Technical Physics, Vol. 47, No. 11, 2002, pp. 1345–1349. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 1–4.

THEORETICAL AND MATHEMATICAL PHYSICS

Heterogeneous Combustion of a Pair of Interacting Particles with an Arbitrarily Shaped Surface

E. R. Shchukin and G. Yu. Yalamov

Moscow Pedagogical University, ul. Radio 10a, Moscow, 107005 Russia

e-mail: in.info@g23.relcom.ru

Received January 30, 2002; in final form, May 6, 2002

Abstract—The heterogeneous combustion of two immobile interacting particles with arbitrarily shaped surfaces is theoretically described under conditions of rapid chemical reaction, where the concentration of oxidizer molecules on the surface of the particles can be considered to be zero. The problem is solved for an arbitrary temperature dependence of molecular transport coefficients when molecular transport in the vicinity of the particles takes place by diffusion. Analytical formulas found allow for the direct characterization of the combustion of individual spheroidal particles and interacting spherical particles in several particular cases. At the same volumes, heavily elongated and oblate particles are shown to burn significantly faster. The approach of the particles significantly influences the time of burning the finer particle if its size is far less than that of the coarser one. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Heterogeneous combustion occurs when molecules of a gaseous medium enter into a chemical reaction with molecules of a solid particle. The process can take place, e.g., in the furnaces of thermal power plants [1, 2] or in zones where aerodisperse systems are clarified by laser radiation [3–6]. Aerosol particles may have an arbitrary shape of their surface and be arbitrarily spaced. Therefore, problems related to the effect of the particle surface shape and approach on the rate and time of combustion are of great importance. Golovin *et al.* [7] have recently found the rate of the high-temperature combustion of an individual spheroidal carbon particle for the case of temperature-independent transport coefficients. However, there is a need to elaborate a combustion theory for arbitrarily shaped particles when there is a significant temperature drop near them, where the molecular transport coefficients depend on temperature. It is also of interest to study the effect of the interaction between a pair of particles with arbitrary surface shapes on their heterogeneous combustion.

First, we consider the free combustion of two particles such that heat generation inside the particles is absent and the effect of radiative heat exchange on the combustion can be neglected. The combustion of the two arbitrarily shaped interacting particles occurs in an *N*-component gas mixture. The chemical reaction follows the scheme

$$\nu_1 A_1 + \nu^{(p)} A^{(p)} = \sum_{j=2}^{N-1} \nu_j A_j,$$

where the coefficients ν_t are equal to the stoichiometric coefficients of the reaction in magnitude ($\nu_1 > 0$, $\nu^{(p)} > 0$, $\nu_j < 0$, $j = 2, \ldots, N-1$), and the superscript p denotes the stoichiometric coefficient of the particle material.

When the atoms of a carbon particle are oxidized by oxygen molecules to carbon monoxide ($O_2 + 2C = 2CO$), the symbols A_1, A_2, and $A^{(p)}$ denote O_2, CO, and C molecules, respectively. The relative oxidizer concentration on the particle surface is $C_1|_S = C_1^{(s)} = 0$. Temperature drops in the vicinity of the particle can be significant. The sizes of the particles and the distance between them are assumed to be sufficiently small; hence, the effect of gravitational convection on heat-and-mass transfer cannot be neglected. The combustion of the particle proceeds in the quasi-steady-state mode. In the vicinity of the particles,

$$\sum_{j=1}^{N-1} C_j \ll 1,$$

where $C_j = n_j/n$, $n = \sum_{j=1}^{N} n_j$, and n_j are the concentrations of the gaseous component molecules.

Under these conditions, the basic mechanism of molecular transfer is molecular diffusion. It should be noted that, at $C_1^{(s)} = 0$, the approach of the particles and the shape of their surfaces affect the rate of combustion most strongly.

With the above assumptions, the distributions of C_j, gas temperature T_e, and particle temperature $T_i^{(f)}$ in the particle–gaseous environment system are described by the set of equations

$$\nabla \cdot \mathbf{q}_j = 0, \quad \nabla \cdot \mathbf{q}_T^{(e)} = 0, \quad \nabla \cdot \mathbf{q}_T^{(f)} = 0, \qquad (1)$$

with the boundary conditions

$$C_1|_{S_f} = 0, \quad T_e|_{S_f} = T_i^{(f)}|_{S_f}, \quad \mathbf{q}_{\perp j}|_{S_f} = (\nu_j/\nu_1)\mathbf{q}_{\perp 1}|_{S_f},$$
$$[L_P m_P(\nu^{(p)}/\nu_1)\mathbf{q}_{\perp 1}^{(e)} + \mathbf{q}_{\perp T}^{(e)}]|_{S_f} = \mathbf{q}_{\perp T}^{(f)}|_{S_f}, \quad (2)$$

$$C_j|_\infty = C_{j\infty}, \quad T_e|_\infty = T_{e\infty}. \quad (3)$$

Here, $\mathbf{q}_j$, $\mathbf{q}_T^{(e)}$, and $\mathbf{q}_T^{(f)}$ are the densities of the molecular and thermal flows, respectively:

$$\mathbf{q}_j = -D_{jN} n \nabla C_j, \quad \mathbf{q}_T^{(e)} = -\chi_e \nabla T_e,$$
$$\mathbf{q}_T^{(f)} = -\chi_i^{(f)} \nabla T_i^{(f)}; \quad (4)$$

χ_e and $\chi_i^{(f)}$ are the gas and particle thermal conductivities, respectively; the index $\perp$ denotes the projections of the vectors onto the normals to the particle surfaces; L_P is the thermal effect of the chemical reaction; m_P is the mass of particle molecules; and the superscript $f = 1, 2$ denotes the number of the particle.

The coefficients D_{jN} of binary diffusion and the thermal conductivity χ_e depend on T_e. The third of boundary conditions (2) and the first of boundary conditions (3) take into account the stoichiometric relation between the molecular flow densities and the continuity of the thermal flow, respectively.

The solution to boundary-value problem (1)–(3) is

$$C_j = \{C_{j\infty} - C_{1\infty}[\nu_j \psi_j(T_e)/\nu_1 \psi_1(T_{i0})]\},$$
$$\psi_T(T_e) = \psi_T(T_{i0}) U(x_k), \quad T_i^{(1)} = T_i^{(2)} = T_{i0} = \text{const}, \quad (5)$$

where x_k are the coordinates of points in the space;

$$\psi_j(T_e) = \int_{T_{e\infty}}^{T_e} (\chi_e/nD_{jN}) dT_e, \quad j = 1, \ldots, N-1;$$
$$\psi_T(T_e) = \int_{T_{e\infty}}^{T_e} \chi_e dT_e. \quad (6)$$

The values of T_{i0} are found by solving the algebraic equation

$$\psi_1(T_{i0}) = L_P m_P(\nu^{(p)}/\nu_1) C_{1\infty}. \quad (7)$$

Thus, at $C_1^{(s)} = 0$, the quasi-steady-state combustion of the particles proceeds at the same temperature irrespective of their shapes, sizes, and spacing.

The values of $U(x_k)$ are determined by solving the boundary problem

$$\Delta U = 0, \quad U|_{S_f} = 1, \quad U|_\infty = 0, \quad (8)$$

where Δ is the Laplacian operator.

Having determined T_{i0}, we can calculate the flow densities $\mathbf{q}_j$ and $\mathbf{q}_T$ and the total molecular, $Q_j^{(f)}$, and thermal, $Q_T^{(f)}$, flows removed from the surfaces of the particles by the formulas

$$\mathbf{q}_j = \frac{\nu_j}{\nu_1}\mathbf{q}_1, \quad \mathbf{q}_1 = -C_{1\infty}F_1(T_{i0})\mathbf{q}_U,$$
$$\mathbf{q}_T = \Psi_T(T_{i0})\mathbf{q}_U, \quad \mathbf{q}_U = -\nabla U, \quad (9)$$

$$Q_j^{(f)} = \frac{\nu_j}{\nu_1} Q_1^{(f)}, \quad Q_1^{(f)} = -C_{1\infty}F_1(T_{i0})Q_U^{(f)},$$
$$Q_T^{(f)} = \Psi_T(T_{i0})Q_U^{(f)}, \quad Q_U^{(f)} = -\oint_{S_f} \nabla U d\mathbf{s}, \quad (10)$$

where $d\mathbf{s}_f$ is the differential vector element of the surface and $F_1(T_{i0}) = \Psi_T(T_{i0})/\Psi_1(T_{i0})$.

Knowing the temperature T_{i0} and having solved boundary problem (8), we can determine the rate of change of the particle mass (with a known surface shape) by the formula

$$\frac{dM_P^{(f)}}{dt} = m_P \frac{\nu^{(p)}}{\nu_1} Q_1^{(f)}, \quad (11)$$

where $M_P^{(f)} = \rho_P V_P^{(f)}$; ρ_P and $V_P^{(f)}$ are the density and volume of the particle, respectively; and t is time.

When calculating changes in the particle shape and volume, we should take into account the expression for the normal projections of the flow densities for particle molecules removed from the surface: $\mathbf{q}_P^{(f)} = -(\nu^{(p)}/\nu_1)\mathbf{q}_{\perp 1}|_{S_f}$.

Boundary-value problem (8) can be solved numerically in the general case and analytically in some particular cases (for example, in the case of individual spherical, spheroidal, and ellipsoidal particles or two interacting particles). For a spherical particle of radius R,

$$U = R/r, \quad \nabla U = (-R/r^2)\mathbf{n}_r, \quad Q_U^{(1)} = 4\pi R, \quad (12)$$

where r is the radial coordinate and $\mathbf{n}_r$ is the unit vector.

For elongated and oblate individual spheroidal particles,

$$U(\xi) = \ln\frac{\cosh\xi + 1}{\cosh\xi - 1} \Big/ \ln\frac{\cosh\xi_0 + 1}{\cosh\xi_0 - 1},$$
$$\nabla_\xi U = -\frac{2\mathbf{n}_\xi}{c\sinh\xi(\sinh^2\xi + \sin^2\eta)^{1/2}\ln((\cosh\xi_0 + 1)/(\cosh\xi_0 - 1))},$$

$$a > b;$$

$$U(\xi) = \frac{\operatorname{arccot}(\sinh\xi)}{\operatorname{arccot}(\sinh\xi_0)},$$

$$\nabla_\xi U = -\frac{2\mathbf{n}_\xi}{c\cosh\xi(\cosh^2\xi - \sin^2\eta)^{1/2}\operatorname{arccot}(\sinh\xi_0)}, \quad (13)$$

$$a < b;$$

$$Q_{U_0}^{(1)} = 4\pi c I(\xi_0);$$

$$I(\xi_0) = \frac{2}{\ln((\cosh\xi_0 + 1)/(\cosh\xi_0 - 1))}, \quad a > b;$$

$$I(\xi_0) = \frac{1}{\operatorname{arccot}(\sinh\xi_0)}, \quad a < b.$$

Here, a is the length of the semiaxis coinciding with the axis of symmetry of the spheroid; b is the length of the other semiaxis; ξ, η, and φ are the spheroidal coordinates [8]; $c = \sqrt{|a^2 - b^2|}$; and $a = c\cosh\xi_0$ and $b = c\sinh\xi_0$ ($a > b$) or $a = c\sinh\xi_0$ and $b = c\cosh\xi_0$ ($a < b$).

In the case of two spherical particles of radii R_1 and R_2,

$$U(\xi, \eta) = U_1(\xi, \eta) + U_2(\xi, \eta),$$

$$U_{k=1,2}(\xi, \eta) = \sqrt{2(\cosh\xi - \cos\eta)}$$

$$\times \sum_{n=0}^{\infty} d_n^{-1}\left[\left(\sinh(-1)^k(\xi_{2^{2-k}} - \xi)\left(n + \frac{1}{2}\right)\right)\right.$$

$$\left.\times \exp\left(-\left(n + \frac{1}{2}\right)|\xi_k|\right)\right]P_n(\cos\eta),$$

$$Q_U^{(f)} = 4\pi R_f\Theta_f, \quad \Theta_f = \Theta_{f,1} - \Theta_{f,2}, \quad (14)$$

$$\Theta_{f,1} = \Psi_f\sinh|\xi_f|,$$

$$\Theta_{f,2} = \psi_3\sinh|\xi_f|, \quad d_n = \sinh(n + 1/2)(\xi_1 + |\xi_2|),$$

$$\Psi_f = \sum_{h=0}^{\infty}\frac{1}{d}\exp\left[(|\xi_{2^{2-f}}| - |\xi_f|)\left(n + \frac{1}{2}\right)\right],$$

$$\Psi_3 = \sum_{h=0}^{\infty}\frac{1}{d_n}\exp\left[-\left(n + \frac{1}{2}\right)(\xi_1 + |\xi_2|)\right],$$

where ξ, η, and φ are the bispherical coordinates [8], h is the center distance of the particles, $\cosh\xi_f = (h^2 + R_f^2 - R_{2^{2-f}}^2)/2hR_f$ ($\xi_1 > 0$, $\xi_2 < 0$), and $P_n = (\cos\eta)$ are the Legendre polynomials.

Estimations made with the formulas derived showed that heavily elongated and oblate particles burn significantly faster at the same volumes. This is clearly exhibited by the curve (Fig. 1) representing the dependence

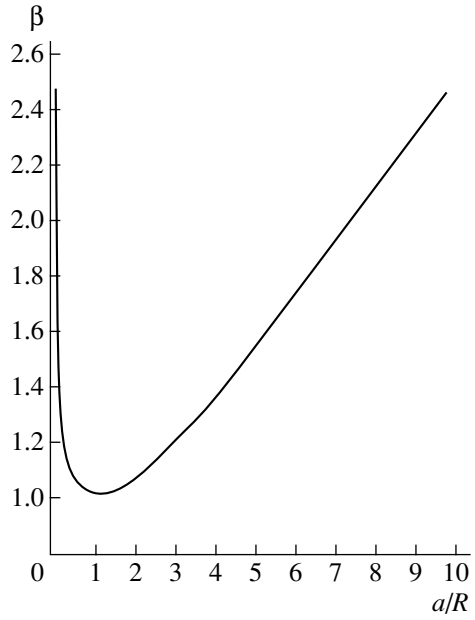


Fig. 1. Dependence of the coefficient β on a/R for spheroidal particles.

of the coefficient $\beta = Q_U^{(1)}/Q_{U_0}^{(1)}$ for individual spheroidal particles on the ratio a/R, where a is the length of the semiaxis of the spheroidal particle and R is the radius of a spherical particle of the same volume. For calculations, the values of $Q_U^{(1)}$ and $Q_{U_0}^{(1)}$ were found by Eqs. (13) and (12), respectively. As the particles approach each other, their rate of combustion decreases (Fig. 2). This is because the flux of oxidizer molecules toward the particle surface decreases when the particles draw near one another. Figure 2 plots the ratio $Q_1^{(1)}/Q_{1\infty}^{(1)}$ against the parameter $S = (h - R_1 - R_2)/R_1$ ($Q_{1\infty}^{(1)} = Q_1^{(1)}$ at $h = \infty$). Curve *1* illustrates the decay of $Q_1^{(1)}$ for particles with $R_1 = R_2$; curve *2*, with $R_2 = 4R_1$.

Let us consider the effect of internal sources and radiative heat exchange on the rate of combustion of interacting particles by the example of two spherical particles with a thermal conductivity χ_i that is much greater than the thermal conductivity χ_e of the gaseous environment. Under these conditions, the particles burn at fixed surface temperatures $T_i^{(1)}$ and $T_i^{(2)}$ independent of the angles η and φ. We describe the combustion of the particles in the quasi-steady-state approximation assuming the concentrations of components involved in the chemical reaction to be small. In this case, just as in the case of free combustion, the distributions of the temperature T_e and of the relative concentrations C_j in the vicinity of the particles are described by the boundary problem

$$\nabla\cdot\mathbf{q}_j = 0, \quad \nabla\cdot\mathbf{q}_T^{(e)} = 0, \quad (15)$$

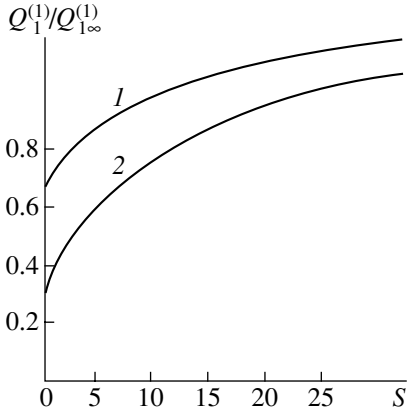


Fig. 2. Dependence of the flow ratio $Q_1^{(1)}/Q_{1\infty}^{(1)}$ on the variable S.

