are, respectively, the radial and axial radii of the equipotential surface lying in the middle between the electrodes of the toroidal capacitor; φ_e is the angle of deflection of the ions in the electric field; l'_e and l''_e are, respectively, the input and output arms of the energy analyzer in terms of r_e ; r_m is the central trajectory radius in the mass analyzer; φ_m is the angle of deflection of the ions in the magnetic field of the mass analyzer; and l'_m and l''_m are, respectively, the input and output arms of the mass analyzer in terms of r_m .

Let us determine the parameters of an ion beam separator built around a cylindrical capacitor and a mass-analyzing magnet producing a uniform sector magnetic field. To do this, we will make use of the condition for direction focusing of ions in a mass-analyzing magnet producing a uniform field and having a parallel beam at the exit. This condition is known [5, 6] to have the form

$$l'_m = \cot \varphi_m, \quad (6)$$

hence, the output arm $l''_m = 0$ at $\varphi_m = 90^\circ$.

Bearing in mind that the electric field parameter $\eta_0 = 2$ for a cylindrical capacitor, assuming that $l''_e = l'_m = 0$, taking into account that $l''_m = 0$ at $\varphi_m = 90^\circ$, and using the condition for direction focusing of the ions ($a_{12} = 0$), we find for the double-focusing system:

$$l'_e = \frac{1}{\sqrt{2}} \cot \sqrt{2} \varphi_e. \quad (7)$$

The angle of deflection for the ions in the electric field of the energy analyzer is found from the condition for velocity focusing. Taking into account that $e = 2\beta +$

γ and $\mu = \beta + \gamma$ ($\beta = (\Delta V)/V$ and $\gamma = (\Delta m)/m$ are the relative changes in the ion velocity and mass), we write the condition for velocity focusing for ions of the same mass ($\gamma = 0$) in the form

$$2a_{13} + a_{14} = 0. \quad (8)$$

Substituting the expressions for the coefficients a_{13} and a_{14} into (8) and having regard to the fact that $\eta_0 = 2$, $l''_e = l''_m = l'_m = 0$, and $\varphi_m = 90^\circ$, we obtain

$$\pm \frac{2 \sin \sqrt{2} \varphi_e}{\sqrt{2}} + 1 = 0. \quad (9)$$

It follows from (9) that with $\sqrt{2} \varphi_e < \pi$, velocity focusing is provided if the electric and magnetic fields deflect the ions in the opposite directions.

Then, we have

$$\sin \sqrt{2} \varphi_e = \frac{\sqrt{2}}{2} \quad \text{or} \quad \varphi_e = 31^\circ 50'.$$

Substituting this value of φ_e into (7) yields the length of the input arm of the energy analyzer: $l'_e = 1/\sqrt{2}$.

If the radii of the central trajectories of the ions in both fields are the same, for example, 100 mm, the parameters of the separator are as follows: $l'_e = 70$ mm, $\varphi_e = 31^\circ 50'$, $l''_e = 0$, $l''_m = l'_m = 0$, and $\varphi_m = 90^\circ$. It is easy to check that for these values of the parameters, the coefficient a_{11} , which specifies the geometric magnifying power of the system, equals unity.

In an energy analyzer built around a cylindrical capacitor, axial focusing of the ion beam is absent; to improve the aperture ratio of the separator, it is therefore necessary to employ a capacitor for which $R_e \neq \infty$. For a spherical capacitor, for example, $R_e = r_e$ and the electric field parameter $\eta_0 = 1$. Then, the conditions for direction and velocity focusings take the form

$$k_e = \cot \varphi_e, \quad (10)$$

$$\pm 2 \sin \varphi_e + 1 = 0. \quad (11)$$

From condition (11) for velocity focusing, we find the angle of deflection of the ions in the energy analyzer, $\varphi_e = 30^\circ$, bearing in mind that the beam is deflected in the opposite directions in the energy and mass analyzers. For this value of φ_e , we determine the length of the input arm of the energy analyzer from condition (10) for direction focusing: $l'_e = \sqrt{3}$. If it is assumed that $r_e = r_m = 100$ mm, the parameters of the separator are $l'_e = 173.2$ mm, $\varphi_e = 30^\circ$, $l''_e = 0$, $l''_m = l'_m = 0$, and $\varphi_m = 90^\circ$. For these values of the parameters, the coefficient a_{11} , and hence, the geometric magnifying power of the system, equals 0.5.

AXIAL FOCUSING OF THE BEAM

In a separator consisting of a cylindrical capacitor and a mass-analyzing magnet producing a uniform field, axial focusing is absent if the ion beam enters into the system and leaves it normally to the boundaries of the electric and magnetic fields. This is because the axial component of the electric field of the energy analyzer and the radial component of the magnetic field of the mass analyzer are absent in this case. With a spherical capacitor used as an energy analyzer, axial focusing in the electrical stage takes place.

The ion-optical properties of such a separator in the axial plane are convenient to estimate with the expression for the axial deflection of the ion with arbitrary initial conditions from the central path near the aperture diaphragm of the beam-forming system:

$$z_k = b_{11}z_0 + b_{12}z'_0, \quad (12)$$

where z_k is the axial deflection of the ion in the image plane of the two-field system expressed in units of r_m , z_0 is the initial axial displacement of the ion beam in units of r_m , and z'_0 is the initial angle of axial divergence.

The coefficients b_{ij} have been obtained by the matrix method [7]. For the given separator, they are

$$b_{11} = \frac{r_e}{r_m} \cos \varphi_e - \varphi_m \sin \varphi_e, \quad (13)$$

$$b_{12} = \frac{r_e}{r_m} (l'_e \cos \varphi_e + \sin \varphi_e) + \varphi_m \cos \varphi_e - l'_e \varphi_m \sin \varphi_e. \quad (14)$$

Specifically, if the angle of axial divergence of the beam at the exit from the ion source is $z'_0 = 0$ and the exit slit height is $h_0 = 2z_0/r_m = 0.1$ mm, the height of the beam near the aperture diaphragm is $h_k = 2r_m z_k = 0.08$ mm.

SECOND-ORDER ABERRATIONS AND RESOLVING POWER

Let us estimate the resolving power of either version of the double-focusing separator. By definition, the resolving power of a mass separator is given by

$$\frac{m}{\Delta m} = \frac{D_m}{2d}, \quad (15)$$

where D_m is the mass dispersion and d is the beam width near the aperture diaphragm.

The mass dispersion is defined by the coefficient a_{14} : $D_m = (r_m a_{14})/2$. Since $a_{14} = 1$ for both designs of the separator, $D_m = r_m/2$.

According to (1), the beam width near the aperture diaphragm is given by

$$d = 2a_{11}r_m z_0 + 2a_{12}r_m z'_0 + 2(2a_{13} + a_{14})\beta. \quad (16)$$

Since our separator features direction and velocity focusings (that is, $a_{12} = 0$ and $2a_{13} + a_{14} = 0$), $d = 2a_{11}r_m z_0$. Since $2z_0 = S_1/r_m$ (S_1 is the width of the exit slit of the ion source), we have $d = a_{11}S_1$. Then, the resolving power of the separator with allowance for second-order aberrations is given by

$$\frac{m}{\Delta m} = \frac{r_m}{2(2a_{11}S_1 + \sum \text{ab.})}, \quad (17)$$

where $\sum \text{ab.}$ is the sum of second-order aberrations.

To estimate the image broadening due to second-order aberrations, we constructed a nine-dimension matrix of ion transfer from the ion source to the aperture diaphragm in the radial plane and computed the aberration coefficients for the two designs of the separator. For the ion beam parameters $y_0 = 5 \times 10^{-4}$, $y'_0 = 2 \times 10^{-3}$, and $\beta = 1 \times 10^{-2}$, the total contribution of second-order aberrations to the beam width at the exit from the mass separator is $\sum \text{ab.} = 0.037$ mm for the system with the cylindrical capacitor and 0.044 mm for the system with the spherical capacitor. Then, the respective values of the resolving power are 182 and 266 if the exit slit of the ion source is $S_1 = 0.1$ mm.

Thus, the separator based on a spherical capacitor offers the higher resolving power. However, the dimensions of this version are somewhat larger because of the longer arm of the energy analyzer.

With this separator incorporated into the equipment for measuring the ion-optical properties of ion sources [8], researchers could measure the relative content of He^{++} ions in ion beams used in microprobe devices.

REFERENCES

1. V. T. Cherepin, *Ion Probe* (Naukova Dumka, Kiev, 1981).
2. H. Liebl, *Nucl. Instrum. Methods Phys. Res. A* **258**, 323 (1987).
3. T. Matsuo, T. Sakurai, and M. Ishihara, *Nucl. Instrum. Methods Phys. Res. A* **298**, 134 (1990).
4. H. Matsuda, *Nucl. Instrum. Methods Phys. Res. A* **298**, 199 (1990).
5. A. S. Kuzema, O. R. Savin, and I. Ya. Chertkov, *Zh. Tekh. Fiz.* **52**, 1182 (1982) [*Sov. Phys. Tech. Phys.* **27**, 715 (1982)].
6. A. S. Kuzema, O. R. Savin, and I. Ya. Chertkov, *Analizers for Magnetic Mass Spectrometers* (Naukova Dumka, Kiev, 1987).
7. S. Penner, *Rev. Sci. Instrum.* **32**, 150 (1961).
8. A. Kalinichenko, V. Khomenko, S. Lebed, *et al.*, *Nucl. Instrum. Methods Phys. Res. B* **122**, 274 (1997).

Translated by V. Isaakyan

**BRIEF
COMMUNICATIONS**

The Effect of Clay Morphology on Water Relaxation

I. V. Lunev, R. R. Nigmatullin, A. Yu. Zavidonov, Yu. A. Gusev, and I. R. Manyurov

Kazan State University, ul. Lenina 18, Kazan, 420008 Tatarstan, Russia

e-mail: Ivan.Lounev@ksu.ru

Received October 27, 2000

Abstract—The frequency dependence of the permittivity of water in calcium kaolinite (clay) is measured. It is shown that two mechanisms contribute to dipole relaxation of water. One refers to water in the free volume of pores in the clay. The other is associated with bound water covering the porous surface. Experimental data are treated in terms of a fractal model of the medium. The frequency dependence of the permittivity in a wide range of water content in the clay is accounted for theoretically. © 2001 MAIK “Nauka/Interperiodica”.