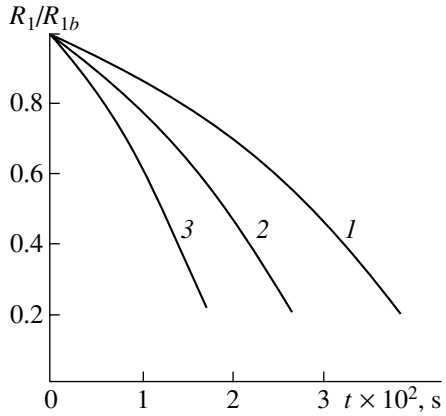


Fig. 3. Time dependence of the dimensionless radius of the finer particle.

$$C_1|_{S_f} = 0, \quad T_e|_{S_f} = T_i^{(f)}, \quad \mathbf{q}_{\perp j}|_{S_f} = (\nu_j/\nu_1)\mathbf{q}_{\perp 1}|_{S_f},$$
$$C_j|_\infty = C_{j\infty}, \quad T_e|_\infty = T_{e\infty}. \tag{16}$$

In the case of small drops in T_e (when the T_e dependence of χ_e can be neglected), the solution to set (15) and (16) is

$$C_j = C_{j\infty} - (\nu_j D_{1N}/\nu_1 D_{1N})C_{1\infty}U(\xi, \eta),$$

$$T_e = T_{e\infty} + (T_i^{(1)} - T_{e\infty})U_1(\xi, \eta) + (T_i^{(2)} - T_{e\infty})U_2(\xi, \eta).$$

At the same time,

$$Q_j^{(f)} = -(\nu_j/\nu_1)nD_{1N}C_{1\infty}Q_U^{(f)},$$

$$Q_T^{(1)} = 4\pi R_1\chi_e[(T_i^{(1)} - T_{e\infty})\Theta_{1,1} - (T_i^{(2)} - T_{e\infty})\Theta_{1,2}],$$

$$Q_T^{(2)} = 4\pi R_2\chi_e[(T_i^{(2)} - T_{e\infty})\Theta_{2,1} - (T_i^{(1)} - T_{e\infty})\Theta_{2,2}].$$

The dependences of $T_i^{(f)}$ and R_f on the time t can be estimated by solving the set of differential equations, including the equations of conservation of heat and mass

$$\gamma_P M_P^{(f)}\frac{dT_i^{(f)}}{dt} = \Delta_f - Q_T^{(f)} - L_P m_P(\nu^{(p)}/\nu_1)Q_1^{(f)} - Q_R^{(f)},$$
$$\frac{dM_P^{(f)}}{dt} = (\nu^{(p)}/\nu_1)m_P Q_1^{(f)}, \tag{17}$$

where Δ_f is the heat evolving in the particle volume per unit time, $Q_R^{(f)}$ is the heat removed from the surface of the particles by radiative heat exchange, and γ_P is the specific heat of the particles.

At large drops of T_e, the set of equations (15) and (16) allows for an analytic solution for a small particle spacing, i.e., for $(h - R_1 - R_2) \ll R_f$. Under these conditions, the temperatures $T_i^{(f)}$ can be taken to be the same ($T_i^{(f)} = T_{i0}$), and we can find the distributions of T_e and C_j by Eq. (5). The temperature T_{i0} can be found by integrating the set of equations

$$\sum_{f=1}^{2}\Delta_f = \sum_{f=1}^{2}\left[\gamma_P M_P^{(f)}\frac{dT_{i0}}{dt} + Q_T^{(f)}\right.$$
$$\left. + L_P m_P(\nu^{(p)}/\nu_1)Q_1^{(f)} + Q_L^{(f)}\right], \tag{18}$$
$$\frac{dM_P^{(f)}}{dt} = (\nu^{(p)}/\nu_1)m_P Q_1^{(f)}.$$

For steady-state combustion, we can determine the temperatures T_{i0} and $T_i^{(f)}$ assuming that the derivatives $dT_i^{(f)}/dt$ and dT_{i0}/dt in Eqs. (17) and (18), respectively, are zero.

Estimations performed using Eqs. (18) showed that the approach of two particles with sharply different radii can significantly cut the burning time for the finer particle. This is well demonstrated from the curves in Fig. 3, representing the time dependence of the relative radius R_1/R_{1b} of carbon particles with the density ρ = 2 g/cm^2. We estimated the burnout time of the particles suspended in air in the radiation field of intensity $I = 4 \times 10^7$ W/m^2 for unit absorption factors. The curves are plotted at the initial radii of the particles R_{1b} = 15 μm, R_{2b} = 50 μm and at $h = R_{1b} + R_{2b}$ (curve *1*), $h = 2R_{2b}$ (curve *2*), and $h = \infty$ (curve *3*).

Thus, in the quasi-steady-state approximation, we solved the problems of the free and complicated (by internal heat generation) combustion of noninteracting and two interacting arbitrarily shaped immobile particles in the case of a rapid chemical reaction on their surfaces. Under the conditions considered, free combustion proceeds at a constant temperature irrespective of the particle shape and spacing. The surface shape may

influence the rate of combustion in the case of heavily elongated and oblate particles. When the particles approach each other, the rate of combustion of the finer particle can decrease significantly if its radius is far less than that of the coarser one.

REFERENCES

1. *Foundations of Practical Theory of Combustion,* Ed. by V. V. Pomerantsev (Énergiya, Leningrad, 1973).
2. L. A. Vulis, *Thermal Conditions of Combustion* (Gosénergoizdat, Moscow, 1958).
3. E. R. Shchukin, Z. L. Shulimanova, and Yu. I. Yalamov, Fiz. Goreniya Vzryva, No. 4, 42 (1982).
4. V. I. Bukatyĭ, I. A. Sutorikhin, and A. M. Shaĭduk, in *Proceedings of II All-Union Conference on Propagation of Laser Radiation in Disperse Media, Obninsk, 1982*, Part II, pp. 141–144.
5. V. I. Bukatyĭ, O. A. Kosykh, and M. Yu. Sverdlov, Fiz. Goreniya Vzryva, No. 4, 51 (1985).
6. D. A. Frank-Kamenetskiĭ, *Diffuison and Heat Transfer in Chemical Kinetics* (Nauka, Moscow, 1967).
7. A. M. Golovin, Yu. G. Degtev, V. V. Kuryatnikov, and V. R. Pesochin, Fiz. Goreniya Vzryva **30** (1), 20 (1994).
8. A. N. Tikhonov and A. A. Samarskii, *Equations of Mathematical Physics* (Nauka, Moscow, 1977; Pergamon, Oxford, 1964).

Translated by K. Shakhlevich

Technical Physics, Vol. 47, No. 11, 2002, pp. 1350–1358. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 5–13.

THEORETICAL AND MATHEMATICAL PHYSICS

Excitation, Generation, and Propagation of Soliton-Like Spin-Wave Pulses in Ferromagnetic Films: Numerical Calculation and Experiment

M. P. Kostylev and N. G. Kovshikov

St. Petersburg State Electrotechnical University (LETI), ul. Prof. Popova 5, St. Petersburg, 197376 Russia

e-mail: eivt@eltech.ru

Received September 18, 2001; in final form, April 26, 2002

Abstract—The excitation, collision, and detection of soliton-like spin-wave pulses in ferromagnetic films are numerically simulated. The theoretical dependence of the pulse peak power at the output of the delay line on the peak power at the input of the input transducer is constructed for the first time. The experiment is compared with measurements. The shape of the nonlinear magnetization pulse excited and propagating in the film is theoretically studied as a function of the carrier frequency position relative to the backward volume spin-wave (BVSW) spectrum. The head-on collision of soliton-like BVSW pulses in ferromagnetic films is simulated for the first time. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Envelope solitons, which appear in theory as solutions to the nonlinear Schrödinger equation, have been the subject of many monographs (see, e.g., [1–4]). In essence, these solitons are nonlinear wave packets retaining their shape during propagation, i.e., not subjected to dispersion spreading. As applied to physical systems, these solitons have been studied for light propagating in a fiber (see, e.g., [5]). Nonlinear excitations have also been investigated in magnetic systems (magnetic solitons) [6–9].

Below, we will consider only one specific type of magnetic solitons, namely, spin-wave (SW) envelope solitons in ferromagnetic films (FFs). Theoretically, the possibility of SW envelope solitons existing in FFs was first estimated in [10, 11]. In these works, as well as in [12], it was shown that an equation of motion for the envelope of a packet of magnetostatic (spin) waves in FFs, which has the form of the nonlinear Schrödinger equation (NSE), can be derived from the equation of motion of the magnetic moment and a set of magnetostatics equations or directly from the spin-wave spectrum with a nonlinear correction to the frequency.

Experimentally, SW envelope solitons were first observed in [13]. These were solitons of high-dispersion dipole–exchange SWs propagating in perpendicularly magnetized FFs with pinned surface spins. These solitons featured a specific envelope shape and the decreased decay of the pulse peak amplitude, which is associated with nonlinear compression (the so-called "self-induced transparency" curve [13]).[1] These solitons were observed in narrow frequency ranges of strong dispersion that correspond to the dipole gap of the SW spectrum. Solitons of weakly dispersive dipole spin waves propagating in perpendicularly magnetized films with unpinned surface spins were first examined in [14, 15]. In [14], the nonlinear compression of SW pulses attendant to the soliton formation was observed; however, the shape of the pulses remained unknown. In [15], the shape of the envelope and its evolution during the soliton formation and propagation were observed. The authors of [16] succeeded in observing the formation of envelope solitons of highly dispersive quasi-surface dipole–exchange SWs in a tangentially magnetized ferromagnetic film with pinned spins. Here, unlike the case of a perpendicularly magnetized film, the solitons of quasi-surface SWs were observed at the low-frequency edges of the dipole gaps of the spectrum. Spin-wave solitons of BVSWs also propagating in tangentially magnetized FFs were first studied in [17].

Theoretically, the formation of SW envelope solitons in YIG films is conventionally explained in terms of the NSE model, which describes the motion of the packet envelope, which has the meaning of the magnetization precession amplitude. To date, the NSE model has also been applied to numerically evaluating the evo-

[1] In the journal JETP Lett. **38**, 413 (1983), Fig. 3 has an error: the abscissa should be the input power.

1063-7842/02/4711-1350$22.00

lution of a nonlinear SW pulse in FFs and the formation of the SW envelope bright soliton [9, 18–21].

Note, however, that the dispersion relation for the NSE in the linear approximation represents a quadratic spectrum:

$$\omega(k) = \omega(k_0) + V_g(k - k_0) + \frac{D}{2}(k - k_0)^2. \qquad (1)$$

Unlike the actual SW spectrum in FFs, this spectrum does not have the cutoff frequency and is symmetric about the point $k = 0$. We are aware of only one theoretical work [22] where this feature of the SW spectrum in FFs was taken into account by introducing an additional nonlocal term into the NSE. The result was an analytical solution in the form of a precessing soliton. Such a solution, however, is not general and does not cover the various situations that may arise when the carrier frequency of the nonlinear pulse approaches the SW cutoff frequency.

Moreover, the homogeneous nonlinear Schrödinger equation ignores the excitation of "nonlinear" variable magnetization pulses by microstrip transducers during the formation of an envelope soliton. Therefore, some physically reasonable boundary or initial conditions are usually used in numerical calculations (see, e.g., [23]).

It is well known that the NSE soliton envelope retains its shape upon pulse collision. For solitons of the forward volume spin wave (FVSW) envelope in FFs, this was first demonstrated in [24, 25]; for solitons of the BVSW envelope, in [26]. Theoretically, the head-on collision of SW envelope solitons in FFs has not yet been studied. Neither has the effect of variably magnetized counterpropagating pulses excited by microstrip transducers on the head-on collision of the solitons.

The aim of this work is to numerically simulate the formation of soliton-like spin-wave pulses in FFs with regard for their actual spectrum, decay, and excitation by microstrip transducers. Analytical results will be compared with measurements. In addition, based on the same model taking into account the actual spectrum, as well as SW decay and excitation, the collision of soliton-like spin-wave pulses in FFs will be numerically evaluated.

It should be noted that we will often use the term "soliton-like pulse," rather than "spin-wave envelope soliton," to stress that our numerical simulation does not rely on the NSE-based model. In place of it, we take advantage of a set of nonlinear differential equations derived from first principles. This set describes the excitation and motion of the spatial Fourier components of the variable magnetization vector in the nonlinear pulse. The associated theory was developed in [27] to study the parametric amplification of SW envelope solitons in FFs by means of parallel magnetic pumping. We used the phenomenological approach. Starting

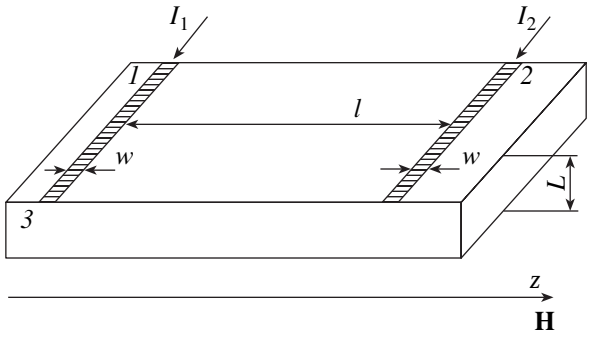


Fig. 1. Structure being simulated. (*1*, *2*) Microstrip transducers and (*3*) ferromagnetic film.

equations were the Landau–Lifshitz equation of motion of magnetic moment and a set of magnetostatics equations.

The set of differential equations obtained in [27] is also applicable to our case if the parametric action on the medium is absent; i.e., it can be applied for describing the excitation and evolution of nonlinear magnetization pulses. Below, we will analyze numerical solutions to a set of coupled nonlinear differential equations derived by modifying the set in [27] to our case.

1. NUMERICAL MODEL

Consider a model of SW delay line (Fig. 1) on FF *3* of thickness *L*. Two parallel SW microstrip transducers (SWMTs) *1* and *2* spaced *l* apart are placed on the film surface. Their width in the *z* direction is *w*. The transducers serve to excite and detect SWs. The film is magnetized to saturation by a permanent magnetic field of strength H. To be specific, we consider the formation and collision of BVSW envelope solitons. This means that the permanent magnetic field **H** is applied in the film plane along the *z* direction and plane spin waves excited by SWMTs propagate in the direction opposite to and along *H* (note that the associated theory can readily be generalized to other FF magnetization configurations).

In transducers *1* and *2*, microwave currents $j_{1,2}(z, t) = J_{1,2}(z, t)\cos(\omega_0 t + \Phi_{1,2})$ pass. The envelopes of their linear densities are $J_{1,2}(t, z)$; their initial phases, $\Phi_{1,2}$; and carrier frequency, ω_0. Let the currents be uniform along the strips. The microwave magnetic fields of the currents excite BVSWs propagating in the form of packets in both directions from the transducers along the permanent magnetic field. If one of the current is absent (namely $J_2 = 0$), SWMT *2* is used as a detector of the BVSW pulse excited by SWMT *1*.

It was already mentioned that the generation, evolution, and collision of BVSW nonlinear packets will be treated with the results of [27]. The set of equations obtained in [27] and describing the motion of the Fou-

rier components of the variable magnetization is simplified in our case to the form

$$\partial m_k^1/\partial t + i[\omega_H + \omega_M/2(1 - P_k)]m_k^1$$
$$+ i\int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \int_{-\infty}^{\infty} dk_3 \Bigg[\sum_{p_1, p_2, p_3 = 1, 2} T_{k_1 k_2 k_3 k}^{p_1, p_2, p_3} m_{k_1}^{p_1} m_{k_2}^{p_2} m_{k_3}^{p_3}$$
$$\times \delta(p_1 + p_2 + p_3 - 4)\delta(k_1 + k_2 + k_3 - k) \Bigg] \qquad (2)$$
$$= |g|\mu_0[J_{1k}(t)\cos(\omega_0 t + \phi_1)$$
$$+ J_{2k}(t)\cos(\omega_0 t + \phi_2)]\frac{k}{|k|}Y_k,$$

$$\partial m_k^2/\partial t - i[\omega_H + \omega_M/2(1 - P_k)]m_k^2$$
$$- i\int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \int_{-\infty}^{\infty} dk_3 \Bigg[\sum_{p_1, p_2, p_3 = 1, 2} T_{k_1 k_2 k_3 k}^{p_1, p_2, p_3} m_{k_1}^{p_1} m_{k_2}^{p_2} m_{k_3}^{p_3}$$
$$\times \delta(p_1 + p_2 + p_3 - 5)\delta(k_1 + k_2 + k_3 - k) \Bigg] \qquad (3)$$
$$= -|g|\mu_0[J_{1k}(t)\cos(\omega_0 t + \phi_1)$$
$$+ J_{2k}(t)\cos(\omega_0 t + \phi_2)]\frac{k}{|k|}Y_k.$$

Here,

$$P_k = 1 - [1 - \exp(-|k|L)/(|k|L)] \qquad (4)$$

is a matrix element of dipole–dipole interaction,

$$Y_k = 1/\sqrt{(2\pi)^{3/2}}[1 - \exp(-|k|L)]/2L \qquad (5)$$

is the overlap integral for the Fourier component of the current magnetic field and the "membrane function" for the BVSW lowest mode, $J_{1k}(t)$ and $J_{2k}(t)$ are the spatial Fourier components of the current pulse envelopes of the exciting transducers, $T_{k_1 k_2 k_3 k}^{p_1, p_2, p_3}$ is the coefficient of nonlinear four-wave interaction between SWs in an FF, $\omega_H = |g|\mu_0 H$, and $\omega_M = |g|\mu_0 M_0$.