The physical effects observed when water interacts with clay materials play a decisive role in construction engineering, geodesy, oil-extracting industry, and other branches of the industry. Information on the interaction between near-surface water molecules and the molecular structure of alumina is of crucial importance. Water-saturated clays containing a great amount of bound H₂O are widely used in construction engineering. The effect of clay saturation by large amounts of water is also exploited in many domains of technology. Yet, clay–water interaction at the molecular and mesoscopic levels has been poorly understood.

Dielectric spectroscopy [1] provides important information on water–alumina interaction at the mesoscopic level. We obtained the low-frequency spectra in the clay–water system for different H₂O concentrations (12, 33, 52, 75, and 92%) at 22°C. Clay (calcium kaolinite) was made in the Laboratory of Disperse Systems (Kiev, Ukraine) headed by academician Ovcharenko [2]. For measurements, a 0.2-mm-thick pellet made by pressing was placed in a vacuum chamber preevacuated for 48 h. Saturation by water was carried out in a desiccator arranged over saturated salt solutions for 48 h. The measurements were performed with the Schlumberger low-frequency dielectric spectrometer in the frequency range of 10^{−4}–10⁵ Hz using a double-electrode titanium measuring cell with an inner diameter of 30 mm. The temperature of the cell was thermostatically controlled with an accuracy of ±0.1°C.

Figure 1 shows the frequency dependence of the imaginary part of the permittivity. These data will be interpreted within a model of fractal medium [3]. In this model, a conducting medium (water) filling pores in clay is described in terms of recap (resistance + capacitance) elements. A recap is a self-similar RC network (Fig. 2) comprising Foster circuits [3]. To produce a recap with an impedance in the form $Z(j\omega) = C_\nu(j\omega)^{-\nu}$

($0 \leq \nu \leq 1$), the components of a self-similar RC network must obey the relationships

$$\frac{R_n}{R_{n+1}} = a, \quad \frac{C_n}{C_{n+1}} = b,$$

where a and b are frequency-independent constants.

Below, we will show that clay contains bound water and water filling the free volume. Hence, the equivalent circuit of the medium can be represented as two parallel-connected recaps (Fig. 3). The conductivity of the resulting recap is given by [3]

$$G(\omega) = C_1(j\omega)^n + C_2(j\omega)^m, \quad (1)$$

where C_1 , C_2 , n , and m are constants.

These constants depend on the ratio $\ln a/\ln b$ and define the self-similarity (fractality) of the medium. To

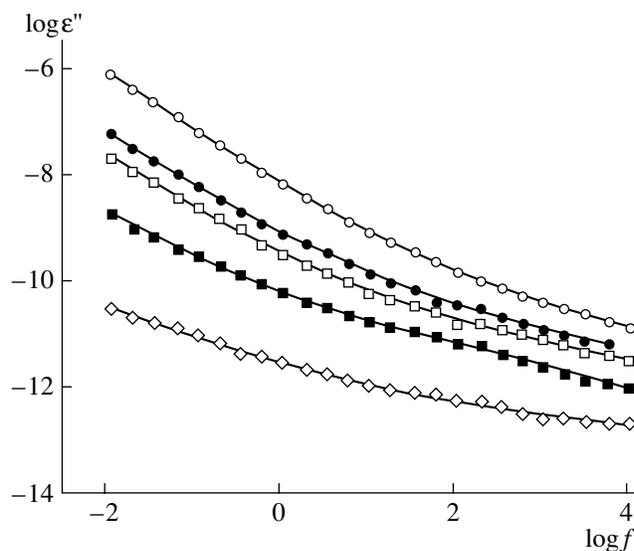


Fig. 1. Imaginary part of the permittivity vs. frequency for a water content of (○) 92, (●) 75, (□) 53, (■) 33, and (◇) 12%. Continuous curves are obtained analytically.

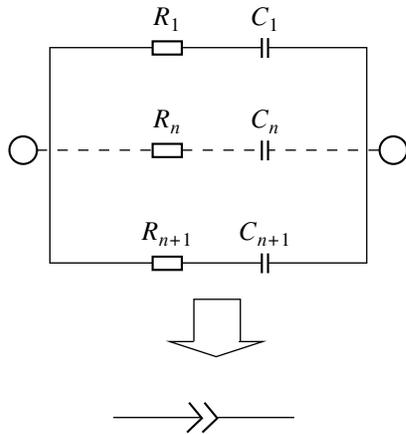


Fig. 2. Recap element formed by self-similar Foster circuits.

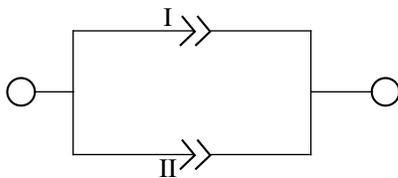


Fig. 3. Equivalent circuit of the clay–water system. I, bound water; II, water in the free volume.

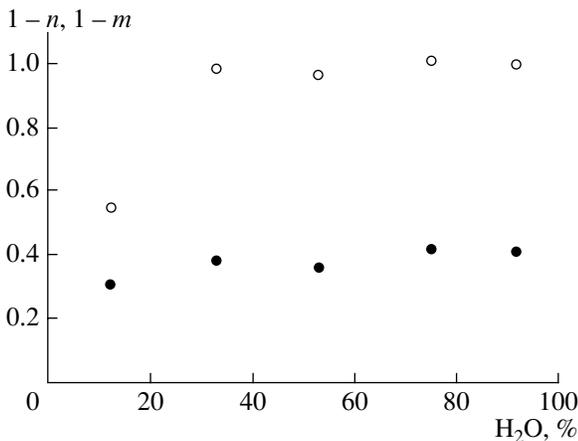


Fig. 4. Exponents (○) $1 - n$ and (●) $1 - m$ vs. water content in the sample.

calculate these parameters, a detailed fractal model is needed (in this work, it is omitted). Therefore, we will calculate them by approximating the experimental data with theoretical formulas using the least squares method (the first approximation). Following the frequency dependence of the imaginary part of the permittivity ϵ'' (Fig. 1), we relate the conductivity $G(\omega)$ and $\epsilon''(\omega)$ as

$$\epsilon''(\omega) = -\text{Re} \left[\frac{jG(\omega)}{G_0\omega} \right], \quad (2)$$

where C_0 is the capacitance of a spectrometer cell.

In view of formulas (1) and (2), the expression for $\epsilon''(\omega)$ can be written as

$$\epsilon''(f) = \frac{A}{f^{1-n}} + \frac{B}{f^{1-m}}, \quad (3)$$

where

$$A = \frac{C_1 \exp(\pi n/2)}{C_0 (2\pi)^{1-n}}, \quad B = \frac{C_2 \cos(\pi n/2)}{C_0 (2\pi)^{1-m}}.$$

Here, f is the linear frequency in hertz ($\omega = 2\pi f$). Using the least squares method, we found the parameters A , B , $1 - n$, and $1 - m$, approximating the experimental data by formula (3). As follows from Fig. 1, our theory adequately approximates the experimental data. Figure 4 shows the exponents $1 - n$ and $1 - m$ vs. humidity. If the humidity exceeds 33%, the exponents $1 - n$ and $1 - m$ are almost humidity-independent within the experimental error and the accuracy of the least squares method. Since in this range, $1 - n \approx 1$, the term A/f^{1-n} in (3) corresponds to Debye relaxation. Consequently, this contribution should be assigned to water filling the free volume. On the other hand, the exponent $1 - m$ noticeably differs from 1 ($1 - m \approx 0.4$). This implies that the contribution B/f^{1-m} in (3) is due to bound water, i.e., to water covering the porous surface of the clay. The fundamental difference between bound water and water in the free volume has been noted in [2]. This difference stems from interaction between water and the porous surface and from the fractal properties of the pores [3, 4]. At a humidity of 12%, the exponent $1 - n$ sharply drops, while $1 - m$ changes insignificantly (Fig. 4). We therefore can assume that water is almost entirely in the bound state at this value of the humidity.

Thus, from our experimental data, it follows that the frequency dependence of the permittivity of a saturating fluid is significantly affected by the fractal geometry of the pores.

REFERENCES

1. M. Shahidi, J. B. Hasted, and A. K. Jonsher, *Nature* **258**, 595 (1975).
2. F. D. Ovcharenko, *Hydrophily of Clays and Clay Minerals* (Akad. Nauk USSR, Kiev, 1961).
3. A. Mehaute, R. R. Nigmatullin, and L. Nivanen, *Fleches du temps et geometrie fractale* (Hermes, Paris, 1998).
4. Yu. D. Feldman, R. R. Nigmatullin, and E. Polygalov, *Phys. Rev. E* **58**, 7561 (1998).

Translated by V. Isaakyan

**BRIEF
COMMUNICATIONS**

On the Theory of Hydrodynamic Phenomena in Quasi-One-Dimensional Systems

S. O. Gladkov

*Moscow State Institute of Radio Engineering, Electronics, and Automation (Technical University),
pr. Vernadskogo 78, Moscow, 117454 Russia*

Received December 18, 2000; in final form, March 7, 2001

Abstract—It is shown that the Navier–Stokes equation for the description of the hydrodynamic properties of structures in d_F -dimensional media can be obtained with the principle of least action. The analysis of the solution to the equations obtained is exemplified by a quasi-one-dimensional flow. © 2001 MAIK “Nauka/Interperiodica”.

Earlier, a way to investigate the dissipative properties of highly inhomogeneous structures was suggested (see, for example, [1, 2]). It was assumed that there exists the hypothetical analogy between such structures and structures that are homogeneous in some geometric space where the principal matrix has a dimensionality $d_F = d + \varepsilon$. Here, $d = 1, 2, \text{ or } 3$ and the value of ε depends on the properties of the filler. It was also assumed that the properties of the matrix can be equivalently described by formally replacing the space of dimensionality d_F by some metric Riemann space with the metric tensor components other than unity.

This work, where we logically elaborate upon the previous investigations, is devoted to the hydrodynamic description of the motion of liquids and gases in spaces with a fractional dimensionality. Our approach is in many ways similar to that adopted in [1] to characterize thermal conduction in quasi-one-dimensional structures.