It should be noted that if the right-hand sides of (2) and (3) are zero (that is, external sources of excitation are absent) and the dispersion relation for BVSWs in FFs $\omega(k)^2 = \omega_H(\omega_H + \omega_M - \omega_M P_k)$ [28] is approximated by quadratic spectrum (1), this set is totally equivalent to the NSE with loss. Also, this set can be reduced to the NSE with loss by diagonalizing the linear part of the equations and subsequently applying inverse Fourier transformation.

2. NUMERICAL SIMULATION OF EXCITATION AND EVOLUTION OF NONLINEAR BVSW PULSES AND COMPARISON WITH EXPERIMENT

By numerically solving the set of Eqs. (2) and (3), we simulated the excitation of nonlinear SW packets by the input microstrip transducer, their propagation in the FF, and the detection of the pulsed signal with the output microstrip transducer.

Initially, the linear parts of the equations were diagonalized to simplify calculations and make the model more "physical" analytically. The statement of the problem in terms of new variables (scalar Fourier amplitudes of the SW packet [3, 9]) halves the number of equations and allows one to get rid of fast-oscillating factors, namely, the carrier frequencies of the input microwave current and the variable magnetization pulse excited by this current in the film.

In addition, the model was discretized by replacing integration over the longitudinal wave number k in Eqs. (2) and (3) by summation over a discrete set of wave numbers. In this case, the set of two integro-differential equations transforms into an infinite set of first-order differential equations. Since such a procedure is equivalent to replacing the real space by the cyclic one, the discrete step Δk of the wave number was taken such that the spacing between the transducers plus the spatial length of the magnetization pulse excited was much less than the cyclic space length $2\pi/\Delta k$. With this condition for Δk satisfied, the requirement that the magnetization pulse spectrum width be much larger than the step of the wave number is fulfilled simultaneously. The number of wave number discrete values taken into account in the calculation was selected such that the range of k used in the calculation was much wider than the SW excitation bandwidth $\delta k = 2\pi/W$ of the microstrip transducer. At the same time, this range should be wider than the k spectrum of the variable magnetization pulse at any stage of its evolution due to the action of a nonlinear medium.

The set of first-order differential equations thus obtained was numerically solved with the fourth-order Runge–Kutta method. Eventually, we obtained the spectra of the spatial scalar Fourier amplitudes of the SW packet at various time instants.

The inverse transition to the variables of Eqs. (2) and (3), i.e., to the Fourier amplitudes of the variable magnetization, and subsequent inverse Fourier transformation made it possible to find the spatial profiles of the nonlinear SW packet at various time instants. Then, from the calculated spectrum of the Fourier amplitudes for the variable magnetization, we evaluated the space–time profile of the microwave field induced by the SW packet in the output transducer. This profile was then converted to the microwave power profile at the output of the output transducer.

The envelopes of microwave pulses of the currents in SWMTs *1* and *2* were set in the form of a trapezoid. Such a form takes into account the nonrectangularity of the current pulse at the output of the microwave modulator, which applied a signal to the input of the SWMT in our experiment. The inclined sides of the trapezoid simulated the leading and trailing edges of an actual input pulse. We also set the half-height width of the

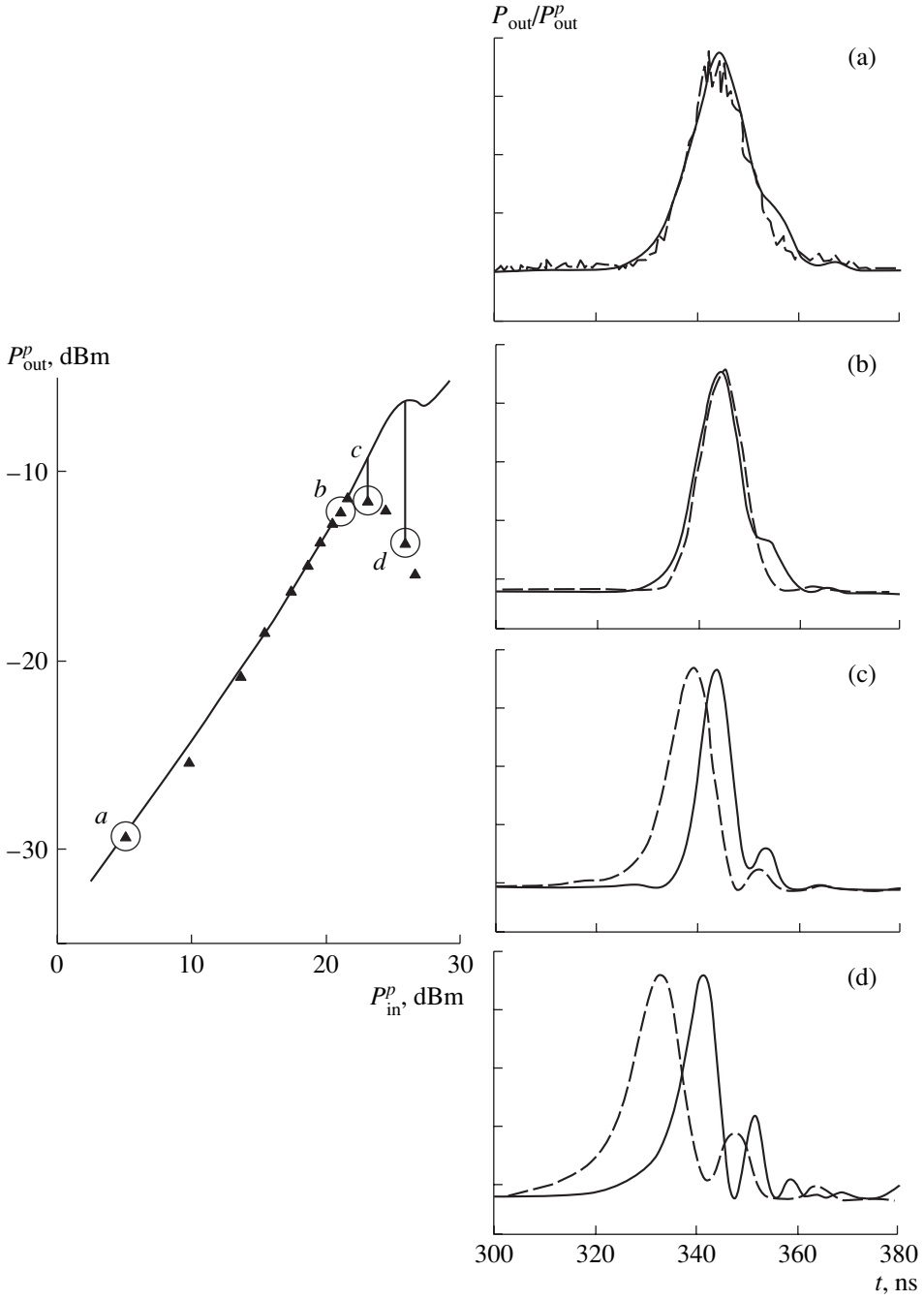


Fig. 2. P^p_{out} vs. P^p_{in} dependence (on the left) and output pulse waveforms (on the right) for a 25-ns-wide input pulse. Continuous curves, calculation; circles on the left and dashed lines on the right, experiment. P^p_{in} = 0.006 (a), 0.23 (b), 0.385 (c), and 0.724 W (d). For the coincidence of the experimental and theoretical plots, the theoretical values of the input power are increased by 2.6 dB, while those of the output power are decreased by 2.6 dB. On the left, the power is normalized to its peak value P^p_{out}. Parameters used in the calculations are L = 5.1 μm, saturation magnetization 1750 Oe, magnetic loss parameter of the film $2\Delta H$ = 0.6 Oe, H = 1407 Oe, transducer widths w = 50 μm, l = 8.37 mm, τ = 25 ns, pulse rise and fall times (measured at the pulse base) 2 ns, and the pulse carrier frequency 5850 MHz.

trapezoidal current pulses τ, as well as their amplitudes I_1 and I_2, initial phases ϕ_1 and ϕ_2, and rise times.

First, using the above numerical model, we studied the formation of an isolated soliton-like SW pulse (the microwave current I_2 was assumed to equal zero). On the left of Fig. 2, the calculated peak power P^p_{out} of the soliton-like pulse envelope at the output of detecting transducer *2* is shown as a function of the microwave power P^p_{in} at the input of exciting transducer *1*. Diagrams (a–d) on the right show the associated waveforms of the pulses coming from the delay line. The calculation was made for output microwave pulses of duration 25 ns. Other parameters are given in the caption to Fig. 2.

The analytical results were compared with measurements made for this configuration. Triangles on the left and dashed lines in the diagrams on the right are experimental data.

It should be noted that our theory is consistent with the well-known linear theory of spin-wave delay line [28] in the continuous weak signal limit. In the experiments, we could determine conversion losses in the input antenna in the linear mode by comparing the power of the signal incident on the input of the structure with that of the microwave signal reflected from it. In this way, the conversion loss per antenna was estimated at –9.4 dB, which is 2.6 dB greater than the calculated value. The difference was found to be associated with the reactance of the antenna grounding circuits, which was initially ignored in the calculation. The inclusion of this reactance eliminates the difference completely. Therefore, the conversion losses evaluated by (2) and (3) were increased by 2.6 dB (i.e., were equal to the experimental value in the linear mode). In this case, as follows from Fig. 2, the analytical and experimental curves $P^p_{\text{out}} = f(P^p_{\text{in}})$ coincide in a wide power range that covers both the linear and nonlinear modes. Coincidence is also observed for the output pulse envelope profiles taken at the same powers. The discrepancy is seen only in the upper part of the curve: the output pulse amplitudes diverge and the instants of output pulse arrival at the detector differ slightly, while the pulse waveforms coincide fairly accurately.

Such a discrepancy can be explained as follows. Our theoretical model includes only one weak nonlinearity of SWs in a ferromagnetic film, namely, four-wave self-action. However, comparing the theory and the experiment, one can suppose that the peak and the descending portion in the experimental curve are associated with some other nonlinear processes taking place at high input powers. These may be four-wave scattering into short spin waves or into BVSW higher thickness modes. The contribution from these processes shows up as an extra nonlinear decay of the pulse, changing the run of the upper part of the curve $P^p_{\text{out}} = f(P^p_{\text{in}})$.

The best criterion indicating that an SW pulse is soliton-like is the constancy of the envelope phase within the pulse. As follows from the phase calculation, the point *b* in Fig. 2 is the closest to the one-soliton mode.

Along with experimental verification, we compared our results with numerical calculations within the NSE model. Earlier, the NSE was used to simulate SW envelope solitons, e.g., in [17]. In [17], as in other works, the time profile of the SW dimensionless scalar amplitude was calculated at a point z_0 of the space with the boundary condition in the form of a rectangular pulse originating at the point $z = 0$ (hereafter, this profile will be referred to as initial, although we are dealing with a boundary condition). Therefore, to correctly contrast the two models, we equated the right-hand sides of (2) and (3), responsible for the SW excitation by external current, to zero and used the same time-dependent boundary conditions as in [17] in the next series of calculations. With such an approach, the only difference between the models is that the dispersion relation is more complicated than (1).

The calculation showed that if the carrier frequency of the initial pulse of the variable magnetization scalar amplitude is away from the upper cutoff frequency of the BVSW spectrum, the output profiles are nearly coincident provided that the pulse is sufficiently long (lasts more than 10 ns as in [17]. Under this condition, the frequency spectrum of the initial pulse has a width that is much narrower than the range of exchangeless BVSWs). This is an expected result since the BVSW spectrum in this case is close to quadratic spectrum (1). With the carrier frequency shifted toward the cutoff frequency, the output pulses had various widths, which indicates the inadequacy of the NSE-based model in this situation.

In the second series of comparative calculations, we used a set of equations with the nonzero right-hand sides. The rectangular pulse of the input current and the initial pulse of the variable magnetization scalar amplitude [17] had the same durations. It turned out that the shape of the magnetization pulse becomes distorted even in the course of excitation (this is especially true for the pulses of width less than 30 ns with a FF thickness and transducer width as in [17]). This is because the microstrip transducer has a finite bandwidth of excited SW wave vectors. This initial distortion affects the shape of the pulse during its subsequent evolution, which makes the discrepancy with the results of the NSE-based model much greater as the carrier frequency approaches the SW cutoff frequency. Note that our results can be compared with those of the NSE-based model only qualitatively, because we had to compare the calculated profiles of the variable magnetization at the same amplitude of the initial pulse of the variable magnetization. However, in our model, the shape of the initial pulse is not specified but is derived analytically from the shape and amplitude of the input

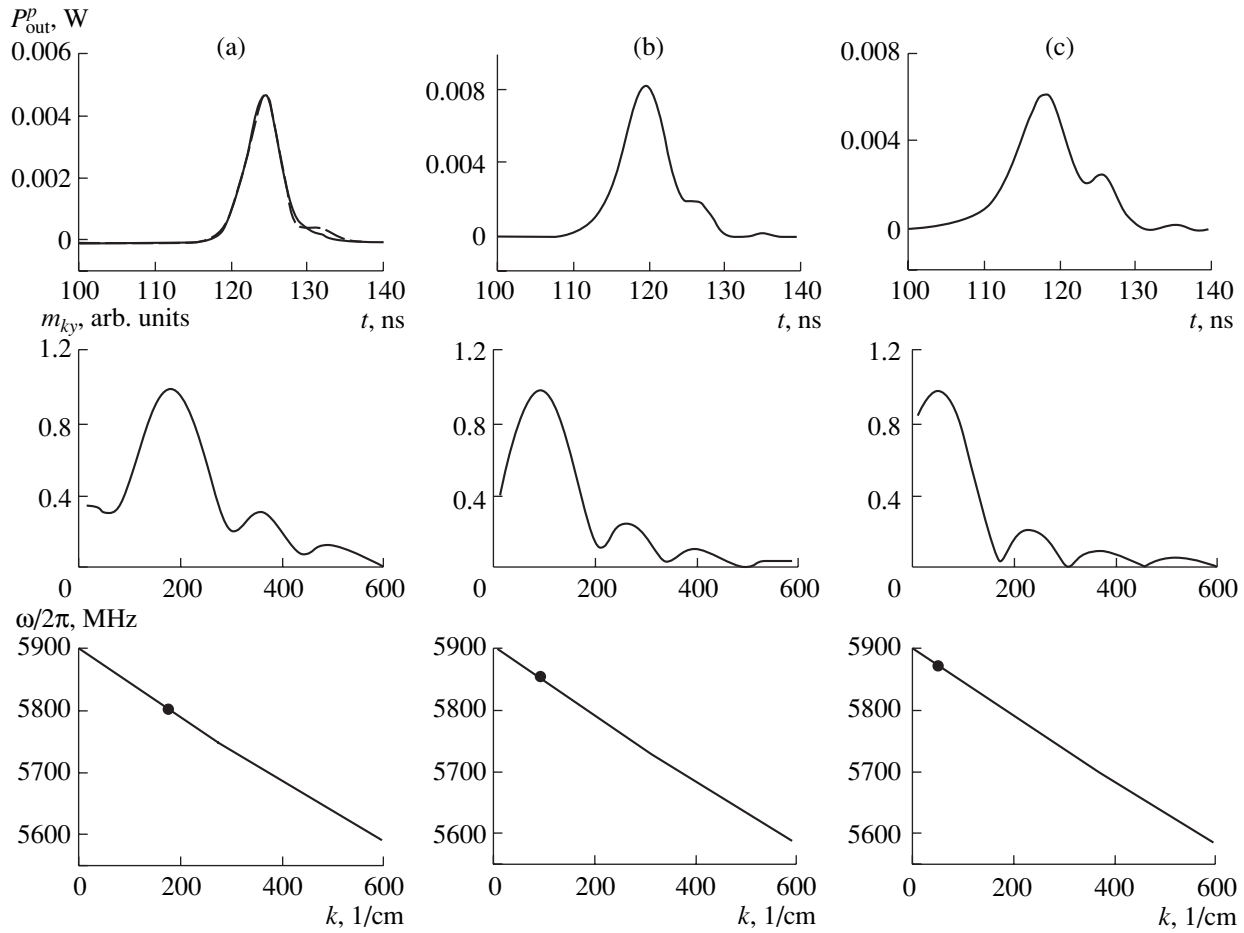


Fig. 3. Calculated diagrams of the variable magnetization pulse envelope for different positions of the carrier frequency Δω relative to the upper boundary of the BVSW spectrum (upper row), the amplitude spectra of the pulse spatial frequencies (middle row), and the dispersion curve for the BVSW lowest mode in the FF (lower row). Circles in the lower row depict the carrier frequency position. The dashed curve in panel (a) (upper row) illustrates experimental data from [17]. To bring the experimental and analytical curves into coincidence, the experimental output power of the pulse is increased by 2.8 dB (this value was obtained by comparing the measured and calculated losses inserted by the prototype in the continuous low-signal regime). The parameters used in the calculation are L = 7.1 μm, saturation magnetization 1750 Oe, magnetic loss parameter of the film $2\Delta H$ = 0.85 Oe, constant magnetic field strength H = 1407 Oe, transducer widths w = 50 μm, l = 4 mm, τ = 13 ns, pulse rise and fall times (measured at the pulse base) 1 ns, and input microwave pulse amplitude 0.25 W (the amplitude of the microwave current of transducer *1* (I_1 = 0.05 A). The carrier frequency equals (a) 5800 MHz ($\Delta\omega/2\pi$ = 100 MHz), (b) 5850 MHz ($\Delta\omega/2\pi$ = 50 MHz), and (c) 5875 MHz ($\Delta\omega/2\pi$ = 25 MHz).

current pulse. Therefore, the initial pulse in our model has a shape that is more complex than a rectangle. Because of this, it is necessary to define its amplitude so that it depends on the pulse shape only slightly and allows for the numerical comparison of the two models. Such an analysis is beyond the scope of this work.

Figure 3 (the upper row) shows the calculated profiles of the envelope of soliton-like SW pulses vs. the pulse carrier frequency position relatively to the BVSW spectrum upper limit for relatively short (half-height width of 14 ns) input microwave pulses. Panel (a) (soliton-like pulse) corresponds to experimental parameters from [17]. Panels (b) and (c) show how the pulse would change if its carrier frequency were closer to the beginning of the BVSW spectrum.

For comparison, the middle row shows the spectra of the spatial Fourier amplitudes of the pulse envelopes; the lower row, dispersion curves for the BVSW lowest mode (filled circles indicate the carrier frequency).

From Fig. 3, it follows that if the carrier wave number is small, the shape of the output pulse is distorted and its duration increases. These distortions appear because of the filtering out of some spatial (and, hence, time) frequencies of the pulse upon its excitation and propagation in a nonlinear medium.

3. NUMERICAL SIMULATION OF HEAD-ON COLLISION BETWEEN SOLITON-LIKE BVSW PULSES

In this section, we study the effect of SW excitation by the SW microstrip transducers on the head-on colli-

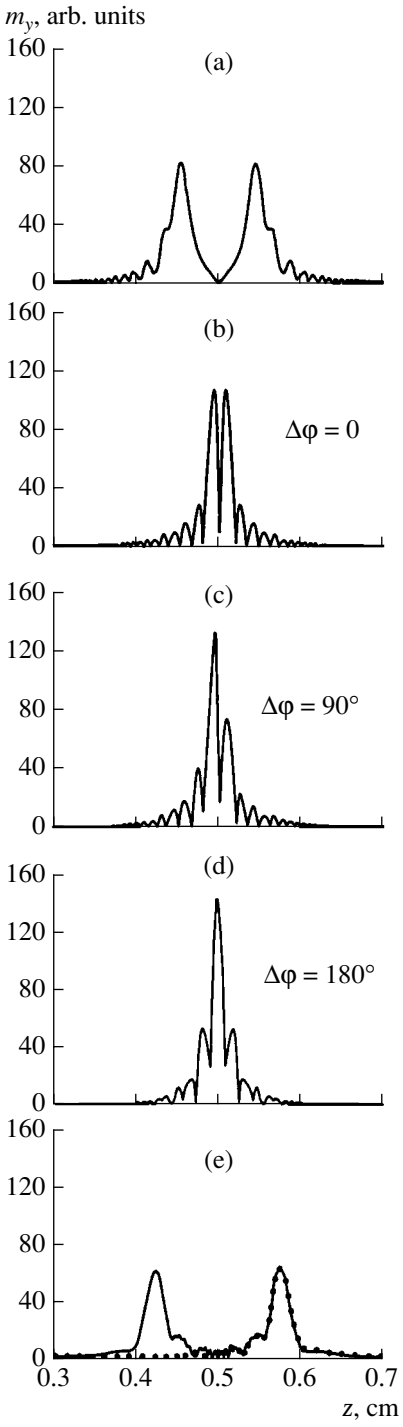


Fig. 4. Numerical simulation of the collision of soliton-like SW pulses in the ferromagnetic film. (a) Profiles of the envelope of variable magnetization pulses immediately before collision, (b) instant of the complete overlap of the pulses in the case of their in-phase excitation ($\Delta\varphi = 0$), (c) instant of the complete overlap of the pulses in the case $\Delta\varphi = 90°$, (d) instant of the complete overlap of the pulses in the case of their antiphase excitation ($\Delta\varphi = 180°$), and (e) pulse profiles after collision. Circles in panel (e) outline the pulse profile in the absence of a counterpropagating pulse. The spacing between the transducers is 8 mm. The other parameters are the same as in Fig. 3a.

sion of soliton-like SW pulses in FFs. Emphasis will be on the effect of the initial phase of exciting pulses of the microwave current from the transducers on the collision scenario.

The calculation was performed with the same model (expressions (2) and (3)) and the same numerical method (see Sect. 2). To simulate the excitation of two counterpropagating pulses of variable magnetization, the right-hand sides of (2) and (3) were set as the Fourier transforms of the microwave currents in microstrip transducers *1* and *2* (Fig. 1).

Figure 4 shows the result of the numerical simulation. The parameters used in the calculation and the pulse shape at the point of collision in the absence of the counterpropagating pulse correspond to panel (a) in Fig. 3. The pulse shapes are shown (a) before collision, (b–d) at the instant of collision, and (e) immediately after collision. In Fig. 4e, filled circles depict the shape of the pulse excited by left transducer *1* for the case when the pulse did not collide with the counterpropagating one (i.e., for $I_2 = 0$) (the time instants are the same). It is seen that the shapes of the pulses that underwent and did not undergo collision coincide within a graphical accuracy throughout the length of the pulses except for the "tail," where the pulses moving apart still continue interfering.

It is known that BVSW pulses generated by SWMTs in opposite directions have an initial phase difference $\Delta\phi = \pi$ (as also follows from equations of motion (2) and (3)). This is because the BVSWs are excited with the transducer microwave current component normal to the FF plane. In the bulk of the film, this component has different signs on the opposite sides of the plane passing through the longitudinal symmetry axis of the transducer normally to the film. Such an antisymmetric distribution of the exciting component of the magnetic field results in the excitation of variable magnetization pulses opposite in phase. For comparison: in a perpendicularly magnetized FF, forward volume spin waves (FVSWs) are excited by the component of the transducer microwave magnetic field that is parallel to the film plane. This component does not change sign in the film. Accordingly, two FVSWs propagating in opposite directions from the transducer have zero phase shift.

Thus, two transducers fed by in-phase currents emit counterpropagating BVSW pulses with a phase difference $\Delta\phi = \pi$. When the pulses collide, the "nonlinear interference" pattern is symmetric about the interference minimum (Fig. 4b) at any time instant. If the input currents are in antiphase, SWs are excited in phase and the interference pattern is symmetric about the interference maximum (Fig. 4c).

Comparing the peak amplitude of the interference pattern (Fig. 4c) with that of the single soliton-like SW pulse at the same time instant, we see that the former differs little from the latter doubled; in other words, ordinary linear interference takes place. To elucidate

why the nonlinear properties of FFs have no effect on the interference pattern, we will perform an estimation in terms of the theory for the envelopes of SW counterpropagating nonlinear pulses [29]. The accuracy of the equations for the envelopes derived in [29] is quite sufficient for these estimates.

Consider the characteristic time of nonlinear pulse compression, the so-called "nonlinear" time T_n, which is introduced in describing SW envelope solitons with the use of the NSE (see, e.g., [17]), and compare it with the time of spatial pulse overlapping. For the interaction of counterpropagating pulses, T_n is found from a set of equations for the envelopes of nonlinearly coupled counterpropagating pulses [29]. It is given by

$$T_n = \pi/(2|S||a|^2), \quad (6)$$

where S is the nonlinear coupling coefficient for counterpropagating pulses and a is the pulse envelope amplitude that has a meaning of the magnetization precession angle (we consider that the counterpropagating pulses have the same amplitude).

Note that the propagation of nonlinear pulses in FFs with the subsequent formation of soliton-like pulses takes place under conditions of magnetic dissipation in a medium. As a result, the pulse power continuously decreases. Therefore, the value of T_n should be calculated from the value of a at the instant the pulses meet in the space. Performing such a calculation with the parameters in Fig. 4, we find that $T_n \approx 500$ ns. Since the time of the spatial overlapping of counterpropagating pulses is on the order of several nanoseconds, the interference pattern cannot be nonlinearly distorted for such a short time compared with T_n.

CONCLUSION

We for the first time simulated the excitation, collision, and detection of soliton-like BVSW pulses in FFs by numerically solving a set of nonlinear differential equations of motion for the spatial Fourier amplitudes of variable magnetization pulses. The analytical results are in good agreement with our experimental data and those taken from the literature.

We for the first time calculated the pulse peak power at the output of the prototype delay line as a function of the peak power at the input of the input transducer.

For the first time, the shape of a nonlinear magnetization pulse generated and propagating in the film was simulated in view of the position of the carrier frequency relative to the BVSW spectrum and spin-wave generation by microstrip transducers.

For the first time, the simulation of the head-on collision of soliton-like BVSW pulses showed that, with single-soliton power levels of colliding pulses, the nonlinearity of the spin system has no time to contribute significantly to the interference pattern because of the small time of pulse overlapping. As a result, we are dealing with usual linear interference.

ACKNOWLEDGMENTS

The authors thank Prof. C.E. Patton (University of Colorado, USA) for experimental data to Fig. 3a and also Prof. B.A. Kalinikos (St. Petersburg State Electrotechnical University, Russia) for encouragement and valuable discussions.

This work was supported by the Russian Foundation for Basic Research (grant no. 99-02-16370), the Ministry of Education of the Russian Federation (program 304.1, project code 269), and INTAS (grant no. 99-1812).

REFERENCES

1. S. P. Novikov, S. V. Manakov, L. P. Pitaevskiĭ, and V. E. Zakharov, *Theory of Solitons: The Inverse Scattering Method* (Nauka, Moscow, 1980; Consultants Bureau, New York, 1984).
2. M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981; Mir, Moscow, 1986), SIAM Studies in Applied Mathematics, no. 4.
3. F. Calogero and A. De Gasperis, *Spectral Transform and Solitons: Tools to Solve and Investigate Nonlinear Evolution Equations* (North-Holland, Amsterdam, 1982; Mir, Moscow, 1985).
4. L. D. Faddeev and L. A. Takhtajan, *Hamiltonian Methods in the Theory of Solitons* (Nauka, Moscow, 1986; Springer-Verlag, Berlin, 1987).
5. G. Agrawal, *Nonlinear Fiber Optics* (Academic, San Diego, 1995; Mir, Moscow, 1996).
6. A. I. Akhiezer and A. E. Borovik, Zh. Éksp. Teor. Fiz. **52**, 508 (1967) [Sov. Phys. JETP **25**, 332 (1967)].
7. A. M. Kosevich, B. A. Ivanov, and A. S. Kovalev, *Nonlinear Waves of Magnetization: Dynamical and Topological Solitons* (Naukova Dumka, Kiev, 1988).
8. V. S. L'vov, in *Solitons*, Ed. by S. E. Trullinger, V. E. Zakharov, and V. L. Pokrovsky (Elsevier, Amsterdam, 1986).
9. V. S. L'vov, *Nonlinear Spin Waves* (Nauka, Moscow, 1987).
10. V. P. Lukomskiĭ, Ukr. Fiz. Zh. **23**, 134 (1978).
11. A. K. Zvezdin and A. F. Popkov, Zh. Éksp. Teor. Fiz. **84**, 606 (1983) [Sov. Phys. JETP **57**, 350 (1983)].
12. A. N. Slavin and I. V. Rojdestvenski, IEEE Trans. Magn. **30**, 37 (1994).
13. B. A. Kalinikos, N. G. Kovshikov, and A. N. Slavin, Pis'ma Zh. Éksp. Teor. Fiz. **38**, 343 (1983) [JETP Lett. **38**, 413 (1983)].
14. P. De Gasperis, R. Marchelli, and G. Miccoli, Phys. Rev. Lett. **59**, 481 (1987).
15. B. A. Kalinikos, N. G. Kovshikov, and A. N. Slavin, IEEE Trans. Magn. **26**, 1477 (1990); Phys. Rev. B **42**, 8658 (1990).
16. B. A. Kalinikos, N. G. Kovshikov, P. A. Kolodin, *et al.*, Solid State Commun. **74**, 989 (1990).

17. M. Chen, M. A. Tsankov, J. M. Nash, and C. E. Patton, Phys. Rev. B **49**, 12773 (1994).
18. A. N. Slavin and G. M. Dudko, J. Magn. Magn. Mater. **86**, 115 (1990).
19. M. Chen, M. A. Tsankov, J. M. Nash, and C. E. Patton, Phys. Rev. B **49**, 12773 (1994).
20. A. D. Boardman, S. A. Nikitov, K. Xie, and H. Mehta, J. Magn. Magn. Mater. **145**, 357 (1995).
21. V. V. Kiselev, A. P. Tankeyev, A. V. Kobelev, and A. G. Shagalov, J. Phys.: Condens. Matter **11**, 3461 (1999).
22. V. V. Kiselev and A. P. Tankeev, J. Phys.: Condens. Matter **8**, 10219 (1996).
23. M. Chen, J. M. Nash, and C. E. Patton, J. Appl. Phys. **73**, 3906 (1993).
24. B. A. Kalinikos and N. G. Kovshikov, in *Digests of INTERMAG'93 Conference, Stockholm, 1993*, p. AE-03.
25. B. A. Kalinikos and N. G. Kovshikov, Pis'ma Zh. Éksp. Teor. Fiz. **60**, 290 (1994) [JETP Lett. **60**, 305 (1994)].
26. M. A. Tsankov, M. Chen, and C. E. Patton, J. Appl. Phys. **76**, 4274 (1994).
27. B. A. Kalinikos and M. P. Kostylev, Zh. Tekh. Fiz. **70** (2), 136 (2000) [Tech. Phys. **45**, 277 (2000)].
28. V. F. Dmitriev and B. A. Kalinikos, Izv. Vyssh. Uchebn. Zaved., Fiz. **31** (11), 24 (1988).
29. B. A. Kalinikos and M. P. Kostylev, IEEE Trans. Magn. **33**, 3445 (1997).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1359–1363. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 14–18.

GASES AND LIQUIDS

Effective Relaxation Time of the Impurity Concentration in a Medium with a Slightly Nonuniform Diffusion Coefficient

A. A. Dubkov, A. A. Mal'tsev, and E. L. Pankratov

Lobachevsky State University, pr. Gagarina 23, Nizhni Novgorod, 603950 Russia
e-mail: dubkov@rf.unn.runnet.ru, maltsev@rf.unn.runnet.ru, elp@rf.unn.runnet.ru

Received January 23, 2002; in final form, May 6, 2002

Abstract—Based on the previously introduced integral criterion and a new approach to estimating the time characteristics of diffusion, the relaxation time of impurity concentration in a medium with a slightly nonuniform diffusion coefficient is determined. The diffusion coefficient profile that cuts the relaxation time is found.

INTRODUCTION

The mathematical theory of diffusion has been developed in great detail (see, e.g., [1–3]). Yet, a number of problems of both theoretical and applied interest have remained unresolved. One is finding the impurity concentration relaxation time in a medium with a space-dependent diffusion coefficient [4]. The urgency of this problem stems from the wide application of diffusion technologies, especially in the fabrication of semiconductors.

In this work, we consider a 1D bounded medium with a slightly nonuniform diffusion coefficient, which adequately reflects the situation in various semiconductors that have different diffusion coefficients in the active (working) regions [5–7]. With the integral criterion introduced previously [8–12] and a new approach to estimating the time characteristics of diffusion processes, we determine the impurity concentration relaxation time in the medium. Based on an expression for the relaxation time, we optimize the spatial dependence of the diffusion coefficient to accelerate impurity "smearing."

STATEMENT OF THE PROBLEM AND SOLUTION METHOD

We consider a 1D medium with reflecting boundaries $x = 0$ and $x = L$ and the slightly nonuniform diffusion coefficient $D = D(x)$. At the time instant $t = 0$, an impurity of unit mass

$$\int_0^L C(x, t)dx = 1$$

with a given distribution of the initial concentration $C(x, 0) = f(x)$ is inserted in the medium. With time, the impurity distribution tends toward the stationary value $C(x, \infty) = 1/L$. Let us define the effective relaxation time of the impurity concentration at a given point of the medium (point of observation) and the diffusion coefficient profile that reduces this time.