To deduce the equations of hydrodynamics in an arbitrary space of dimensionality d_F , we proceed as follows. As is known [3], the Navier–Stokes equation has the form

$$\rho d\mathbf{v}/dt = -\nabla P + \eta \Delta \mathbf{v} + \zeta \nabla (\nabla \cdot \mathbf{v}). \quad (1)$$

Here, ρ is the density of the liquid; P is the pressure; \mathbf{v} is the velocity of a given element of the liquid ($d\mathbf{v}/dt = \partial \mathbf{v}/\partial t + (\mathbf{v} \cdot \nabla) \mathbf{v}$); and η and ζ are the first and second viscosities, respectively.

In any space, the scalar quantities ρ , η , ζ , and P are invariant unlike the vectors. In a d_F -dimensional space, the pressure gradient can be expressed as $\mathbf{A}P$ using fractional differentiation, where the operator of fractional differentiation \mathbf{A} is defined as [4]

$$A_\alpha f = \int_{-\infty}^{\infty} (ik_\alpha)^{1+\varepsilon} f_k e^{ikx} d^3 k / (2\pi)^3. \quad (2)$$

Here, the subscript α means $x, y, \text{ or } z$. It follows from definition (2) that the operator \mathbf{A} is a normalizable linear operator. The sum of two last terms in Eq. (1) can be obtained by taking the variational derivative of the functional

$$L\{\mathbf{v}\} = -(1/2) \times \int [K_1 (\partial v_i / \partial x_k)^2 + K_2 (\partial v_i / \partial x_i)^2] d^3 x, \quad (3)$$

where K_1 and K_2 are constants.

Representing the Navier–Stokes equation as

$$\rho d\mathbf{v}/dt = -\nabla P + \gamma \delta L / \delta \mathbf{v}, \quad (4)$$

where γ is a phenomenological dissipation coefficient, we come to Eq. (1) with

$$\eta = \gamma K_1, \quad \zeta = \gamma K_2. \quad (5)$$

It was assumed hypothetically [4] that a space of a fractional dimensionality d_F is equivalent (isomorphic) to some continuous Riemann space with a certain metrics. Based on this assumption, we substitute

$$L\{\mathbf{v}\} = -(1/2) \int [K_1 v_k^i v_i^k + K_2 (v_i^i)^2] g^{1/2} d^3 x \quad (6)$$

for expression (3). Here, g is the determinant of the metric tensor for this metric space (see below) and v_k^i is the k th covariant derivative of the velocity \mathbf{v} .

Since $v_k^i = \partial v^i / \partial x^k + \Gamma_{ks}^i v^s$, where Γ_{ks}^i is the Christoffel symbol, Eq. (6) yields

$$L\{\mathbf{v}\} = -(1/2) \int \{ K_1 (\partial v^i / \partial x^k + \Gamma_{ks}^i v^s) (\partial v^k / \partial x^i + \Gamma_{is}^k v^s) + g^{-1} K_2 [(\partial / \partial x^i) (g^{1/2} v^i)]^2 \} g^{1/2} d^3 x. \quad (7)$$

Varying Eq. (7) with respect to v^i , we obtain

$$\delta L = \int \{ K_1(\partial/\partial x^k)[g^{1/2}(\partial v^k/\partial x^i + \Gamma_{is}^k v^s)]\delta v^i - K_1\Gamma_{is}^k(\partial v^s/\partial x^k + \Gamma_{kp}^s v^p)\delta v^i g^{1/2} + K_2(\partial/\partial x^i)[(1/g^{1/2})(\partial/\partial x^k)(g^{1/2} v^k)]\delta v^i g^{1/2} \} d^3x.$$

Hence, Eq. (4) is reduced to

$$\begin{aligned} & \rho g_{is}[\partial v^s/\partial t + v^k(\partial v^s/\partial x^k + \Gamma_{kp}^s v^p)] \\ & - A_i P + \eta(\partial/\partial x^k)[g^{1/2}(\partial v^k/\partial x^i + \Gamma_{is}^k v^s)] \\ & - \eta\Gamma_{ik}^s(\partial v^k/\partial x^s + \Gamma_{sp}^k v^p)g^{1/2} \\ & + \zeta g^{1/2}(\partial/\partial x^k)[(1/g^{1/2})(\partial/\partial x^k)(g^{1/2} g_{is} v^s)]. \end{aligned} \tag{8}$$

To obtain the complete system of equations, the continuity equation

$$\partial\rho/\partial t + g^{-1/2}(\partial/\partial x^i)(g^{1/2} v^i) = 0 \tag{9}$$

should be added to Eq. (8).

It was hypothesized that the physical properties of complex structures can be described in terms of fractional-dimension spaces specified by a certain metrics. To exemplify the use of Eqs. (8) and (9), we shall consider a quasi-one-dimensional flow. Let the metric be

$$dl^2 = g(x)dx^2 + dy^2 + dz^2, \tag{10}$$

where

$$g(x) = g_{xx} = (x/L_0)^{\varepsilon\nu} \tag{11}$$

is the only other-than-zero component of the metric tensor. Here, $\varepsilon < 0$ is the true fractional dimensionality of the space, L_0 is some characteristic size, and ν is a factor responsible for the fulfillment of specific conditions. The meaning of ν is as follows. Let a quasi-one-dimensional structure be specified in a continuous metric space and let its prototype in our real world be, say, a fir-tree. Then, the lengths of the structure and its prototype calculated with metrics (11) must be equal as scalar values. From this condition, ν is determined. Such a requirement, in principle, meets the reality. Further, we shall write simply ε instead of the product $\varepsilon\nu$. For metrics (11), only one Christoffel symbol is not zero:

$$\Gamma_{xx}^x = -(1/2g)\partial g/\partial x = -\varepsilon/2x. \tag{12}$$

Then, Eq. (8) for the component v^x takes the form

$$\begin{aligned} & g(x)\rho[\partial v^x/\partial t + v^x(\partial v^x/\partial x)(1 - \varepsilon/2x)] \\ & = -A_x P + \eta g^{1/2}[\partial^2 v^x/\partial x^2 + (\varepsilon/2x)\partial v^x/\partial x \\ & + \varepsilon(1 - \varepsilon)v^x/2x^2] \\ & + \xi g^{1/2}(\partial/\partial x)[g^{-1/2}(\partial(v^x g^{1/2})/\partial x)]. \end{aligned} \tag{13}$$

In the case of an incompressible liquid, by virtue of Eq. (9) we obtain

$$(\partial/\partial x^i)(g^{1/2} v^i) = 0; \tag{14}$$

hence, the term proportional to ξ in Eq. (13) disappears.

Equation (14) yields

$$v^x = Cx^{-\varepsilon/2}, \tag{15}$$

where $C = \text{const}$.

Let us solve Eq. (13) simultaneously for small velocities, when the term $v^x(\partial v^x/\partial x)(1 - \varepsilon/2x)$ can be neglected (small Reynolds numbers); for the steady-state case; and for small ε . Putting $v^x = v$, we have

$$\begin{aligned} & \partial^2 v^x/\partial x^2 + (\varepsilon/2x)\partial v^x/\partial x + \varepsilon(1 - \varepsilon)v^x/2x^2 \\ & = (1/\eta L_0^\varepsilon) \int_{-\infty}^{\infty} (ik)^{1+\varepsilon} P_k e^{ikx} dk/2\pi, \end{aligned} \tag{16}$$

where, in the right-hand side, we used definition (2), which determines the action of the linear operator of fractional differentiation on a scalar function and P_k is the Fourier transform of pressure $P(x)$.

To solve the equation obtained, let us assume that the pressure varies according to the law

$$P(x) = P_0 \cos ax, \tag{17}$$

where a is a constant having the dimension of reciprocal length.

This means that the Fourier transform is given by

$$P_k = \frac{2P_0\gamma(a^2 + k^2 + \gamma^2)}{(a^2 + k^2 + \gamma^2)^2 - 4a^2k^2}, \tag{18}$$

where γ is some formally introduced quantity providing the convergence of the integral.

For such P_k , we calculate the integral in the right-hand side of Eq. (16). In fact,

$$\begin{aligned} J & = (\gamma P_0/\pi\eta L^\varepsilon) \int_{-\infty}^{\infty} (ik)^{(1+\varepsilon)} \\ & \times \frac{e^{ikx}(a^2 + k^2 + \gamma^2)dk}{(a^2 + k^2 + \gamma^2)^2 - 4a^2k^2}. \end{aligned} \tag{19}$$

Integral (19) is easy to calculate using the theory of residues. Eventually (after passing to the limit $\gamma \rightarrow 0$), we obtain

$$J = -(P_0 a^{1+\varepsilon}/\eta L^\varepsilon) \sin(ax - \pi\varepsilon/2) e^{i\pi(1+\varepsilon)}. \tag{20}$$

The solution of homogeneous equation (16) for small ε leads to

$$v(x) = Ax^{1-\varepsilon} + Bx^{-\varepsilon/2}. \tag{21}$$

Now if we fix x , setting it equal to some value x_0 , and draw a plane perpendicular to that of Fig. 1a, we

apparently obtain the picture shown in Fig. 1b. A one-dimensional structure of “holes” from which the liquid flows out is clearly seen. This structure, roughly speaking, is a Cantor set where holes of different diameters can be considered as points, the largest hole corresponding to the basic trunk.

Next, solving Eq. (16) by the method of variation of constants, we find in view of Eq. (20)

$$v(x) = Ax^{1-\varepsilon} + Bx^{-\varepsilon/2} + D[x^{1-\varepsilon}/(1-\varepsilon/2)] \times \int x^\varepsilon \sin(ax - \pi\varepsilon/2) dx - D[x^{-\varepsilon/2}/(1+\varepsilon/2)] \int x^{1+\varepsilon/2} \sin(ax - \pi\varepsilon/2) dx, \tag{22}$$

where $D = (P_0 a / \eta)(a / L_0)^\varepsilon e^{i\pi(1+\varepsilon)}$.

In the absence of pressure ($P_0 = 0$), solutions (22) and (15) must coincide. This means that we should set $A = 0$ and $B = C$. Finally,

$$v(x) = Cx^{-\varepsilon/2} + D[x^{1-\varepsilon}/(1-\varepsilon/2)] \times \int x^\varepsilon \sin(ax - \pi\varepsilon/2) dx - D[x^{-\varepsilon/2}/(1+\varepsilon/2)] \int x^{1+\varepsilon/2} \sin(ax - \pi\varepsilon/2) dx. \tag{23}$$

If $\varepsilon = 0$ in Eq. (23), we obtain the natural result

$$v_{\varepsilon=0}(x) = C - (P_0/a\eta) \sin ax. \tag{24}$$

The limiting case, i.e., Eq. (24), confirms the validity of the mathematics and Eq. (23), which implies that in quasi-one-dimensional structures, the flow rate is not constant in the absence of external factors (in particular, pressure). The explanation of this is trivial: the interaction between different parts of d_F -dimensional structures always exists. Note that our theory (in particular, quasi-homogeneous) is also of applied value, since the hydrodynamic flow of liquids along randomly branched pipelines with a gradually decreasing radius has not been amenable to theoretical treatment.

The only condition required for the mathematical characterization of the physical properties of such systems is their similarity in some range of the space scale. The suggested heuristic physical description of such complex branched systems may be useful in many applications.

In fact, results obtained by this method can be used in those fields of science and technology that are seem-

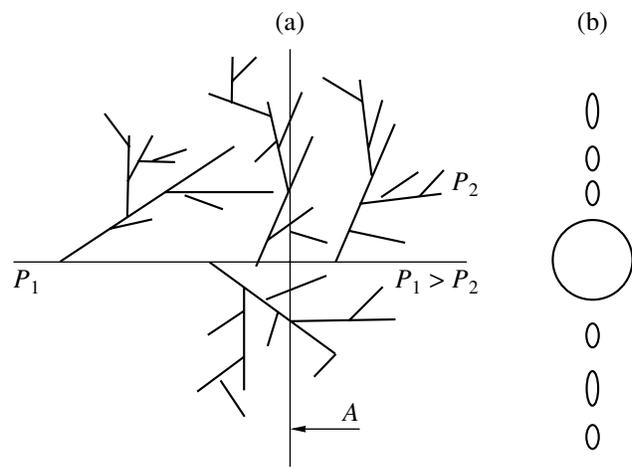


Fig. 1. (a) Schematic representation of a structure with $\varepsilon > 0$ and (b) its view in the A section at $x = \text{const}$.

ingly far from hydrodynamics, for example, in fiberoptic technology, which deals with branched optical fibers, and in the theory of crystallization.

CONCLUSION

We suggested the way to describe hydrodynamic phenomena in spaces of arbitrary dimensionality and applied d_F -dimensional hydrodynamic equation (8) to the quasi-one-dimensional case as an example.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant no. 99-01-00011).

REFERENCES

1. S. O. Gladkov, Zh. Tekh. Fiz. **67** (7), 8 (1997) [Tech. Phys. **42**, 724 (1997)].
2. S. O. Gladkov, Perspekt. Mater., No. 1, 5 (2000).
3. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics, Vol. 6: Fluid Mechanics* (Nauka, Moscow, 1988; Pergamon, New York, 1987).
4. S. O. Gladkov, *Physics of Composites: Thermodynamic and Dissipative Properties* (Nauka, Moscow, 1999).

Translated by M. Astrov

BRIEF
COMMUNICATIONS

An Electrostatic Spectrograph in the Form of a Truncated Cylinder

L. P. Ovsyannikova and T. Ya. Fishkova

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

e-mail: L. Ovsyannikova@pop.iofe.rssi.ru

Received February 15, 2001

Abstract—An electrostatic spectrograph based on an energy analyzer in the form of a truncated cylinder is designed. The parameters of the device can vary in a wide range. The position of the focal line is found for the beam energies differing by one order of magnitude. The linear and specific energy dispersions are determined. The performance of the spectrograph suggested is compared with that of a conventional spectrograph with two planar electrodes. © 2001 MAIK “Nauka/Interperiodica”.

The energy spectrum of charged particle beams is usually determined with a plane capacitor. Capacitor-based spectrographs are widely used for finding electron spectra, particularly, in studying the ion and atom fluxes emitted by a hot plasma (see, e.g., [1]). When the capacitor operates under the mirror conditions with the 45° entrance angle for the central beam path, first-order angular focusing is observed in the dispersion plane provided that the source and the detectors are placed on the lower capacitor plate [2]. The disadvantages of the plane capacitor are (1) the absence of focusing in the plane normal to the dispersion plane, (2) the need for placing the detectors immediately on the capacitor plane, and (3) the fact that the plane capacitor, being opened on the sides, has the self-field of scattering and is subjected to external fields.

A mirror analyzer in the form of a truncated cylinder [3] is free of these disadvantages. It is the aim of this work to study its operation in the spectrograph mode (Fig. 1) and to compare its parameters with those of the plane capacitor. Our analysis relies on formulas derived in [3]. In the 2D approximation, the expression for the potential distribution between the flat and cylindrical electrodes was more conveniently written as

$$\phi(x, y) = V/b \{ \pi/2 - \arctan[(a^2 - x^2 - y^2)/(2ax)] \}, \quad (1)$$
$$a = \sqrt{p(2-p)}, \quad b = \pi/2 - \arctan[(1-p)/a],$$

where V is the potential difference between the electrodes and p is the interelectrode spacing along the normal (the segment height).

Hereafter, p , x , y , and other geometric parameters are expressed in terms of the radius of curvature R of the cylindrical electrode. Note that in [3], the expression for the proportionality coefficient b , which is responsible for the value of the potential difference between the electrodes, was in error. Yet, this error does

not influence the form of the formulas derived in the work cited. The values of the parameters calculated for a segment of height p are merely valid for an analyzer with the $(2-p)$ -high segment.

In this work, we study the performance of a truncated-cylinder energy analyzer operating as a spectrograph when its dimensions are varied in a wide range ($0.2 \leq p/R \leq 1.8$) and the energy spread of the beam is as high as one order of magnitude. As for the plane mirror configuration, the entrance angle of the central beam path is set equal to $\theta_0 = 45^\circ$ and the source is assumed to be placed on the planar grounded electrode. Conditions for first-order focusing in the dispersion plane are considered. For these conditions, Fig. 2 demonstrates the d_i vs. eV/ϵ and h_i vs. eV/ϵ curves, where d_i is the entrance–exit length of the central paths for the beams with various energies; h_i is the shortest spacing between the detectors and the planar electrode; e and ϵ are the particle charge and energy, respectively; eV/ϵ is the electrostatic field force; and i is the channel no. Both d_i and h_i grow with segment height and rapidly diminish with decreasing energy. Starting from $eV/\epsilon \approx 3$, the focal point of the beam falls on the planar electrode of the spectrograph.

Under the angular focusing conditions, the linear energy dispersion is $D = L/(2\cos^2\theta_0)$, where L is the analyzer base (source–detector distance), for any electrostatic analyzer [4]. For the entrance angle $\theta_0 = 45^\circ$, $D_i = L_i = d_i + h_i$.

To estimate the specific dispersion, which is the measure of the resolving power and is defined as the ratio of the linear dispersion to the highest aberration term, we calculated the second-order aberration coefficients in the dispersion plane. For a truncated cylinder,

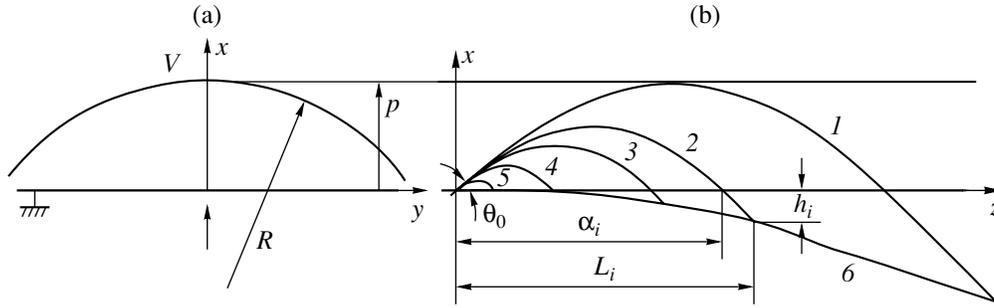


Fig. 1. Electrostatic spectrograph of charged particles in the form of a truncated cylinder. (a) Cross section and (b) the dispersion plane showing the central beam trajectories in the truncated cylinder with a segment height $p = 0.4R$. The analyzer force eV/ϵ is (1) 0.5, (2) 0.75, (3) 1.0, (4) 2.0, and (5) 5.0. Curve 6 is the focal line, at which the detectors are placed.

they are given by

$$C_2/R = 3/2[(h_0 + h)(\cot\theta_0 - \tan\theta_0)/\sin^2\theta_0 - d\tan^2\theta_0] - ab\epsilon/(eV)\cos^3\theta_0/\sin\theta_0\{1 - 2[b\epsilon/(eV)\sin\theta_0]^2\} + 8/a^3[b\epsilon/(eV)]^{5/2}\sin^2\theta_0\cos^3\theta_0 \quad (2)$$

$$\times \int_0^{x_m} x(2a^2 + 3x^2)\sqrt{\arctan(x_m/a) - \arctan(x/a)}dx.$$

The formulas for L (the entrance–exit length for the central trajectory of the beam), $h_0 + h$ (the sum of the distances from the sources and the detector to the electrode plane), and x_m (the coordinate of the point where the central trajectory turns) are given in [3]. The maximal energy of the charged particles at which the turn point of the central trajectory lies on the cylindrical electrode is independent of the segment height:

$$\epsilon_{\max} = eV/\sin^2\theta_0. \quad (3)$$

At the entrance angle $\theta_0 = 45^\circ$, $\epsilon_{\max} = 2$ eV, which corresponds to an analyzer field force $eV/\epsilon = 0.5$.

Figure 3 shows the aberration coefficients calculated by formula (2) for different geometries of a truncated cylinder, as well as the second-order aberration coefficients for a plane mirror when the interelectrode spacing equals the cylinder radius. The latter are given by the simple analytic form [2] $C_2 = 2dp/R$ (p is the interelectrode spacing). It is seen that if the segment height $p \geq R$, the aberration coefficients vanish at high energies in the analyzer force range $eV/\epsilon = 0.7-1.1$. However, their energy dependence in this range is very sharp: when the analyzer operates in the spectrograph mode at energies $\epsilon = (0.5-2.0)$ eV, the aberration coefficient in the dispersion plane reaches the value $C_2 = (100 - 10)R$. Therefore, if the segment height exceeds the cylinder radius, the analyzer can operate as a spectrograph only in the narrow energy range $\epsilon = (0.2-0.5)$ eV. In this case, the maximum distance x_m between the trajectories and the planar electrode is much smaller than p ; that is, the analyzer aperture is used inefficiently.