As is known, the space–time distribution of the impurity concentration $C(x, t)$ satisfies the diffusion equation

$$\frac{\partial C(x, t)}{\partial t} = \frac{\partial}{\partial x}\left[D(x)\frac{\partial C(x, t)}{\partial x}\right] = -\frac{\partial G(x, t)}{\partial x}, \quad (1)$$

where $G(x, t)$ is the impurity flux. This equation is complemented by the initial, $C(x, 0) = f(x)$, and boundary, $G(0, t) = G(L, t) = 0$, conditions.

The space dependence of the diffusion coefficient can be represented as the sum of its mean,

$$D_0 = \frac{1}{L}\int_0^L D(x)dx,$$

and a small correction taking into account the difference between the law of diffusion coefficient variation and the mean value of the diffusion coefficient; that is,

$$D(x) = D_0(1 + \varepsilon g(x)), \quad 0 < \varepsilon \ll 1, \quad |g(x)| \le 1.$$

In view of such a representation, the diffusion equation takes the form

$$\frac{\partial C(x, t)}{\partial t} = D_0\frac{\partial}{\partial x}\left[(1 + \varepsilon g(x))\frac{\partial C(x, t)}{\partial x}\right]. \quad (2)$$

The introduction of the small parameter ε makes it possible to seek a solution to Eq. (2) in the form of the power series in ε [12].

A solution to diffusion equation (2) (the transient process to settling the equilibrium value of the concentration) has a complicated time dependence. To find the concentration relaxation time, it is necessary to use some criterion that allows one to determine the desired

time from a given transient. The relaxation time can be found with an "equiareal rectangle" as in [8–12]:

$$\Theta(x) = [C(x,0) - C(x,\infty)]^{-1} \times \int_0^\infty [C(x,t) - C(x,\infty)]dt. \quad (3)$$

Such an approach to finding the relaxation time is applicable to transients that are monotonic in time. From the practical viewpoint, of most interest is the situation where the initial distribution is set at one of the boundaries of the medium and the relaxation time is looked for at a point near the other boundary. In this case, transients are always monotonic in time.

We will seek a solution to Eq. (2) in the form of the expansion

$$C(x,t) = \sum_{k=0}^{\infty} \varepsilon^k C_k(x,t). \quad (4)$$

Substituting (4) into (2) and equating the coefficients by the same powers of ε on both sides of the diffusion equation, we arrive at a set of equations for the functions $C_k(x, t)$:

$$\frac{\partial C_0(x,t)}{\partial t} = D_0 \frac{\partial^2 C_0(x,t)}{\partial x^2},$$

$$\frac{\partial C_k(x,t)}{\partial t} = D_0 \frac{\partial^2 C_k(x,t)}{\partial x^2} + D_0 \frac{\partial}{\partial x}\left[g(x) \frac{\partial C_{k-1}(x,t)}{\partial x} \right], \quad k \geq 1. \quad (5)$$

Similarly, substituting (4) into the initial and boundary conditions, we arrive at a set of conditions for the functions $C_k(x, t)$:

$$\frac{\partial C_k(0,t)}{\partial x} = \frac{\partial C_k(L,t)}{\partial x} = 0, \quad k \geq 0, \quad (6)$$

$$C_0(x,0) = f(x), \quad C_k(x,0) = 0, \quad k \geq 1. \quad (7)$$

The linear relationship (3) between the relaxation time and the concentration $C(x, t)$ allows us to calculate the relaxation time in the form of the sum

$$\Theta(x) = \Theta_0(x)\left(1 + \sum_{k=1}^{\infty} \varepsilon^k \tau_k(x) \right), \quad (8)$$

where

$$\Theta_0(x) = [Lf(x) - 1]^{-1} \int_0^\infty [LC_0(x,t) - 1]dt \quad (9)$$

is the zeroth-order approximation for relaxation time (3).

The functions $\tau_k(x)$ entering into (8) and defined by the expression

$$\tau_k(x) = \int_0^\infty C_k(x,t)dt \Big/ \int_0^\infty [C(x,t) - C(x,\infty)]dt \quad \text{for } k \geq 1, \quad (10)$$

will hereafter be called relative corrections to the relaxation time in a homogeneous medium (to the zeroth-order approximation). In view of the first equation in set (5) and relationship (3), the zeroth-order approximation of the relaxation time can be represented as

$$\Theta_0(x) = [6D_0 d(x)(1 - Lf(x))]^{-1}, \quad (11)$$

where

$$d(x) = [6LxF_0(x) - 3F_2(L) + 6L(F_1(L) - F_1(x)) - 2L^2 - 3x^2]^{-1},$$

$$F_i(x) = \int_0^x y^i f(y)dy.$$

The first-order correction to the relaxation time can be obtained by substituting solutions to the first and second equations in set (5) into relationship (10):

$$\tau_1(x) = \frac{6\left[\int_x^L g(v)[v - LF_0(v)]dv - \int_0^L g(v)v(vL^{-1} - F_0(v))dv \right]}{3F_2(L) - 6L[F_1(L) - F_1(x)] + 2L^2 + 3x^2 - 6LxF_0(x)}. \quad (12)$$

Since the parameter ε is small and the absolute value of the function $g(x)$, which introduces the diffusion coefficient variation, is limited, we can reject terms nonlinear in ε in (4) and (8) and consider only the linear approximation.

CALCULATION OF THE CORRECTIONS TO THE RELAXATION TIME

Let at the initial time instant the impurity be totally concentrated at a point near the left-hand boundary of the medium: $f(x) = \delta(x)$. Note that, because of initial conditions (7), the delta function $\delta(x)$ should be considered as one-sided [13]. Of most interest is the correction $\tau_1(x)$ at the point $x = L$, which is easy to calculate from (12). Placing the origin at the point $x = L/2$, one readily obtains

$$\tau_1(L) = 6L^{-3} \int_{-L/2}^{L/2} g(z + L/2)(z^2 - L^2/4)dz. \quad (13)$$

From (13), it follows that only the even (relative to the middle of the medium) component of the space

dependence of the diffusion coefficient makes a contribution to the correction $\tau_1(L)$.

By way of example, consider a medium with a planar layered spatial structure [5, 6]. For such a medium, the space dependence of the diffusion coefficient can be conveniently represented with Walsh functions $Wal(k, x)$, where k is the Walsh function order [14]. Only even-order Walsh functions contribute to the correction $\tau_1(L)$.

It can be shown that spectral methods of representing the disturbance profiles of the diffusion coefficient allow for the calculation of the correction $\tau_1(x)$ for any profile $g(x)$. In this case, the computing procedure can be greatly simplified. Let us represent the spatial dependence of the diffusion coefficient as the series over some orthogonal basis of functions $\{g_m(x)\}$:

$$D(x) = D_0\left[1 + \varepsilon\sum_{m=1}^{\infty}\beta_m g_m(x)\right], \qquad (14)$$

where β_m are the coefficients of expansion of the function $g(x)$ over the orthogonal basis.

In this case, since (12) is linear in $g(x)$, the correction $\tau_1(x)$ is the superposition of the corrections $\eta_{1m}(x)$ depending on the functions $g_m(x)$; that is,

$$\tau_1(x) = \sum_{m=0}^{\infty}\beta_m\tau_{1m}(x). \qquad (15)$$

Each of the corrections $\tau_{1m}(x)$ is calculated from (12) subject to $g(x) = g_m(x)$. For example, in the case of the trigonometric basis $\{\cos 2\pi mx/L, \sin 2\pi mx/L\}$, one readily finds from (13) the corrections $\tau_{1m}(L)$ for individual spatial harmonics:

$$\tau_{1\cos m}(L) = 3/(\pi m)^2, \quad \tau_{1\sin m} = 0, \qquad (16)$$

where $\tau_{1\cos m}(L)$ and $\tau_{1\sin m}(L)$ are corrections corresponding to the functions $g(x) = \cos 2\pi mx/L$ and $g(x) = \sin 2\pi mx/L$.

Figure 1 shows the dependence of the correction $\tau_1(L)$ on the number m of a cosinusoidal spatial harmonic. As the frequency of the diffusion coefficient disturbance profile rises (see the table and Fig. 1), the effect of nonuniformity on the relaxation time diminishes because of its averaging. The rapid convergence of spectral expansions allows one to reject fine features of the diffusion coefficient disturbance profile during its optimization and approximate the function $g(x)$ by the first one or two spectral components.

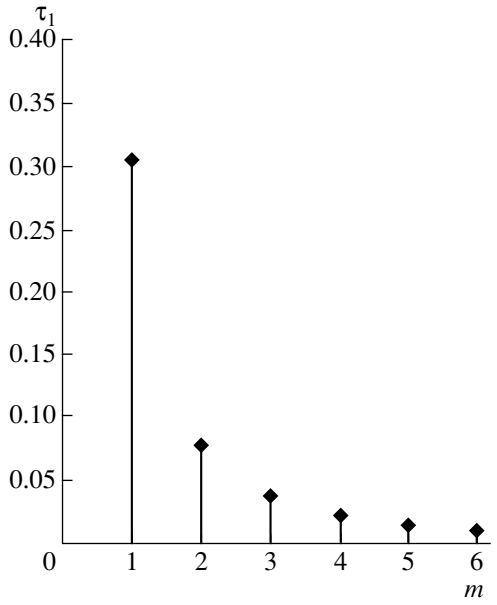


Fig. 1. First-order correction to the relaxation time that corresponds to the nonuniform diffusion coefficient $D(x) = D_0[1 + \mu\cos(2\pi mx/L)]$ vs. the frequency m of its spatial dependence.

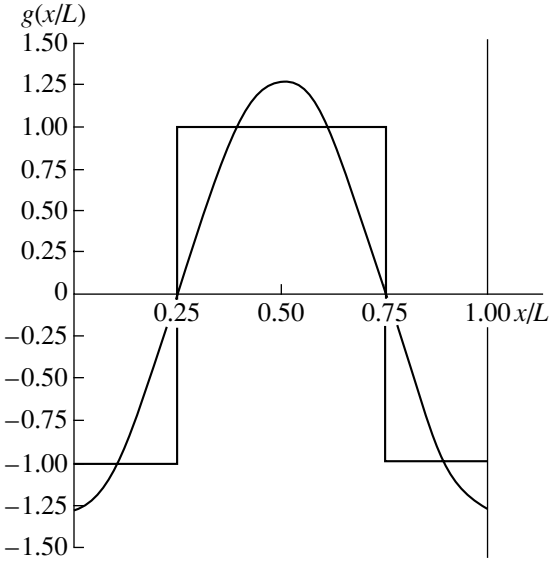


Fig. 2. Three-layer and harmonic structures accelerating the diffusion process.

IMPURITY DIFFUSION ACCELERATION BY SELECTING THE OPTIMAL SPATIAL STRUCTURE OF THE MEDIUM

The above analysis shows that only even (relative to the middle of the medium) profiles of $g(x)$ contribute to the correction $\tau_1(L)$. If the law of diffusion coefficient variation is approximated by asymmetric (relative to the middle of the medium) sinusoidal functions or odd-order Walsh functions, the diffusion process is described by the mean of the function $D(x)$. The difference between the relaxation time in an inhomogeneous medium and its zeroth-order approximation is maximum in a three-layer medium (see the table and Fig. 1). The relaxation time depends essentially on the alternation of the layers, which is reflected by the dependence

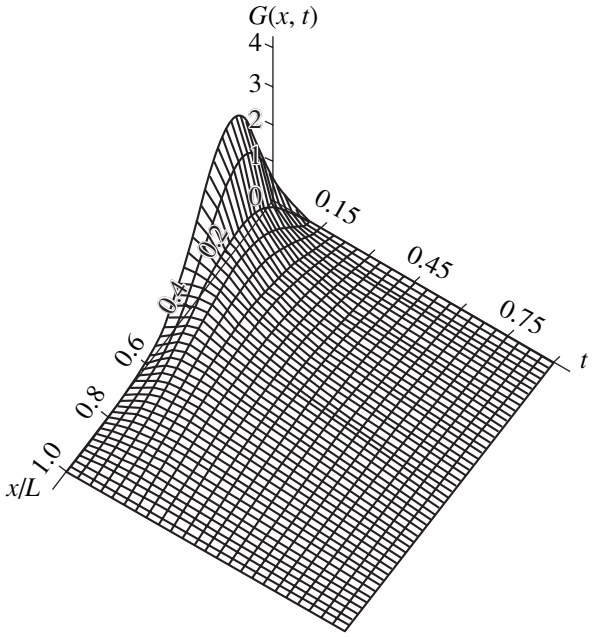


Fig. 3. Space–time distribution of the impurity flux $G(x, t)$ for the initial distribution $f(x) = \sqrt{\frac{2}{\pi W}}\exp\left(-\frac{x^2}{2W}\right)$ with a small parameter W ($W = 0.04$).

of the correction $\tau_1(L)$ on the sign of $g(x)$ in (13). Figure 2 shows the diffusion coefficient profiles in the case when the diffusion proceeds faster than in the homogeneous medium. The change of sign in this profile retards the process. Correction (13) for the three-layer medium shown in Fig. 2 equals $\tau_1(L) = -3/8$. The correction corresponding to the harmonic profile of the diffusion coefficient shown in Fig. 2 has a greater value: $\tau_{1\cos m}(m = 1) \approx -3/\pi^2$. Physically, the obtained dependence of the relaxation time on the phase φ of the harmonic perturbation profile of the diffusion coefficient and on the shift α of the Walsh function is related to the space–time distribution of the impurity flux. At the initial time instant, the flux is the greatest near the point $x = 0$. Since the flux is very intense, the impurity travels the interval $0 \le x \approx L/4$, "overlooking" the differences between the current diffusion coefficient and its mean value. Subsequently, for a time $T \ll \Theta_0$, the maximal value of the flux shifts toward the middle of the medium, decreasing in absolute value and the diffusion is the most intense at the center of the medium. Hence, to cut the concentration relaxation time, it is necessary to raise the diffusion coefficient in the interval $L/4 \le x \le 3L/4$. At the boundary $x = L$, the reflection of the impurity from the boundary becomes significant. For this reason, the resultant flux (the superposition of the forward and reflected fluxes) in the interval $3L/4 \le x \le L$ is low and this interval is gradually filled with the impurity. This fill affects the relaxation time insignificantly. Figure 3 exemplifies the space–time impurity flux distribution obtained by numerical simulation.

Corrections to the relaxation time for several even-order Walsh functions

Function	First-order correction to the relaxation time, τ_1
$Wal(2, L)$	3/8
$Wal(4, L)$	3/32
$Wal(6, L)$	3/128

The applicability domain of the linear (with respect to ε) approximation used to calculate the relaxation time can be found by numerically simulating the diffusion process. For the impurity initial distribution and the diffusion coefficient profiles considered in this work, the calculation of the relaxation time in the linear (in ε) approximation was shown to be valid up to $\varepsilon \approx 0.1$.

CONCLUSION

Based on a new approach to finding the time characteristics of diffusion processes, we studied the dependence of the impurity concentration relaxation time in a weakly inhomogeneous medium on the space variation of the diffusion coefficient. The relaxation time was estimated with linear (in the diffusion coefficient) approximation. Such an estimation yields simple and clear relationships for the relaxation time and facilitates greatly the optimization of the diffusion coefficient spatial profile derived from the spectral representation of the diffusion coefficient. The optimization of the diffusion coefficient spatial structure allowed us to find the diffusion coefficient profile that accelerates impurity diffusion compared with the case of a homogeneous medium with the diffusion coefficient D_0.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 00-15-96620 and 02-02-17517) and by INTAS (grant no. 2001–0450).

REFERENCES

1. A. I. Raĭchenko, *Mathematical Theory of Diffusion as Applications* (Naukova Dumka, Kiev, 1981).
2. J. Crank, *The Mathematics of Diffusion* (Oxford Univ. Press, Oxford, 1975).
3. J. P. Stark, *Solid State Diffusion* (Wiley, New York, 1976; Énergiya, Moscow, 1980).
4. B. A. Zon, S. B. Ledovskiĭ, and A. N. Likholet, Zh. Tekh. Fiz. **70** (4), 38 (2000) [Tech. Phys. **45**, 419 (2000)].
5. I. P. Stepanenko, *Foundations of Microelectronics* (Sov. Radio, Moscow, 1980).