If $p \leq 0.3R$, the distance between the focal line and the planar electrode is small, $0.01 \leq h_i/R \leq 0.1$, and the energies detected differ only by a factor of 2 or 3.

Based on the aforesaid, we selected an analyzer with the distance between the cylindrical and planar electrodes $p = 0.4R$. For this geometry, the beam central trajectories for a tenfold difference in the beam energies

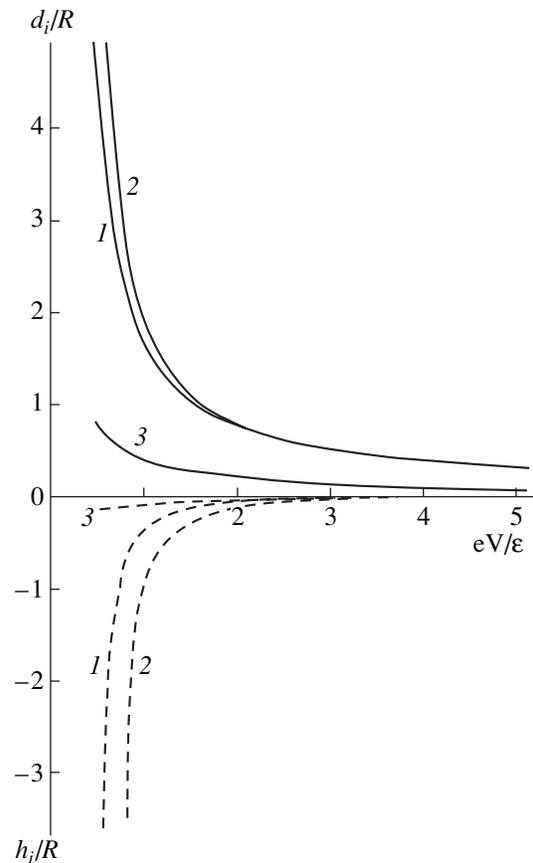


Fig. 2. d_i/R vs. eV/ϵ (continuous curves) and h_i/R vs. eV/ϵ (dashed curves), where d_i is the entrance–exit length of the central paths for the beams and h_i is the spacing between the detectors and the planar electrode. The interelectrode spacing $p =$ (1) R , (2) $1.8R$, and (3) $0.2R$.

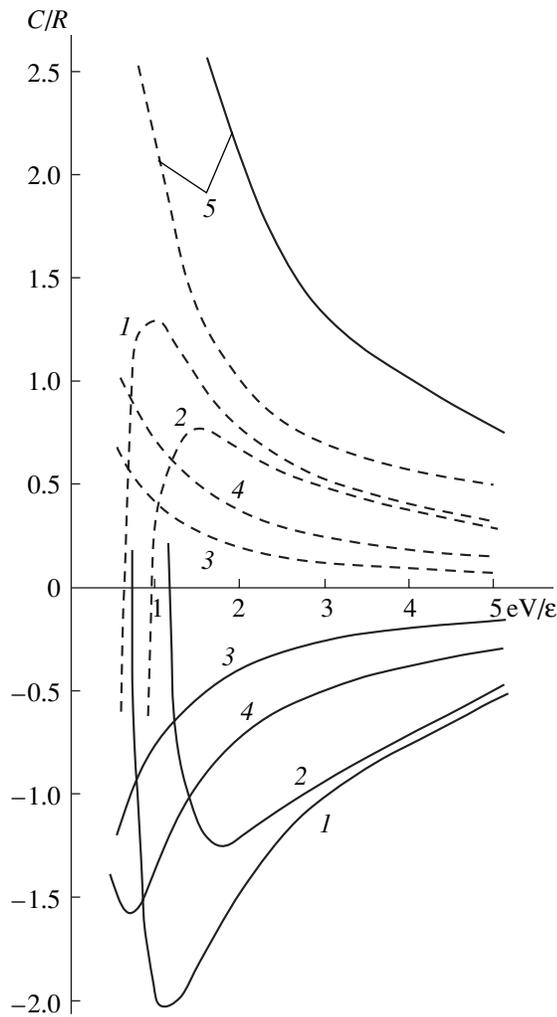


Fig. 3. Aberration coefficients in the dispersion plane (continuous curves) and coefficients that define the beam spread in the vertical plane (dashed curves) vs. analyzer force for $p = (1) R$, (2) $1.8R$, (3) $0.2R$, and (4) $0.4R$. Curves 5 refer to the plane capacitor.

are depicted in Fig. 1. At high energies, $\varepsilon = (0.4-2.0)$ eV, the focal line is outside the field, unlike the case of a plane capacitor.

The linear energy dispersions for both (truncated-cylinder and plane-mirror) analyzers are nearly the same. The specific dispersions $\delta = D/C_2$ in the low-energy channels are also almost coincident. However, in the high-energy channels, $\varepsilon = (2-1)$ eV, the specific

dispersion of the truncated cylinder is 1.3–3.0 times as large as that of the plane mirror, for which $\delta = 0.5$ for each of the channels.

It is known that a plane mirror analyzer does not provide focusing in the vertical plane (normal to the dispersion plane). Therefore, the image width in the focus is $\Delta y = C_\gamma \tan \gamma$, where $C_\gamma = d$ and γ is the angular half-aperture of the beam in the vertical plane.

We calculated the coefficients C_γ for the truncated cylinder analyzer by numerically solving the second-order differential equations derived in [3]. The numerical results are depicted in Fig. 3 by the dashed curves. At high energies, these coefficients are much smaller than for the plane mirror analyzer. Hence, the beam is focused in the vertical plane. At low energies, this effect is insignificant.

It should be noted that focusing in the vertical plane and the lower aberration in the dispersion plane increase the luminosity of the spectrograph. This is of special importance for the identification of hot-plasma high-energy components, since the particle flux density rapidly drops with energy. An increased luminosity improves the identification of the high-energy spectra of the particles detected.

Thus, a double-electrode electrostatic spectrograph in the form of a truncated cylinder was designed and its optimal geometry was justified. It offers a lower aberration level in the dispersion plane compared with a plane capacitor and the focusing effect in the vertical plane. In addition, our analyzer having the planar and cylindrical electrodes is closed on the sides, which eliminates the self-field of scattering and excludes the penetration of external fields.

REFERENCES

1. V. S. Koïdan, *Prib. Tekh. Éksp.*, No. 3, 63 (1971); A. B. Izvozchikov *et al.*, *Zh. Tekh. Fiz.* **62** (2), 157 (1992) [*Sov. Phys. Tech. Phys.* **37**, 201 (1992)].
2. V. P. Afanas'ev and S. Ya. Yavor, *Electrostatic Energy Analysers for Charged-Particle Beams* (Nauka, Moscow, 1978), pp. 54–58.
3. T. Ya. Fishkova and I. A. Korishch, *Zh. Tekh. Fiz.* **56**, 367 (1986) [*Sov. Phys. Tech. Phys.* **31**, 221 (1986)].
4. T. Ya. Fishkova, *Zh. Tekh. Fiz.* **57**, 1358 (1987) [*Sov. Phys. Tech. Phys.* **32**, 804 (1987)].

Translated by V. Isaakyan

BRIEF
COMMUNICATIONS

Self-Diffusion at Grain Boundaries with a Disordered Atomic Structure

V. N. Perevezentsev

*Blagonravov Mechanical Engineering Research Institute (Nizhni Novgorod Branch), Russian Academy of Sciences,
ul. Belinskogo 85, Nizhni Novgorod, 603024 Russia*

e-mail: pevn@uic.nnov.ru

Received March 13, 2001

Abstract—A change in the free energy of a grain boundary is analyzed in the case when lattice vacancies come to the boundary and are then delocalized in its disordered atomic structure. It is shown that the free energy of the boundary is minimized at some excess atomic volume $\Delta v_b = \Delta v_b^*$, whose value depends on the energy of vacancy formation in the crystal lattice and the boundary energy. The formation of a metastable localized grain-boundary vacancy as a result of thermal fluctuations of the density in a group of $n_0 = \Omega_v/\Delta v_b$ atoms (Ω_v is the vacancy volume), followed by the jump of an adjacent atom into this vacancy, is taken as an elementary event of grain-boundary diffusion. Expressions for the activation energy of diffusion and the diffusion coefficient are derived for equilibrium ($\Delta v_b = \Delta v_b^*$) and nonequilibrium ($\Delta v_b > \Delta v_b^*$) grain boundaries. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

In spite of a variety of papers on experimental and theoretical investigation into diffusion along disordered grain boundaries (which are ordinary grain boundaries with the misorientations far from specific angles [1]), commonly accepted concepts of mechanisms behind self-diffusion along them are still lacking [2, 3]. In recent years, interest in this problem has quickened because of anomalously high coefficients of diffusion along grain boundaries that were discovered in submicrocrystalline materials obtained by high-rate plastic deformation [4]. It has been shown [5, 6] that coefficients of self- and heterodiffusion in submicrocrystalline materials may be one order of magnitude (or even more) higher than in the coarse-grain state. These anomalies of grain-boundary diffusion are assumed to be associated with the nonequilibrium structure of grain boundaries [4].

In this work, we suggest a unified approach to describing diffusion along equilibrium and nonequilibrium grain boundaries with the disordered structure.

MODEL

We will consider a grain boundary (GB) as a thin (of thickness $\delta_0 \sim (2-3)a_b$, where a_b is the interatomic spacing) layer of an amorphous material sandwiched in misoriented grains. It is obvious from general considerations that the GB specific free energy depends on the volume occupied by GB atoms. If the number of the atoms remains unchanged, there is some GB volume at which the energy is minimal. With this in mind and assuming,

in the first approximation, that the GB energy and volume are uniformly distributed among GB atoms, the free energy F_b of a GB atom can be expanded in powers of $(v_b - v_b^0)$, where v_b is the GB atomic volume and v_b^0 is the atomic volume for the minimal-energy boundary. Up to second-order terms, we can write

$$F_b = F_b^0 + \frac{1}{2}k'(v_b - v_b^0)^2, \quad (1)$$

where the coefficient k' is related to the compression coefficient $1/K_b = -(1/v_b)(\partial v_b/\partial p)$ by the relationship $k' = K_b/v_b^0$. The second term in the right-hand side of (1) describes the elastic energy of the GB atom. Note that the linear term in (1) is absent.

Consider how the GB energy changes when the boundary captures vacancies from the volume. The jump of a lattice vacancy to the boundary means that one GB atom is removed and a GB vacancy arises. If the atomic structure does not rearrange by relaxation, such an unrelaxed vacancy has an energy $F_{vf}^b = F_{vf}^{cr} - F_b^0$, where F_{vf}^{cr} is the free energy of vacancy formation in the crystal lattice. However, the computer simulation of usual boundaries [7–9] shows that vacancies in disordered atomic structures are unstable and are rapidly delocalized, that is, “smeared” as a result of relaxation displacements of atoms within a region measuring several tens or hundreds of interatomic spacings. In terms of our model, vacancy delocalization increases the excess volume $(v_b - v_b^0)$ and, according to formula

(1), raises the elastic energy of GB atoms. If the concentration C_b of lattice vacancies at a GB is sufficiently high, elastically strained regions due to vacancy delocalization overlap and one can assume that eventually the elastic energy will be uniformly distributed and depend only on the deviation Δv_b of the atomic volume v_b from its initial value v_b^0 ; that is, $\Delta v_b = (v_b - v_b^0) = C_b v_b v_b^0$. The associated change ΔE_e^b in the GB elastic energy per unit area can be written as

$$\Delta E_e^b = \frac{K_b}{2v_b^0} \left(\frac{\delta_0}{v_b^0} - C_b \delta_0 \right) (C_b v_b v_b^0)^2. \quad (2)$$

Without considering the contribution from the configurational vibrational entropy, the total change ΔF of the GB free energy as a result of the relaxation of the GB atomic structure and vacancy delocalization (at $C_b v_b^0 \ll 1$) is given by

$$\Delta F = -C_b \delta_0 (F_{vf}^{cr} - F_b^0) + \frac{1}{2} \delta_0 K_b C_b^2 v_b^2. \quad (3)$$

From the condition $\partial(\Delta F)/\partial(C_b v_b) = 0$, one can find the "equilibrium" excess atomic volume $\Delta v_b^* = C_b^* v_b v_b^0$, at which the GB energy is minimal:

$$\Delta v_b^* = \frac{(F_{vf}^{cr} - F_v^0) v_b^0}{K_b v_b}. \quad (4)$$

At $\Delta v_b < \Delta v_b^*$, the absorption of a lattice vacancy by a GB is energetically favorable (the vacancy flux is directed toward the boundary), while at $\Delta v_b > \Delta v_b^*$, the energetically favorable process is the injection of vacancies into the grain volume (the associated mechanism is discussed below). A GB at which the excess atomic volume equals Δv_b^* will be referred to as an equilibrium boundary.

Now consider the mechanism of GB self-diffusion. Let GB atoms have some excess volume Δv_b . Thermal fluctuations may locally increase the density of the GB material. It is easy to check that if the fluctuations cover a group of n_0 atoms, where $n_0 = \Omega_v / \Delta v_b$, and decrease the atomic volume v_b to v_b^0 , an unrelaxed vacancy of volume Ω_v will appear (hereafter, we assume that $\Omega_v = v_b^0$). In this case, the elastic energy of atoms covered by the fluctuations is released and the free energy increases by the energy of the unrelaxed vacancy ($F_{vf}^{cr} - F_b^0$). Hence, within this mechanism, the activa-

tion energy of GB vacancy generation can be expressed as

$$\begin{aligned} F_{vf}^b &= F_{vf}^{cr} - F_b^0 - n_0 \frac{K_b}{2v_b^0} (v_b - v_b^0)^2 \\ &= (F_{vf}^{cr} - F_b^0) - \frac{K_b \Delta v_b}{2}. \end{aligned} \quad (5)$$

The expression for the free energy of activation of GB diffusion is obtained if the activation energy F_{vm}^b of GB atom jump into the resulting vacancy is added to (5):

$$F_d^b = F_{vf}^b + F_{vm}^b = (F_{vf}^{cr} - F_b^0 + F_{vm}^b) - \frac{K_b \Delta v_b}{2}. \quad (6)$$

COEFFICIENT OF DIFFUSION ALONG EQUILIBRIUM AND NONEQUILIBRIUM GRAIN BOUNDARIES

Let us calculate the coefficient of diffusion for an equilibrium grain boundary. Since the excess atomic volume Δv_b is defined by expression (4) in this case, we obtain from (4) and (6) (at $v_b/v_b^0 \cong 1$)

$$F_d^b = \frac{F_{vf}^{cr} - F_b^0 + 2F_{vm}^b}{2}. \quad (7)$$

Consequently, the expression for the coefficient of diffusion along an equilibrium GB is given by

$$D_b = D_{b_0} \exp\left(-\frac{H_{vf}^{cr} - H_b^0 + 2H_{mv}^b}{2kT}\right), \quad (8)$$

where

$$D_{b_0} = \frac{1}{z_b} a_b^2 v_b \exp\left(\frac{S_{vf}^{cr} - S_b^0 + 2S_{mv}^b}{k}\right). \quad (9)$$

Here, z_b is the coordination number; v_b is the frequency of atomic oscillations at the boundary; k is the Boltzmann constant; H_{vf}^{cr} , H_b^0 , and H_{mv}^b are the enthalpies; S_{vf}^{cr} , S_b^0 , and S_{mv}^b are the entropies in the associated expressions for the free energy ($F = H - TS$). As follows from (8), the energy of GB diffusion activation $Q_b = (H_{vf}^{cr} - H_b^0 + 2H_{mv}^b)/2$ is related to the enthalpy of lattice vacancy formation H_{vf}^{cr} . Expressing H_{vf}^{cr} through the energy of volume diffusion activation Q_v ($H_{vf}^{cr} = Q_v - H_{vm}^{cr}$, where H_{vm}^{cr} is the enthalpy of activation of vacancy migration in the crystal lattice), one can relate Q_b and Q_v as

$$Q_b = \frac{Q_v + (2H_{vm}^b - H_{vm}^{cr}) - H_b^0}{2}. \quad (10)$$

Let us estimate the parameters entering this expression. We will use their typical values: $Q_v = 18kT_m$ (T_m is the melting point) [2], $H_b^0 = \gamma_b^0 v_b^0 / \delta_0 \cong 1.3kT_m$ (γ_b^0 is the GB specific surface enthalpy), $\gamma_b^0 a_b^2 \cong 4kT_m$ [10], $\delta_0 \cong 3a_b$, and $H_{vm}^{cr} \cong (0.45-0.50)Q_v = (8.1-9.0)kT_m$ [11]. Since literature data for H_{vm}^b are lacking, we will assume that its contribution to the energy of GB diffusion activation (by analogy with volume diffusion) is 0.45–0.50; that is, $H_{vm}^b \cong (0.45-0.5)Q_b \cong (4.0-4.5)kT_m$. Substituting these values to (10) yields $Q_b = (9.1-10.5)kT_m$. Thus, expression (10) is in good agreement with the well-known empirical relationship $Q_b \approx Q_v/2$.

The preexponential D_{b_0} is still harder to estimate, since reliable literature data on the parameters entering (9) are absent. A rough estimation can be made if we put $S_{vf}^{cr} \cong 2k$ [12], $S_b^0 \cong (\tilde{S}_b^0 a_b^2) / \delta_0 \cong 0.66k$ ($\tilde{S}_b^0 a_b^2 = 2k$ is the entropy per unit surface area of the boundary [10]), $S_{mv}^b \cong 0.5S_{mv}^{cr} \cong 1k$, $Z_b \cong 6$, $a_b \cong 3 \times 10^{-8}$ cm, $v_b \sim 10^{13}$ 1/s. Then, $D_{b_0} = 1.7 \times 10^{-2}$ cm²/s, which coincides with experimental value of D_{b_0} by order of magnitude [2].

The activation energy of diffusion along a nonequilibrium boundary (a boundary at which atoms have an excess volume $\Delta v_b > \Delta v_b^*$) is obtained by substituting $\Delta v_b = \Delta v_b^* + \Delta \tilde{v}_b$ (where $\Delta \tilde{v}_b$ is the nonequilibrium excess volume of an atom) into (6). The expression for the coefficient of diffusion D_b' along a nonequilibrium GB is then given by

$$D_b' = D_b \exp\left(\frac{K_b \Delta \tilde{v}_b}{2kT}\right), \quad (11)$$

where D_b is the coefficient of diffusion along an equilibrium boundary.

Thus, at $\Delta \tilde{v}_b > 0$, the diffusion coefficient for a nonequilibrium boundary exceeds that for an equilibrium boundary. In particular, with $T = T_m/2$, $K_b = 2G(1 - \mu)/3(1 - 2\mu)$ (where G is the GB shear modulus and μ

is Poisson's ratio), $G a_b^3 \cong 50kT_m$, and $\mu \cong 1/3$, the coefficient D_b' exceeds D_b by one order of magnitude at $\Delta \tilde{v}_b \cong 2 \times 10^{-2} v_b$. Note that the value of $\Delta \tilde{v}_b$ depends on the concentration of nonequilibrium defects (vacancies, dislocations, etc.) that arrive at the boundary during either deformation or post-deformation recovery processes. The effect of nonequilibrium defects on GB diffusion will be considered in subsequent publications.

ACKNOWLEDGMENTS

The author thanks L.N. Perevezentseva for the assistance in preparing the article.