6. V. G. Gusev and Yu. M. Gusev, *Electronics* (Vysshaya Shkola, Moscow, 1991).
7. K. V. Shalimova, *Physics of Semiconductors* (Énergoatomizdat, Moscow, 1985).
8. D. A. Garanin, Phys. Rev. E **54**, 3250 (1996).
9. V. T. Coffey, D. S. F. Crothers, and Yu. P. Kalmykov, Phys. Rev. E **55**, 4812 (1997).
10. A. N. Malakhov, Chaos **7**, 488 (1997).
11. A. N. Malakhov and E. L. Pankratov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **44** (4), 367 (2001) [Radiophys. Quant Electron. **44**, 339 (2001)].
12. A. A. Mal'tsev and E. L. Pankratov, in *Proceedings of the 5th Scientific Conference on Radiophysics*, Ed. by A. V. Yakimov (Talam, Nizhni Novgorod, 2001), pp. 211–212.
13. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1968; Nauka, Moscow, 1977).
14. I. S. Gonorovskiĭ, *Radio Circuits and Signals* (Nauka, Moscow, 1977).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1364–1368. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 19–23.

GASES AND LIQUIDS

Enskog Kinetic Equation Modified for a Dense Soft-Sphere Gas

V. I. Kurochkin

Samara State University, Samara, 443011 Russia

e-mail: kurochk@ssu.samara.ru

Received October 30, 2001; in final form, February 27, 2002

Abstract—The Enskog kinetic equation is modified for dense soft-sphere gases and gas mixtures when the diameter of the particles depends on their relative velocity. Analytical expressions for the transport coefficient of a monoatomic dense gas are derived, and the viscosity coefficient is calculated based on smoothed experimental data. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The Boltzmann kinetic equation is valid only for low- and normal-density gases. It was derived under a number of assumptions, specifically, that the distribution function (DF) is time-independent on the time scale on the order of the collision time and space-independent on the scale on the order of molecule size. In addition, multiple collisions were ignored. These assumptions are valid for point particles and become incorrect if the particle size is comparable to the particle free path. In medium- and high-density gases, for which the density parameter $\varepsilon = n\sigma^3$ (where n is the concentration of the particles and σ is their characteristic size) is no longer negligibly small and may even be comparable to unity, the assumptions involved in the derivation of the Boltzmann equation become invalid. Enskog was the first to modify the Boltzmann equation for dense gases [1]. He took into account the nonlocality of the DF in the collision integral within distances on the order of the molecule size, the constraint of gas particles, and the screening effect and, hence, is valid for the model of hard elastic spheres. Other models (kinetic equations) for medium- and high-density gases currently available [2–5] are, as rule, very complicated and cannot be generalized for gas mixtures. Therefore, the Enskog theory can be applied with advantage in calculations that do not require too high an accuracy. A significant drawback of the Enskog kinetic equation is the improper temperature dependence of the transport coefficient; therefore, when using this equation in practice, researchers have to relate the Enskog theory to experimental data for the gas compressibility with some uncertainty. The basic reason for the improper temperature dependence in the Enskog theory is the roughness of the hard-sphere model, which ignores one of the fundamental properties of gases, namely, their compressibility. At low density of a gas, the Boltzmann equation allows for the accurate calculation of the transport coefficients even within the hard-sphere model, since the purely kinetic mechanism of momentum and energy transfer prevails in this case. Collisional momentum and energy transfer is significant in medium-density gases and dominates in dense gases; therefore, the form of the interaction potential in the near-zone field plays a decisive role in this case. This disadvantage of the hard-sphere model is eliminated in the "macroscopic" soft-sphere model [6–8], which includes the temperature dependence of the hard-sphere diameter. This dependence is derived from the temperature dependences of the transport coefficients at normal densities or by any other way. In equilibrium statistical mechanics, the following approximation is used [9]:

$$\frac{\sigma}{\sigma_0} = \left(\frac{T_0}{T}\right)^{1/s}, \tag{1}$$

where s usually equals 10–12 and T_0 and σ_0 are constants (characteristic temperature and size of the particles).

This approximation is valid for a gas where particles interact as repulsive centers with a potential

$$\varphi(r) = \varepsilon\left(\frac{\sigma_0}{r}\right)^s. \tag{2}$$

In this case, $T_0 = \varepsilon/k_B$, where k_B is the Boltzmann coefficient and ε is a force constant. Such an expedient provides satisfactory agreement between theoretical and experimental data. At the same time, it shows that the interaction potential parameters are self-consistent with the kinetic equation and do not agree with the theory "from outside." The generalization of the soft-sphere theory for a gas mixture is difficult because of a great number of parameters to be found as a function of temperature. In the absence of a unified reliable method for finding these dependences, such models are of minor value.

The aim of this work is to justify and generalize the macroscopic soft-sphere model involved in the Enskog

theory for the microscopic level and construct a simple kinetic model of dense soft-sphere gases that allows for the extension to gas mixtures.

KINETIC EQUATION

The idea of the macroscopic soft-sphere model is that the diameter of colliding particles depends on the interaction parameters unlike hard spheres. In the simplest case, one can use the power dependence of the relative diameter on the relative impact velocity g:

$$\frac{\sigma}{\sigma_0} = \left(\frac{g_0^2}{g^2}\right)^{\mu}. \tag{3}$$

Comparing (3) with a formula for the distance of closest approach between particles with potential (2), we can identify the quantity g_0^2 as $4\varepsilon/m$ and the quantity $1/\mu$ as the exponent s in initial potential (2). The scattering differential cross section in this case can be written as

$$d\sigma = \sigma_0^2\left(\frac{4\varepsilon}{mg^2}\right)^{2\mu}(\mathbf{g}\cdot\mathbf{k})d\mathbf{k}, \tag{4}$$

where $\mathbf{k}$ is the unit vector directed from the center of one particle to that of the other at the collision instant.

Generalizing the result obtained for gas mixtures, we find for the differential cross section of scattering particles of sort i by those of sort j

$$d\sigma_{ij} = \sigma_{0ij}^2\left(\frac{2\varepsilon_{ij}}{m_{ij}g_{ij}^2}\right)^{2\mu}(\mathbf{g}_{ji}\cdot\mathbf{k})d\mathbf{k}, \tag{5}$$

where $\mathbf{g}_{ji} = \mathbf{v}_j - \mathbf{v}_i$ is the relative velocity and $m_{ij} = m_i m_j/(m_i + m_j)$ is the reduced mass.

The Enskog equation for a mixture of medium-density soft-sphere gases formally has the same form as for hard spheres [10] with the exception that the differential cross section is expressed by formula (5):

$$\frac{\partial f_i}{\partial t} + \mathbf{v}_i\cdot\frac{\partial f_i}{\partial \mathbf{r}_i} = \int\{\chi_{ij}(\mathbf{r}_i, \mathbf{r}_j')f_i'(\mathbf{r}_i)f_j'(\mathbf{r}_j') - \chi_{ij}(\mathbf{r}_i, \mathbf{r}_j)f_i(\mathbf{r}_i)f_j(\mathbf{r}_j)\}d\sigma_{ij}d\mathbf{v}_j. \tag{6}$$

Here, f_i is the DF for particles of sort i, $\mathbf{r}_j = \mathbf{r}_i - \sigma_{ij}\mathbf{k}$, and $\mathbf{r}_j' = \mathbf{r}_i + \sigma_{ij}\mathbf{k}$. The primed DF depends only on "past-impact" velocity; hence,

$$\mathbf{v}_i' = \mathbf{v}_i + 2\mu_{ji}(\mathbf{g}_{ji}\cdot\mathbf{k})\mathbf{k}, \quad \mathbf{v}_j' = \mathbf{v}_j - 2\mu_{ji}(\mathbf{g}_{ji}\cdot\mathbf{k})\mathbf{k}, \tag{7}$$

where $\mu_{ji} = m_i/(m_i + m_j)$.

The function χ_{ij} in Eq. (6) is the locally equilibrium pair correlation function [7], well-known in statistical mechanics. Equation (6) differs from the conventional modified Enskog equation for gas mixtures [6, 7] in that the center-to-center distance σ_{ij} between particles at the collision instant depends on the particle relative velocity according to (3). Equation (5) is nonlocal. To derive a local equation (for a single point $\mathbf{r} = \mathbf{r}_i$), it is necessary to expand all the functions in Eq. (6) into the Taylor series, as in [1], and leave pair derivatives with respect to coordinates. Then, the collision integral on the right of (6) can be represented as the sum of local integrals:

$$I_{\text{cm}} = I_0 + I_1 + I_2. \tag{8}$$

Here,

$$I_0 = \int\{f_i'f_j' - f_if_j\}\chi_{ij}d\sigma_{ij}d\mathbf{v}_j, \tag{9}$$

$$I_1 = \frac{1}{2}\int\{f_i'f_j' + f_if_j\}\mathbf{k}\cdot\nabla\chi_{ij}d\sigma_{ij}d\mathbf{v}_j, \tag{10}$$

$$I_2 = \int\{f_i'\nabla f_j' + f_i\nabla f_j\}\cdot\mathbf{k}\chi_{ij}d\sigma_{ij}d\mathbf{v}_j. \tag{11}$$

These integrals differ from those in [1] by the form of the differential cross section and also by the fact that the function χ_{ij} is under the integral, since it depends on the particle size. Instead of χ_{ij}, one usually takes the pair correlation function. For a gas mixture, it can be calculated, e.g., by the Mansuri–Karnahan–Sterling–Leland formula [11]. For a single-component gas, one can use the expansion of the pair correlation function in the form [1]

$$\chi = \chi(bn) = 1 + 0.625bn + 0.287(bn)^2 + 0.11(bn)^3, \tag{12}$$

where

$$b = \frac{2\pi}{3}\sigma^3,$$

and n is the density of the particles.

SOLUTION OF THE KINETIC EQUATION FOR A SINGLE-COMPONENT GAS BY THE CHAPMEN–ENSKOG METHOD

We will solve Eq. (6) for a single-component gas in the near-equilibrium state. Our goal is to derive transport equations, obtain expressions for the flux vectors, and calculate the transport coefficients. Multiplying Eq. (6) by the additive coefficients 1, $m\mathbf{v}$, and $(1/2)mv^2$ in view of (8)–(11) and integrating over velocities, we come to conventional equations for mass, momentum, and energy transfer [1], where the viscous stress tensor and the heat flux vector have the form

$$\mathbf{P} = \mathbf{P}_K + \mathbf{P}_{\Phi 1} + \mathbf{P}_{\Phi 2}, \quad \mathbf{q} = \mathbf{q}_K + \mathbf{q}_{\Phi 1} + \mathbf{q}_{\Phi 2}. \tag{13}$$

Here,

$$\mathbf{P}_K = mn\langle\mathbf{CC}\rangle, \quad \mathbf{q}_K = \frac{1}{2}mn\langle\mathbf{CC}^2\rangle, \tag{14}$$

$$\mathbf{P}_{\Phi 1} = \frac{1}{2}mn^2\langle\langle(\mathbf{g}\cdot\mathbf{k})\mathbf{kk}\chi\sigma\rangle\rangle, \tag{15}$$

$$\mathbf{P}_{\Phi 2} = \frac{1}{2}mn^2\langle\langle(\mathbf{g}\cdot\mathbf{k})\mathbf{k}\mathbf{k}[\mathbf{k}\cdot\nabla\ln(f/f_1)]\chi\sigma\rangle\rangle, \quad (16)$$

$$\mathbf{q}_{\Phi 1} = \frac{1}{4}mn^2\langle\langle[(\mathbf{g}\cdot\mathbf{k})^2 + 2(\mathbf{g}\cdot\mathbf{k})(\mathbf{C}\cdot\mathbf{k})]\mathbf{k}\chi\sigma\rangle\rangle, \quad (17)$$

$$\mathbf{q}_{\Phi 2} = \frac{1}{4}mn^2\langle\langle[(\mathbf{g}\cdot\mathbf{k})^2 + 2(\mathbf{g}\cdot\mathbf{k})(\mathbf{C}\cdot\mathbf{k})]\mathbf{k} \times[\mathbf{k}\cdot\nabla\ln(f/f_1)]\chi\sigma\rangle\rangle. \quad (18)$$

$\mathbf{C} = \mathbf{v} - \mathbf{u}$ is the intrinsic velocity of the particles; $\mathbf{u}$ is the hydrodynamic velocity; and the subscripts K and Φ refer to contributions due to the molecule flux (kinetic energy) and collisions [1], respectively. Ordinary and double broken brackets mean averaging over one DF and two DFs and cross section, respectively:

$$\langle\phi\rangle = \frac{1}{n}\int f\phi d\mathbf{v}, \qquad \langle\langle\phi\rangle\rangle = \frac{1}{n^2}\int\int\int f f_1\phi d\sigma d\mathbf{v} d\mathbf{v}_1. \quad (19)$$

Relationships (15)–(18) differ from those given in [1] in that χ and σ are functions of the relative velocity and stand under the integral.

Expanding the DF in the small parameter which is the Knudsen number, that is, assuming that $f = f^{(0)}(1 + Kn\varphi)$, we find that $f^{(0)}$ is the locally Maxwellian DF and, for the first-order correction to φ,

$$-n^2 I_0(\varphi) = Df^{(0)} - I_1(f^{(0)}) \equiv \mathbf{A}\cdot\nabla\ln T + \mathbf{B} : \nabla\mathbf{u}. \quad (20)$$

Here,

$$\mathbf{A} = f^{(0)}\left[\left(\omega^2 - \frac{3}{2} - \frac{1}{nk_\mathrm{B}}\frac{\partial p}{\partial T}\right)\mathbf{C} - \mathbf{M}\right], \quad (21)$$

$$\mathbf{B} = f^{(0)}\left[\frac{m}{T}\{\mathbf{C}\mathbf{C}\} + \left(\frac{2}{3}\omega^2 - \frac{p}{nk_\mathrm{B}T}\right)\mathbf{I} - \mathbf{W}\right], \quad (22)$$

$$\mathbf{M} = \int f_1^{(0)}\left(\omega_1^2 + (\omega_1')^2 - 3 + T\frac{\partial\ln\chi}{\partial T}\right)\mathbf{k}\chi\sigma d\sigma d\mathbf{v}_1, \quad (23)$$

$$\mathbf{W} = \int f_1^{(0)}\left[\frac{m}{k_\mathrm{B}T}(\mathbf{C}_1 + \mathbf{C}_1')\right]\mathbf{k}\chi\sigma d\sigma d\mathbf{v}_1, \quad (24)$$

p is pressure, $\mathbf{I}$ is unit tensor, $\boldsymbol{\omega}_i = \mathbf{C}_i/V_\mathrm{T}$, $V_\mathrm{T} = \sqrt{(2k_\mathrm{B}T)/m}$, and expression

$$\{\mathbf{ab}\} = \frac{1}{2}(\mathbf{ab} + \mathbf{ba}) - \frac{1}{3}(\mathbf{a}\cdot\mathbf{b})\mathbf{I} \quad (25)$$

is valid for any two vectors $\mathbf{a}$ and $\mathbf{b}$.

Note that the vector $\mathbf{M}$ and the tensor $\mathbf{W}$ entering into relationships (21) and (22) do not admit, unlike [1], direct integration over $\mathbf{v}_2$. However, such a procedure is not required to calculate the transport coefficients.

A solution to Eq. (23) can be represented in the form

$$\varphi = -\frac{1}{n}[\mathbf{a}\cdot\nabla\ln T + \mathbf{b} : \nabla\mathbf{u}], \quad (26)$$

where the vectors $\mathbf{a}$ and $\mathbf{b}$ are looked for, as usual, as the expansion in Sonine polynomials.

In the first-order approximation, the expansion in Sonine polynomials for these vectors takes the form

$$\mathbf{a} = a_1\left(\frac{5}{2} - \omega^2\right)\mathbf{C}, \quad \mathbf{b} = b_0\{\boldsymbol{\omega}\boldsymbol{\omega}\}. \quad (27)$$

The coefficients a_1 and b_0 are found by integrating Eq. (23) with appropriate weights [1]. Eventually, we have

$$b_0 = \frac{\langle\mathbf{B} : \{\boldsymbol{\omega}\boldsymbol{\omega}\}\rangle}{[\{\boldsymbol{\omega}\boldsymbol{\omega}\}, \{\boldsymbol{\omega}\boldsymbol{\omega}\}]}, \qquad a_1 = \frac{\left\langle\mathbf{A}\cdot\left(\frac{5}{2} - \omega^2\right)\mathbf{C}\right\rangle}{\left[\left(\frac{5}{2} - \omega^2\right)\mathbf{C}, \left(\frac{5}{2} - \omega^2\right)\mathbf{C}\right]}, \quad (28)$$

where integral brackets are defined as

$$[F, G] = \frac{1}{n^2}\int f^{(0)} f_1^{(0)}[F' + F_1' - F - F_1] \times[G' + G_1' - G - G_1]\chi d\sigma d\mathbf{v} d\mathbf{v}_1. \quad (29)$$

When directly calculating the means in relationships (28), one should pass to variables (relative velocity and center-of-mass velocity), as is usually done in calculating integral brackets. The direct calculation by formulas (28) yields

$$b_0 = -\frac{5\sqrt{2}}{8\sqrt{\pi}V_\mathrm{T}\sigma_0^2 R_0}\left(1 + \frac{2}{5}n_* R_1\right), \quad (30)$$

$$a_1 = -\frac{15\sqrt{2}}{32\sqrt{\pi}V_\mathrm{T}\sigma_0^2 R_0}\left(1 + \frac{3}{5}n_* R_4\right). \quad (31)$$

Here,

$$n_* = \frac{2\pi}{3}n\sigma_0^3, \quad R_0 = \frac{\sqrt{\pi}}{24}r_2^{(5)}, \quad R_1 = \frac{5}{6}r_3^{(2)} - \frac{1}{5}r_3^{(4)}, \quad R_4 = \frac{25}{36}r_3^{(2)} - \frac{8}{15}r_3^{(4)} + \frac{1}{9}r_3^{(6)}, \quad (32)$$

$$r_k^{(q)} = \frac{8\tau^{-k\mu}}{\sqrt{\pi}}\int_0^\infty \exp(-\gamma^2) \times\chi(n_*\tau^{-3\mu}\gamma^{-6\mu})\gamma^{2+q-2k\mu}d\gamma, \quad (33)$$

and $\tau = (k_\mathrm{B}T)/\varepsilon$ is the reduced dimensionless temperature.

FLUX VECTORS AND TRANSPORT COEFFICIENTS

Substituting the DF obtained into expressions (16)–(21) for flux vectors and performing necessary calculations, we find

$$\mathbf{P} = p\mathbf{I} - 2\eta\{\nabla\mathbf{u}\} - \kappa(\nabla\cdot\mathbf{u}), \quad \mathbf{q} = -\lambda\nabla T, \quad (34)$$

where the hydrostatic pressure p is given by the equation of state

$$p = \left(n + \frac{1}{3}n_*^2 r_3^{(2)}\right)k_B T. \quad (35)$$

From relationship (35), it follows that the second virial coefficient for the equation of state is given by

$$B = \frac{8\sqrt{\pi}}{9}\sigma_0^3\Gamma\left(\frac{5-6\mu}{2}\right)\tau^{-3\mu}, \quad (36)$$

which coincides with the expression for the second virial coefficient for repulsion center potential (2).

According to (16), the shear viscosity coefficient η and the thermal conductivity λ include both usual (kinetic) contributions and those due to collisional momentum and energy transfer; that is, $\eta = \eta_K + \eta_{\Phi 1} + \eta_{\Phi 2}$ and $\lambda = \lambda_K + \lambda_{\Phi 1} + \lambda_{\Phi 2}$, where

$$\eta_K = \frac{1}{2}b_0 T, \quad \eta_{\Phi 1} = \frac{1}{5}b_0 k_B T n_* R_2, \quad \eta_{\Phi 2} = \frac{48}{25\pi}\eta_0 n_*^2 R_3, \quad (37)$$

$$\lambda_K = -\frac{5}{4}a_1 V_T^2, \quad \lambda_{\Phi 1} = -\frac{3}{4}a_1 V_T^2 n_* R_5, \quad \lambda_{\Phi 2} = \frac{32}{25\pi}\lambda_0 n_*^2 R_3, \quad (38)$$

where

$$R_2 = \frac{2}{15}r_3^{(4)}, \quad R_3 = r_4^{(3)}, \quad R_5 = \frac{11}{45}r_3^{(4)} - \frac{5}{18}r_3^{(2)}, \quad (39)$$

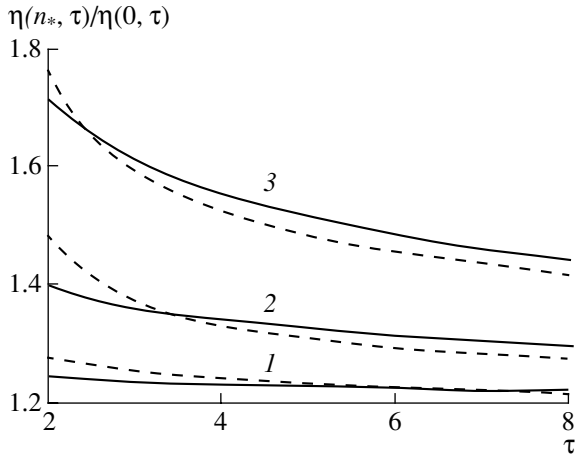
$$\eta_0 = \frac{5\sqrt{mk_B T}}{16\sqrt{\pi}\sigma_0^2}, \quad \lambda_0 = \frac{75k_B\sqrt{k_B T}}{64\sqrt{\pi m\sigma_0^2}}. \quad (40)$$

Eventually, for the shear viscosity coefficient, volume viscosity coefficient, and thermal conductivity, we have

$$\frac{\eta}{\eta_0} = \frac{\left(1+\frac{2}{5}n_* R_1\right)\left(1+\frac{2}{5}n_* R_2\right)}{R_0} + \frac{48}{25\pi}n_*^2 R_3, \quad (41)$$

$$\frac{\kappa}{\eta_0} = \frac{16}{5\pi}n_*^2 R_3, \quad (42)$$

$$\frac{\lambda}{\lambda_0} = \frac{\left(1+\frac{3}{5}n_* R_4\right)\left(1+\frac{3}{5}n_* R_5\right)}{R_0} + \frac{32}{25\sqrt{\pi}}n_*^2 R_3. \quad (43)$$



Reduced viscosity coefficient vs. reduced temperature for $n_* = 0.2$ (*1*), 0.5 (*2*), and 0.8 (*3*). Continuous curves, calculation; dashed curves, smoothed experimental data.

It should be noted that the coefficients R_i are selected so that $R_i = 1$ at $\chi = 1$ for hard spheres ($\mu = 0$). Therefore, expressions (41)–(43) for the transport coefficients within the hard-sphere model coincide with those derived by Enskog. These coefficients specify the temperature dependence of corrections to the transport coefficients in the case of a dense gas, while the coefficient R_0 introduces a difference between the temperature dependences of the viscosity and thermal conductivity for hard- and soft-sphere normal-density gases. The expression for R_0 coincides with the temperature dependence of $\Omega^{(2,2)*}$, the Chapman–Cowling reduced integral for the repulsion center potential; that is, $R_0 \sim \Omega^{(2,2)*} \sim \tau^{-k\mu}$ at $n_* \longrightarrow 0$.

The analytical reduced viscosity coefficient vs. reduced temperature τ is plotted in the figure at different gas densities. In the calculations, we took $\mu = 1/12$. For comparison, smoothed experimental data for inert gases and nitrogen [12] are also shown. At $\tau > 2$, the discrepancy is 4–7%. Thus, our model describes adequately the transport properties of a monoatomic dense gas at sufficiently high temperatures.

CONCLUSIONS

(1) A model of the Enskog kinetic equation for soft-sphere dense gases and gas mixtures is represented.

(2) Analytical expressions for the transport coefficients of a monoatomic soft-sphere dense gas are derived.

(3) The calculated viscosity coefficient in the reduced temperature range $2 < \tau < 8$ for the reduced density $n_* = 0.2$, 0.5, and 0.8 is compared with experimental data. The discrepancy does not exceed 7%.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 01-03-32729).

REFERENCES

1. J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport in Gases* (North-Holland, Amsterdam, 1972; Mir, Moscow, 1976).
2. A. A. Vasserman and I. P. Khasilov, Teplofiz. Vys. Temp. **27**, 35 (1989).
3. V. I. Kurochkin and S. V. Tsaplin, Teplofiz. Vys. Temp. **31**, 903 (1993).
4. D. G. Friend and J. C. Rainwater, Chem. Phys. Lett. **107**, 590 (1984).
5. V. A. Rabinovich and S. B. Kiselev, in *Thermophysical Properties of Substances and Materials* (Izd. Standartov, Moscow, 1978), Vol. 12, p. 124.
6. V. I. Kurochkin, Teplofiz. Vys. Temp. **33**, 161 (1995).
7. J. M. J. Coremans and J. M. J. Beenakeer, Physica (Amsterdam) **26**, 653 (1960).
8. A. B. Medvedev, Teplofiz. Vys. Temp. **33**, 227 (1995).
9. H. C. Andersen, J. D. Weeks, and D. Chandler, Phys. Rev. A **4**, 1597 (1971).
10. H. Beijeren and M. H. Ernst, Physica (Amsterdam) **68**, 437 (1973).
11. J. M. Kincaid, M. Lopez de Haro, and E. G. D. Cohen, J. Chem. Phys. **79**, 4509 (1983).
12. N. J. Trappeiers, A. Botzen, C. A. Ten Seldam, *et al.*, Physica (Amsterdam) **31**, 1681 (1965).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1369–1372. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 24–27.

GASES AND LIQUIDS

Dynamo-Optical Effect in Homogeneous Newtonian Fluids

S. V. Stebnovskii

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

e-mail: root@hydro.nsc.ru

Received March 15, 2002

Abstract—It is experimentally found that homogeneous polar fluids (water, glycerol, ethanol, etc.) become optically anisotropic under shear. As a result, the permittivity, which is a macroscopic characteristic of a fluid, becomes anisotropic as well. Results obtained indicate that quiescent polar fluids have signs of an ordered structure, which readily disappears under weak shear stresses and then relaxes to the initial state for several hours (depending on the fluid temperature). © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The performance of this study has been dictated by the following reason. In our earlier works [1–3] devoted to the stability of the structure of disperse liquids in force fields of arbitrary intensity, the nonequilibrium character of disperse systems (emulsions and suspensions) was revealed in the case when there are no gradient temperature and concentration fields and the system is not exposed to external force fields (including the gravitational field). It was shown that if the initial spacing between dispersed elements is on the order of their sizes, the particles approach and touch each other irrespective of the system's dimensions (the particle sizes are 0.1–1.0 cm and the typical approach velocity is $\sim 10^{-4}$ cm/s). The law of particle interaction was found to be the same as in the classical problem of the elasticity theory concerning two interacting spherical cavities in a stressed elastic space. Therefore, one can assume that the approach of dispersed elements is caused by the weak structural strength (elasticity of shear) of the liquid matrix under weak and slow deformation. This argument was supported by experimental results [4]. In [4], it was found that at low rates of shear, water behaves as a medium with a shear modulus of the order of 10^{-6} Pa. (In those experiments, a cylindrical water volume was slowly rotated around its symmetry axis in a cylindrical vessel. The liquid volume rotated by inertia for a time and then executed weak elastic torsional damped vibrations, indicating that water has a strength of shear.)

Since we consider a fundamental property of fluids, namely, the presence of a nonzero though very weak structural strength in the quasi-stationary state, it seems to be reasonable to study the structural ordering of fluids by an experimental technique other than that used in [4]. This problem is solved in this paper.

VALIDATION OF INVESTIGATION METHOD

It is well known [5, 6] that the permittivity of a condensed medium is a macroscopic characteristic of its molecular structure. Therefore, if a medium is structurally ordered, the variation of its structure under mechanical load must show up through the variation (at least, local) of the permittivity ε. Since the refractive index in a transparent medium is related to the permittivity as $n = \varepsilon^{1/2}$, elastic strains will influence the optical properties of structured transparent media through the tensor ε_{ik}. That is, an optically isotropic medium becomes anisotropic due to the deformation. The photoelasticity method, which is widely used in studies of the elastic stress distribution in a strained transparent solid, is based on this effect. In addition, optical anisotropy arises in moving non-Newtonian fluids, namely, suspensions and colloidal solutions with nonspherical particles, if a velocity gradient is present [7]. In this case, the orientation of nonspherical disperse elements changes in the region of the velocity gradient owing to the shear deformation of the medium. This causes the structure of the medium and, hence, ε to change. The refractive index also changes, and the optical isotropy breaks. This phenomenon is referred to as the dynamo-optical effect or the Maxwell effect.

The question arises of whether the dynamo-optical effect takes place in homogeneous Newtonian fluids. The answer to this question will provide valuable information about the structure of fluids. Indeed, if a fluid has an ordered structure, its permittivity [8] $\varepsilon = 1 + 4\pi Np$ (where N is the number of molecules or molecular clusters per unit volume and p is the polarizability, which describes the structure of molecules or their aggregates) must change under macroscopic shear. Hence, the value of n will also change in the deformation zone.

1063-7842/02/4711-1369$22.00

EXPERIMENT AND RESULTS

(1) In this paper, the dynamo-optical effect in homogeneous Newtonian fluids that have various polarizability of molecules is experimentally studied. Figure 1 shows the experimental setup. Here, *1* is a $6 \times 8 \times 3.5$-cm transparent thermostatic vessel filled with fluid *2* to be studied, and *3* is a vibrator (rod) with fixed plates *4*, which are set in vibration by solenoid *5*. The optical inhomogeneity of the fluid is detected by Schlieren–Aufnahmegerat-80 instrument *6* and CCD-TR780E digital video camera *7*.

The experiments were conducted for the following liquids: L1, transformer oil (permittivity $\varepsilon = 2.3$ and shear viscosity $\mu = 0.16$ Pa s at $T = 17°C$); L2, glycerol ($\varepsilon = 42.4$, $\mu = 1.48$ Pa s, $T = 18°C$); L3, ÉD-20 epoxy resin ($\varepsilon = 3.7$, $\mu = 28$ Pa s, $T = 17°C$); L4, distilled water ($\varepsilon = 81$, $\mu = 10^{-3}$ Pa s, $T = 17°C$); L5, ethanol ($\varepsilon = 25.7$, $\mu = 1.3 \times 10^{-3}$ Pa s, $T = 17°C$); and L6, Freon 10 ($\varepsilon = 2.23$, $\mu = 0.97 \times 10^{-3}$ Pa s, $T = 16°C$).

When vibrator *5* is turned on, a three dimensional flow with a velocity gradient forms near the edges of plates *4* in liquid medium *2* (the frequency ω of the vibrator and the amplitude δ of plate vibrations are varied). If regions with optical anisotropy (waviness) appear in the gradient flow, they are detected by video recorder *7* with the aid of schlieren instrument *6*, which operates by the Toepler scheme.

Figure 2 shows the video records of the process that were obtained under the following conditions: $\omega = 60.5$ Hz, $\delta = 0.1$–0.2 cm, $T = 17°C$, and the vibrator operating time $\widetilde{\Delta t} = 5$ s. In Fig. 2, panels (a–d) refer to L1, L2, L3, and L4, respectively. The first frame in each of the records shows the quiescent state before the vibrator is turned on; the second one corresponds to the instant the vibrator is turned on to excite the three-dimensional flow with the velocity gradient (zero time, $t = 0$); the third frame was taken 3 s after the vibrator had been turned on; and the fourth one shows the state when the vibrator was turned off ($t = \widetilde{\Delta t}$).

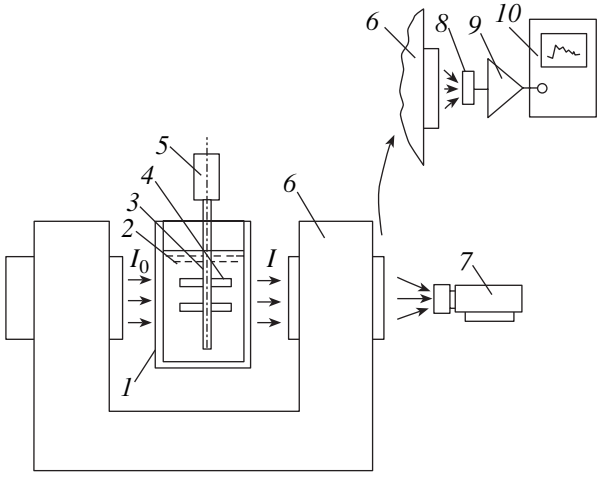


Fig. 1. Experimental set for studying the dynamo-optical effect in a fluid.

From the video records, one can draw the following conclusions. In medium L1, the flow with the gradient velocity field does not break optical anisotropy visually (i.e., does not cause the dynamo-optical effect): internal waviness (the inhomogeneity of the refractive index n and, hence, ε) is not observed (Fig. 2a). In medium L2 (Fig. 2b), regions with a high optical anisotropy appear in the gradient flow; the waviness heals up in 15–16 h at $T = 17°C$ and $\widetilde{\Delta t} = 5$ s. In medium L3, (Fig. 2c), waviness is not observed during the formation of the gradient flow. In medium L4 (Fig. 2d), optically inhomogeneous regions appear. However, as a result of the low viscosity of the liquid, they are swept away by the flow toward the free surface of the liquid (a flow of such a type is caused by the influence of the rigid bottom of vessel *1*). The healing of optical inhomogeneities in liquid L4 occurs 3–4 s after their formation. The dynamo-optical effect is also observed in medium L5 and does not take place in medium L6. All the experiments were performed with a fixed sensitivity of the schlieren instrument to optical inhomogeneities in the liquids.