REFERENCES

1. A. N. Orlov, V. N. Perevezentsev, and V. V. Rybin, *Grain Boundaries in Metals* (Metallurgiya, Moscow, 1980).
2. I. Kaur and W. Gust, *Fundamentals of Grain and Interphase Boundary Diffusion* (Stuttgart, 1989; Mashinostroenie, Moscow, 1991).
3. V. Naundorf, M. P. Macht, A. S. Bakai, *et al.*, *J. Non-Cryst. Solids* **250-252**, 679 (1999).
4. R. Z. Valiev and I. V. Aleksandrov, *Nanostructural Materials Received by Severe Plastic Deformation* (Logos, Moscow, 2000).
5. R. Z. Valiev, T. M. Razumovskii, and V. I. Sergeev, *Phys. Status Solidi A* **139**, 321 (1993).
6. R. Wurschum, A. Kubler, S. Gruss, *et al.*, *Ann. Chim. (Paris)* **21**, 471 (1996).
7. H. Gleiter, *Mater. Sci. Eng.* **52**, 91 (1982).
8. R. W. Ballufi, T. Kwak, P. D. Bristowe, *et al.*, *Scr. Metall.* **15**, 951 (1981).
9. A. V. Vekman, *Cand. Sci. Dissertation*, Barnaul, 2000.
10. J. W. Prowan and O. A. Bamiro, *Acta Metall.* **25**, 309 (1977).
11. B. S. Bokshtein, S. Z. Bokshtein, and A. A. Zhukhovitskiĭ, *Thermodynamics and Kinetics of Diffusion in Solids* (Metallurgiya, Moscow, 1974).
12. *Physical Metallurgy*, Ed. by R.W. Cahn and P. Haasen (North-Holland, Amsterdam, 1983; Metallurgiya, Moscow, 1987), Vol. 3.

Translated by V. Isaakyan

BRIEF
COMMUNICATIONS

Emission from a DC Glow Discharge in a He/H₂O Mixture

A. K. Shuaibov, A. I. Dashchenko, and I. V. Shevera

Uzhgorod National University, Podgornaya ul. 46, Uzhgorod, 88000 Ukraine

e-mail: ishev@univ.uzhgorod.ua

Received March 27, 2001

Abstract—The electrical and optical characteristics of a longitudinal dc glow discharge in a cylindrical discharge tube in mixtures of helium with saturated water vapor at room temperature are investigated. In the UV range, a broad band with a maximum at $\lambda_{\max} = 309.6$ nm and $\Delta\lambda = 9$ nm prevails. The H α 656.3-nm, H β 486.1-nm, and HeI lines in the range 440–670 nm are the main diagnostic spectral lines. The helium partial pressure and the glow discharge current are optimized to achieve the maximum intensities of the 309.6-nm band and HeI and H α spectral lines. The results obtained are of interest for the development of an ecologically safe radiation source based on the products of the decomposition of water molecules and clusters in plasma. © 2001 MAIK “Nauka/Interperiodica”.

INTRODUCTION

Low-temperature plasmas of rare gas mixtures with a small admixture of water vapor are selective radiation sources based on the products of the decomposition of water molecules (mainly OH radicals) [1, 2]. The creation of an ecologically safe lamp based on the OH(A–X; 0–0) 306.4-nm transition pumped by a glow or an RF discharge was reported in [3]. In He(Ar)/H₂O mixtures, the intensity of the OH(A–X) band increases with increasing water-vapor partial pressure from 1 to 130 Pa. At larger water-vapor concentrations, when (H₂O)_m, (OH)_m, and (OH)_m · (H₂O)_n cluster molecules begin to play an important role, the discharge emission has not been studied, although the cluster plasma is of interest for lamp development [4].

In this paper, we present the results of studies of the emission from a dc glow discharge in the water vapor–helium plasmas at $P(\text{H}_2\text{O}) = 2.0$ – 2.5 kPa.

EXPERIMENTAL

A glow discharge was ignited in a quartz tube with an inner diameter of 7 mm and an anode–cathode spacing of 50 mm. The design of the discharge tube is described in [5]. The power deposited in the discharge was 40–60 W and the discharge current varied within the range 2–50 mA. Saturated water vapor at room temperature ($T = 17$ – 21°C) was produced by evaporating distilled water from a reservoir placed in the lower part of a 10-l buffer chamber. The discharge tube with open ends was placed in the upper part of the buffer chamber. The pressure of the saturated water vapor varied within the range 2.0–2.5 kPa. The intensity of the emission band was determined as the area below the spectrum curve corrected for the relative spectral sensitivity of the recording system. The residual air pressure in the buffer chamber did not exceed 10–15 Pa. The spectral

resolution was 0.2 nm. The plasma emission intensity was measured with an accuracy of 7–10%.

ELECTRICAL AND OPTICAL CHARACTERISTICS

A dc glow discharge in a He/H₂O mixture at $P(\text{He}) = 1$ – 16 kPa was quite homogeneous over the discharge tube length. As the helium partial pressure increased, the diameter of the discharge plasma decreased from 4–5 to 2 mm. The normal mode of the discharge was obtained at currents $I_{\text{ch}} \geq 30$ mA, whereas at low currents, the discharge operated in the subnormal mode (Fig. 1). In the normal mode, the ignition voltage and quasi-steady discharge voltage increased with helium partial pressure. In our experiments, the U_{ch} values were one order of magnitude

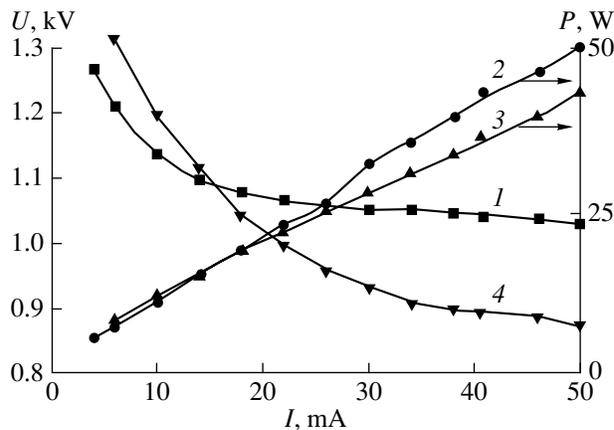
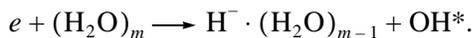


Fig. 1. (1, 4) Current–voltage characteristics and (2, 3) the power deposited in a longitudinal glow discharge vs. the discharge current for $P(\text{He})/P(\text{H}_2\text{O}) = (1, 2)$ 16/2.5- and (3, 4) 1.0/2.5-kPa mixtures.

higher and I_{ch} values were one order of magnitude lower than those in [3], although the discharge tube geometry and the power $P_{\text{ch}} = I_{\text{ch}}U_{\text{ch}}$ were approximately the same.

The UV and visible emission spectra from the plasma are shown in Fig. 2. The main constituents of the UV spectrum of He/H₂O plasma are a broad band with a maximum at $\lambda_{\text{max}} = 309.6$ nm and two weaker bands with $\lambda_{\text{max}} = 286$ and 262 nm. The most intense band has a fine structure with peaks at 307.0, 309.6 (the main one), 312.4, and 315.2 nm. The closest to the recorded peaks are the edges of the OH(A-X; 0-0) 306.4-nm (R-branch), OH(A-X; 0-0) 308.9-nm (Q-branch), OH(A-X; 1-1) 312.2-nm, and OH(A-X; 2-2) 318.5-nm bands [6-8]. In [9], the broad emission band in the wavelength range 300-310 nm with peaks at 307.0, 308.3, and 309.1 (the main one) nm was also recorded in an RF discharge in water vapor at $P(\text{H}_2\text{O}) = 100$ Pa and was assigned to the spontaneous decay of (H₂O)₂ cluster molecules, although no justification for such an assignment was presented. The edges of the 286- and 262-nm bands correlate to the well-known transitions of OH radicals: OH(A-X; 1-0) 281.1-282.9 nm, OH(B-A) 278 nm, and OH(A-X; 2-0) 260.9-262.2 nm. The experiments with a glow discharge in He/H₂O = (0.5-16)/2.5-kPa mixtures show that, in the VUV spectral region, the emission is concentrated in a broad band with $\lambda_{\text{max}} = 185$ nm. A decrease in the water-vapor partial pressure to 0.1-0.2 kPa shifts the maximum of the VUV band emission from 185 to 180 nm, which almost corresponds to the edge of the OH(C-A) 179-nm emission band [3]. It is seen that the recorded peaks of the UV emission correlate with edges of the most intense bands of the OH radical and are slightly shifted to the long-wavelength side of the spectrum. Thus, the characteristic emission of a plasma of saturated water vapor can be attributed to the emission of (OH)₂^{*} hydroxyl dimers. At elevated water vapor pressures, a rapid hydration of OH dimers to the cluster molecules like (OH)₂^{*} · (H₂O)_m (where $m \geq 1$) occurs [10]. In a saturated water vapor plasma, the excited OH radicals are produced via the reaction



When OH and OH* radicals collide, the creation and subsequent hydration of (OH)₂^{*} dimers can occur. In the 300- to 400-nm wavelength range, the bands belonging to the 2⁺ molecular nitrogen system (nitrogen is present in the gas mixture as a minor admixture) were observed. In the visible region (440-680 nm), the H_α, H_β, and H_γ spectral lines, as well as the HeI 667.8-nm, 587.6-nm, 501.0-nm, and 447.1-nm lines, were observed. These spectral lines can be used to measure the n_e and T_e values by the emission spectroscopy. The intensity of the HeI 667.8-nm line is much higher than that of the HeI 587.6-nm line,

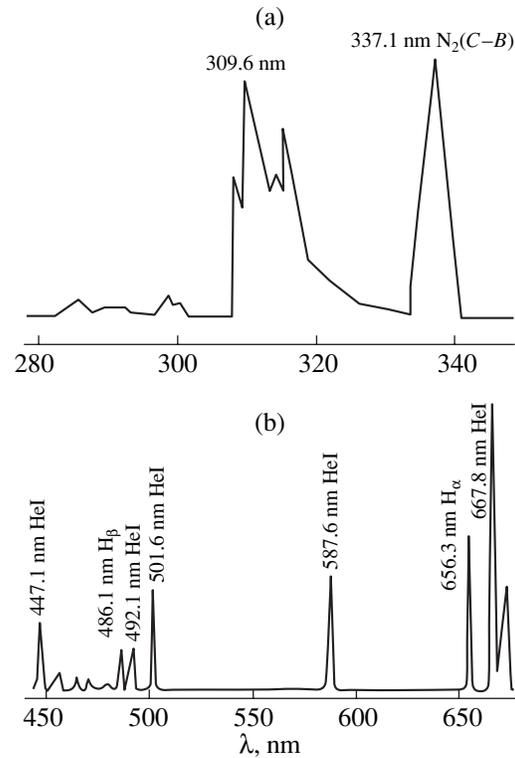


Fig. 2. (a) UV and (b) visible emission spectra from the plasma of a He/H₂O mixture.

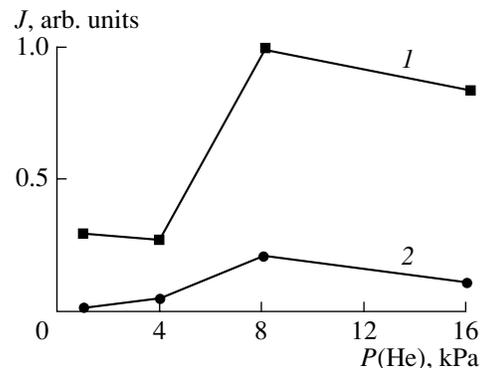


Fig. 3. Intensities of the emission bands with $\lambda_{\text{max}} =$ (1) 309.6 and (2) 337.1 nm for a glow discharge in a He/H₂O mixture vs. the helium partial pressure for $P(\text{H}_2\text{O}) = 2.5$ kPa and $I_{\text{ch}} = 50$ mA.

although, according to the table data [11], the ratio of their intensities should be 0.12. The upper levels corresponding to these lines have the same energy ($\epsilon_{\text{up}} = 23.07$ eV), whereas the energies of the lower levels ϵ_0 are quite different (21.22 and 20.96 eV, respectively) [11].

The increase in the intensity of the HeI red line may be related to the depopulation of the lower level in collisions with water molecules and the products of water dissociation in the discharge.

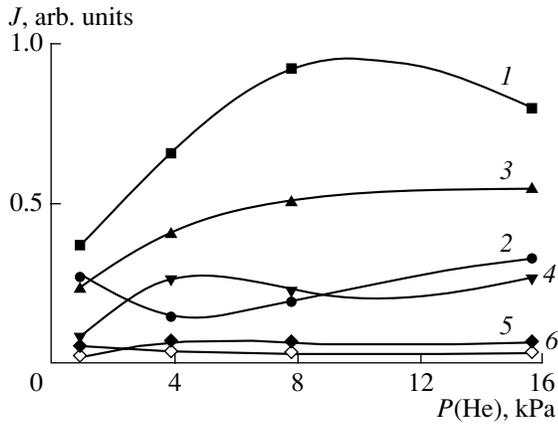


Fig. 4. Intensities of the helium and hydrogen line emission from a glow discharge in a He/H₂O mixture vs. the helium partial pressure for $P(\text{H}_2\text{O}) = 2.5$ kPa and $I_{\text{ch}} = 50$ mA: (1) HeI 667.8-nm, (2) H α 656.3-nm, (3) HeI 587.6-nm, (4) HeI 501-nm, (5) H β 486.1-nm, and (6) HeI 491-nm lines.

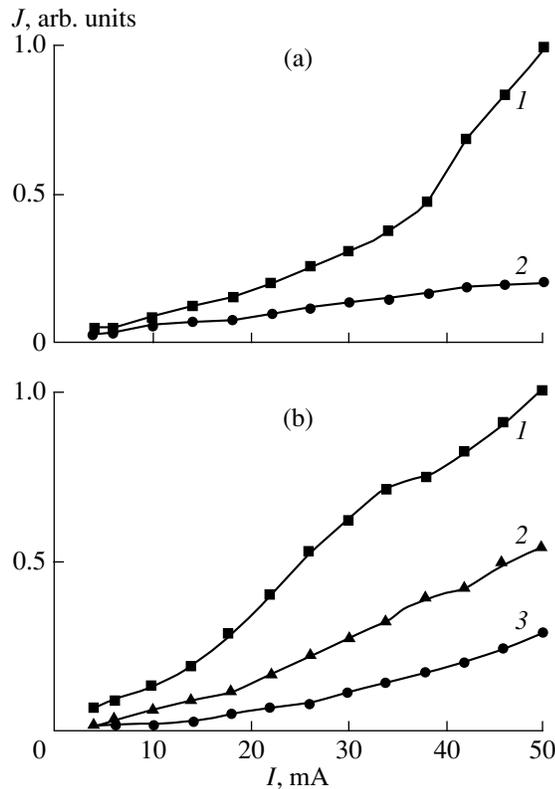


Fig. 5. Intensities of (a) the bands with $\lambda_{\text{max}} = (1)$ 309.6 and (2) 337.1 nm and (b) (1) the HeI 667.8-nm, (2) H α 656.3-nm, and (3) H β 486.1-nm spectral lines for a glow discharge in the He/H₂O = 8.0/2.5-kPa mixture vs. the discharge current.

In the range $P(\text{He}) = 1.0\text{--}4.0$ kPa, when the discharge is quite uniform along the discharge tube radius, the intensity of the emission band with $\lambda_{\text{max}} = 309.6$ nm slightly decreases with increasing helium partial pressure (Fig. 3). The intensity of the admixture molecular

emission in the N₂(C–B; 0–0) 337.1-nm band, which has a broad maximum at $P(\text{He}) = 8$ kPa, is one order of magnitude lower than that of the 309.6-nm characteristic band. As the helium partial pressure increases, the plasma column contracts and the intensity of the 309.6-nm band significantly increases and reaches its maximum at $P(\text{He}) = 8$ kPa.

Figure 4 shows the typical dependences of the intensities of helium and hydrogen linear emission from a glow discharge in a He/H₂O mixture on the helium partial pressure. For the H α line, the emission intensity as a function of the helium partial pressure, $J = f(P[\text{He}])$, has a minimum at $P(\text{He}) = 4.0$ kPa; the intensity of this spectral line is approximately one order of magnitude higher than that of the H β line. The intensity of the H β line steadily decreases with helium partial pressure. For the HeI 667.8-nm, 501.0-nm, and 491.1-nm lines, the dependence $J = f(P[\text{He}])$ has a maximum at $P(\text{He}) = 4\text{--}8$ kPa, whereas for the HeI 587.6-nm line, the intensity steadily increases up to $P(\text{He}) = 16$ kPa. The dependences of the intensities of the band with $\lambda_{\text{max}} = 309.6$ nm and the HeI 667.8-nm spectral line on the helium partial pressure are qualitatively the same in the helium pressure range 4–16 kPa. This indirectly indicates that, at elevated pressures of a He/H₂O mixture, the excited helium atoms play a significant role in the production of the excited OH radicals and (OH)*-based clusters.

Figure 5 presents the typical dependences of the intensities of the characteristic bands of water vapor plasma and admixture molecules (N₂) and the intensities of HI and HeI spectral lines on the discharge current. For the band with $\lambda_{\text{max}} = 309.6$ nm, the dependence is almost linear at low currents, whereas at $I_{\text{ch}} \geq 15\text{--}20$ mA, the intensity increases more rapidly. For the 337.1-nm band, the dependence is linear over the entire range of the discharge currents. The most intense spectral lines are the HeI 667-nm and HI 656.3-nm lines; the dependences of their intensities on I_{ch} are almost linear. The emission intensities of other HeI lines ($\lambda = 587.6$, 506.1, and 491 nm) are 10–50 times lower than that of the HeI red line; the dependences of their intensities on the discharge current are similar.

CONCLUSIONS

The results of investigations of a longitudinal dc glow discharge in He/H₂O mixtures at $P(\text{H}_2\text{O}) = 2.5$ kPa can be summarized as follows.

(i) In the wavelength range 200–400 nm, the emission band with $\lambda_{\text{max}} = 309.6$ nm and width $\Delta\lambda = 9$ nm is the most intense.

(ii) The H α 656.3-nm and H β 486.1-nm lines, as well as the HeI 667.8-nm, 587.6-nm, 501.0-nm, and 491.1-nm lines, can be used to determine the plasma density n_e and temperature T_e by the emission spectroscopy.

(iii) The maximum intensity of the 309.6-nm band is attained at a helium partial pressure of 8 kPa; in the discharge current range 20–50 mA, the intensity increases nonlinearly with the current without any tendency to saturate.

(iv) A simple UV radiation source based on the characteristic band with $\lambda_{\max} = 309.6$ nm and operating in the regime of the slow replacement of the $P(\text{He})/P(\text{H}_2\text{O}) = 8.0/2.5$ kPa gas mixture at $I_{\text{ch}} \geq 50$ mA can be created.

REFERENCES

1. A. A. Mavlyutov, A. I. Mis'kevich, and B. S. Salamakha, *Opt. Spektrosk.* **76**, 946 (1994) [*Opt. Spectrosc.* **76**, 847 (1994)].
2. A. K. Shuaibov and I. V. Shevera, *Zh. Prikl. Spektrosk.* **67**, 242 (1997).
3. A. Ya. Vul', S. V. Kidalov, V. M. Milenin, *et al.*, *Pis'ma Zh. Tekh. Fiz.* **25** (1), 10 (1999) [*Tech. Phys. Lett.* **25**, 4 (1999)]; *Pis'ma Zh. Tekh. Fiz.* **25** (8), 62 (1999) [*Tech. Phys. Lett.* **25**, 321 (1999)].
4. B. M. Smirnov, *Usp. Fiz. Nauk* **170**, 495 (2000).
5. A. K. Shuaibov, A. I. Dashchenko, and I. V. Shevera, *Zh. Prikl. Spektrosk.* **68**, 275 (2001).
6. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).
7. R. W. B. Pears and A. G. Gaydon, *The Identification of Molecular Spectra* (Chapman and Hall, London, 1963).
8. L. Wallace, *Astrophys. J., Suppl. Ser.* **7**, 245 (1962).
9. M. Soskida, in *Abstracts of the 32nd Conference on European Group for Atomic Spectroscopy, Vilnius, 2000*, p. 314.
10. A. A. Vostrikov and D. Yu. Dubov, *Zh. Tekh. Fiz.* **57**, 760 (1987) [*Sov. Phys. Tech. Phys.* **32**, 459 (1987)].
11. A. R. Striganov and G. A. Odintsova, *Tables of Spectral Lines of Atoms and Ions: Handbook* (Énergoatomizdat, Moscow, 1982).

Translated by N. Ustinovskii