From these results, it follows that the regions with an observable optical inhomogeneity form in the high-permittivity liquids and their formation does not depend on viscosity. Indeed, in high-viscosity liquid L1, $\varepsilon = 3.2$ and waviness is absent; in high-viscosity liquid L2, $\varepsilon = 42.4$ and waviness is observed; in high-viscosity liquid L3, $\varepsilon = 3.7$ and waviness is absent; in low-viscosity liquid L4, $\varepsilon = 81$ and waviness is observed; in low-viscosity L5, $\varepsilon = 25.7$ and waviness is observed; and in low-viscosity L6, $\varepsilon = 2.3$ and waviness is absent. (It is worth noting that in the low-ε media, optical homogeneity seems to be slightly broken in the gradient flow region; however, this effect is negligible and is beyond the sensitivity of the schlieren instrument.) Since ε is a macroscopic characteristic of the molecular structure of a medium, one can suppose that in regions where the optical inhomogeneity appears under the influence of a shear flow, so does the inhomogeneity of the permittivity, since $\varepsilon = n^2$, and the structure of the medium changes. Thus, quiescent liquids L2, L4, and L5 have an ordered structure, which breaks in the gradient flow field.

From our experiments, it follows that when the operating time of the vibrator in the water increases to $\Delta t = 15$–20 s, the waviness formation in the gradient flow stops. If the vibrator operates for 15–20 s, is turned off, and then is turned on again in Δt^*, repeatedly applying a shear stress, the dynamo-optical effect does not take place if $\Delta t^* < 15$–25 min. Weak waviness starts forming with increasing Δt^*, and only at $\Delta t^* > 2$–3 h (depending on the liquid temperature) is the dynamo-optical effect fully restored: regions with the same degree of optical anisotropy (waviness) as at the initial excitation appear in the gradient flow. Therefore, one

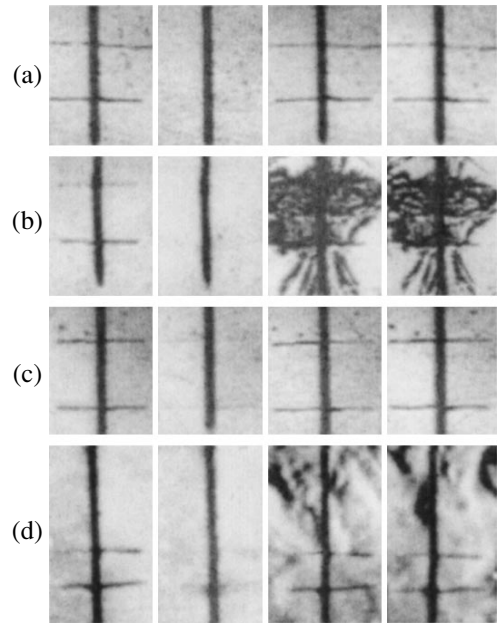


Fig. 2. Video records of the fluid behavior under shear for (a) transformer oil, (b) glycerol, (c) epoxy resin, and (d) water.

can assume that water offers thixotropic properties. This means that it has an ordered structure at rest, which breaks under weak shear stresses and then is recovered after a time (Δt^{**} = 2–3 h).

The time of dynamo-optical effect recovery Δt^{**} in glycerol is several days (Δt^{**} depends essentially on temperature; therefore, a thermostat keeping the liquid temperature constant for several days is required).

As was noted, the time taken to heal up the waviness in glycerol (at T = 15–17°C) is on the order of 15–16 h, whereas for water and ethanol, the time is 4–5 s. Thus, while the waviness intensity depends on the permittivity of a fluid, the recovery time of its structure depends on the viscosity.

(2) In our experiments, we also considered the time variation of the total luminous flux $I(t)$ passed through the water sample under vibration. The basic idea in this case was as follows. In the first approximation, the attenuation of a monochromatic luminous flux in an optically transparent homogeneous liquid obeys the Bouguer–Lambert–Beer law $I = I_0\exp(-\beta_\lambda l)$, where $\beta_\lambda = \alpha_\lambda + \zeta_\lambda$ is the coefficient of luminous flux attenuation from I_0 to I by a liquid layer of depth l; α_λ, coefficient of light absorption by the liquid; ζ_λ, coefficient of light scattering; and λ, wavelength of the monochromatic radiation (monochromatic light with λ = 540 nm was used in the experiments). Since α_λ is very low for transparent liquids, one can assume that $\beta_\lambda \simeq \zeta_\lambda$. It is well known [8] that the mechanism of molecular light scattering by a pure (without impurities) liquid is collective fluctuations of the refractive index n or, in view of the fact that $n = \varepsilon^{1/2}$, collective fluctuations of ε. This agrees with the Einstein formula

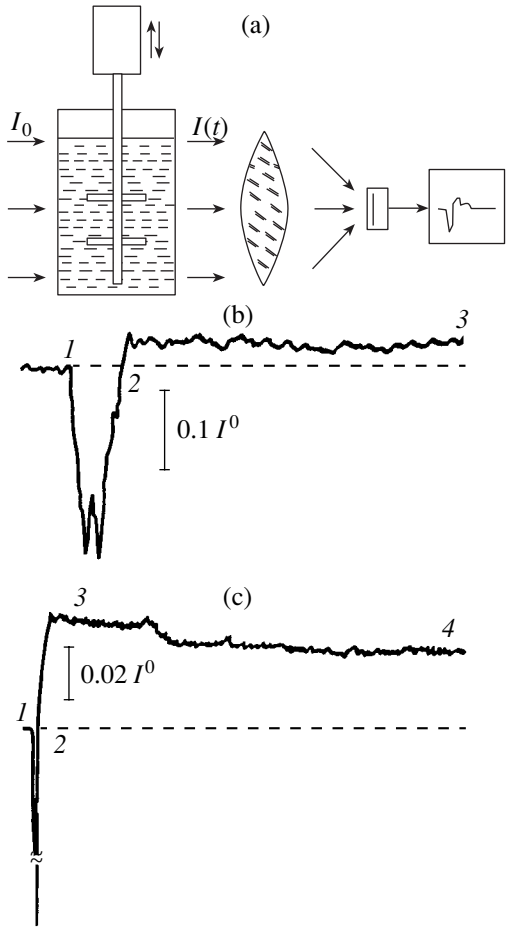


Fig. 3. Intensity variation of the luminous flux at the exit from the water sample under shear. (a) Scheme for detecting the luminous flux $I(t) = I^0 + \Delta I(t)$, (b) oscillogram of the luminous flux $\Delta I(t)$ at the initial stage of the process at a sweep time t_s = 1 min, and (c) oscillogram of $\Delta I(t)$ at t_s = 6.6 min.

$$I = I_0\frac{\pi^2}{2\lambda^4 x^2}V^* V(\overline{\Delta\varepsilon})^2(1+\cos^2\Theta).$$

According to this formula, the intensity of luminous flux scattering at an angle Θ to its direction is proportional to $(\overline{\Delta\varepsilon})^2$. Here, $\Delta\varepsilon$ is the fluctuation of the permittivity in a volume V^* of sizes much smaller than λ that contains many molecules and x is the distance from the scattering volume to the point of observation. Thus, the more homogeneous the structure, the lower the light scattering coefficient in a liquid medium. For example, if molecules in the structure volume were distributed perfectly uniformly and, hence, ε were the same throughout the volume (its fluctuation is absent, $\Delta\varepsilon \longrightarrow 0$), the scattering coefficient would tend to zero. In

other words, the light scattering coefficient of a homogeneous liquid must be lower and the light attenuation must be weaker than in a liquid consisting of polar molecule aggregates or other ordered mesostructures.

In order to measure the luminous flux $I(t)$ passed through the water, photodiode *8* was placed at the exit from schlieren instrument *6* instead of video recorder *7* (Fig. 1). The signal of the photodiode (or the intensity of the luminous flux $I(t)$) was amplified by amplifier *9* and detected by S9-16 oscillograph *10*, operating in the low-frequency regime. Oscillograms in Fig. 3 show the variations of the luminous flux intensity $I(t) = I^0 + \Delta I(t)$ at the exit from the water sample, whose ordered structure breaks under vibration (at $T = 17°C$). Here, I^0 is the initial luminous flux at the exit from the water sample without vibration and $\Delta I(t)$ is the variation of the luminous flux intensity according to the variation of light scattering in the stressed sample. In both oscillograms, the zero level corresponds to the constant luminous flux intensity I^0. The dip between points *1* and *2* corresponds to the formation of waviness in the water sample (due to gradient flows induced by the vibrator), i.e., to more intense light scattering. The portions *2–3* in Fig. 3b and *2–4* in Fig. 3c correspond to an increased transparency of the sample after the breakdown of the ordered mesostructures and, hence, to a decreased light scattering coefficient in the more homogeneous liquid. The value of $\Delta I(t)$ decreases slowly, approaching zero within $\Delta t^{**} \simeq 2.5$ h. It follows that the structural properties of the water are fully recovered within the time Δt^{**}, which must depend on the temperature T, frequency ω, and plate vibration amplitude δ, i.e., on the intensity of the gradient flow induced in the liquid and specifying the breakdown of its structure (finding these dependencies is the subject of a separate study). It was found that the dynamo-optical effect in water is observed at temperatures below 18.5–18.6°C. At $T \geq$ 19°C, the effect is negligibly small; i.e., the water structure is disordered.

CONCLUSION

A series of experiments conducted with the aid of a special technique that differs essentially from that used in [4] made it possible to observe the dynamo-optical effect in homogeneous polar fluids. When at rest, these fluids have a highly ordered structure, which is in agreement with [1–4]. Using water as an example because of its significance in biochemical and biophysical processes, it was shown that after intense breakdown, the liquid structure is recovered within several hours (the relaxation time depends on the liquid temperature).

REFERENCES

1. S. V. Stebnovskii, Zh. Tekh. Fiz. **51**, 2177 (1981) [Sov. Phys. Tech. Phys. **26**, 1280 (1981)].
2. S. V. Stebnovskiĭ, Prikl. Mekh. Tekh. Fiz. **40** (3), 53 (1999).
3. S. V. Stebnovskiĭ, Prikl. Mekh. Tekh. Fiz. **40** (4), 156 (1999).
4. R. A. Apakashev and V. V. Pavlov, Izv. Akad. Nauk, Mekh. Zhidk. Gaza, No. 1, 3 (1997).
5. P. Debye, *Polare Molekeln* (Verlag von S. Hirzel, Leipzig, 1929), p. 123.
6. K. L. Wolf, *Teoretishe Chemie* (J. A. Barth Verlag, Leipzig, 1943), p. 374.
7. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Fizmatgiz, Moscow, 1959; Pergamon, Oxford, 1960).
8. G. S. Landsberg, *Optics* (Nauka, Moscow, 1976), p. 926.

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 11, 2002, pp. 1373–1379. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 28–34.

GASES AND LIQUIDS

Parametric Buildup of the Instability of a Charged Flat Liquid Surface Imposed on Kelvin–Helmholtz Instability

A. I. Grigor'ev, A. S. Golovanov, and S. O. Shiryaeva

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

e-mail: grig@uniyar.ac.ru

Received April 26, 2002

Abstract—The instability of capillary gravitational waves that is developed at the charged flat interface between media is studied for the case when the upper medium moves parallel to the interface with a velocity that has constant and time-dependent components. It is shown that the Mathieu–Hill equation, describing the temporal evolution of the capillary wave amplitudes in such a system, has unstable solutions at those values of physical parameters (electric field strength and wind velocity) meeting the conditions for Saint Elmo fire initiation in the atmosphere. © *2002 MAIK "Nauka/Interperiodica".*

(1) Saint Elmo fire is among the inadequately explored phenomena of atmospheric electricity. The statistical processing of observations regarding this phenomenon under natural conditions [1, 2] suggests that its conventional interpretation as a powerful corona discharge is to some extent incorrect. Almost 80% of the associated works relate Saint Elmo fire with electrostatic instability in the storm electric field of a charged film or water drops covering objects near which the phenomenon is observed. The same has been indicated by various experiments [3–6]. In [1, 7] (see also [2]), a model of Saint Elmo fire was proposed where it is related to the ejection of highly dispersed heavily charged droplets from the liquid surface exposed to an external electric field. The high electric fields of the self-charges of the droplets cause corona discharges in their vicinities and, hence, glow. The characteristic linear size of Saint Elmo fire tongues depends on the distance passed by a droplet until the corona discharge in its vicinity is quenched.

Nevertheless, theoretical estimates show that the water surface subjected to an electric field must become unstable at near-ground electric field strengths one order of magnitude higher than that observed under natural conditions [8]. According to observations [1, 2], the intensity of Saint Elmo fire grows appreciably with increasing wind velocity. This suggests that Kelvin–Helmholtz instability is favorable for loosening the critical conditions for the development of the charged liquid surface instability [9]. However, the reduction of the critical strength of the external electric field obtained in [9] proved to be insufficient. The situation was not improved when a finite thickness of the upper medium was taken into account [10]. The latter factor introduces the linear interaction between the waves and an additional vibrational instability of the interface. It remains to study the effect of the time dependence of the wind velocity, which is typical of stormy weather—a prerequisite for Saint Elmo fire to occur. Such a dependence, according to [11, 12], may promote the parametric instability of a charged liquid surface at external electric fields meeting actual stormy conditions.

(2) We shall solve the stability problem for a tangential discontinuity between two immiscible ideal incompressible liquids with different densities ρ_1 and ρ_2. Either liquid occupies the semi-infinite space in the gravity field, and the upper liquid moves parallel to the interface with a time-dependent velocity $\mathbf{U} = \mathbf{U}(t)$. We assume that the upper liquid (parameters referring to it are marked by subscript 1) is dielectric with a permittivity $\varepsilon_1 = \varepsilon$ and the lower liquid (associated parameters are marked by subscript 2) is perfectly conducting. The upper liquid is exposed to a uniform electrostatic field $\mathbf{E}_0$ directed parallel to the free-fall acceleration $\mathbf{g}$. The electric field $\mathbf{E}_0$ induces a charge at the unperturbed flat interface, which is described by the equation $z = 0$ in the Cartesian coordinate system with unit vectors $\mathbf{n}_z \parallel \mathbf{g}$ and $\mathbf{n}_x \parallel \mathbf{U}$. The charge has a constant density $\sigma = E_0/4\pi$. The thermal motion of molecules induces capillary wave motion of very small amplitude ξ_0 in both media ($\xi_0 \sim (kT/\alpha)^{1/2}$, where k is the Boltzmann constant, T is absolute temperature, and α is the surface tension coefficient at the interface). The capillary motion causes a perturbation $\xi(x, t)$ of the interface; therefore, the equation for the interface takes the form $z = \xi(x, t)$.

Let the velocity field potentials of the upper and lower liquids be $\psi_1(\mathbf{r}, t)$ and $\psi_2(\mathbf{r}, t)$ and the potential of the total electric field, $\phi(\mathbf{r}, t)$. Then, the problem of the temporal evolution of the capillary wave amplitudes is mathematically stated as [13, 14]

$$\Delta\phi = 0; \quad \phi \equiv E_0 z + \varphi(\mathbf{r}, t); \tag{1}$$

$$z \longrightarrow -\infty: \varphi \longrightarrow 0; \tag{2}$$

1063-7842/02/4711-1373$22.00

$$z = \xi(x; t): \varphi = \text{const}; \tag{3}$$

$$\Delta\psi_i = 0; i = 1, 2; \tag{4}$$

$$z \longrightarrow -\infty \quad \psi_1 - xU(t) \longrightarrow \text{const} = 0; \tag{5}$$

$$z \longrightarrow \infty \quad \psi_2 \longrightarrow \text{const} = 0; \tag{6}$$

$$z = \xi(x; t): \frac{\partial\psi_1}{\partial z} \approx U\frac{\partial\xi}{\partial x} + \frac{\partial\xi}{\partial t}; \quad \frac{\partial\psi_2}{\partial z} \approx \frac{\partial\xi}{\partial t}; \tag{7}$$

$$\rho_1\frac{\partial\psi_1}{\partial t} + \rho_1 g\xi + \frac{1}{2}\rho_1[(\nabla\psi_1)^2 - U^2(t)] = \rho_2\frac{\partial\psi_2}{\partial t} + \rho_2 g\xi + P_E - P_\alpha. \tag{8}$$

Here, k is the wave number; $\varphi(\mathbf{r}, t)$ is the potential of the electric field generated by the interface perturbation $\xi(x, t)$ (both parameters are of the same order of smallness); P_E is the electric field pressure on the interface, which is determined by solving electrostatic problem (1)–(3) and has the form $P_E = 4\pi\sigma^2 k\xi$ [13]; $P_\alpha = \alpha\partial^2\xi/\partial x^2$ is the Laplace pressure under the liquid surface perturbed by the wave motion; and Δ is Laplacian.

Our aim is to derive a differential equation describing the temporal variation of the amplitudes of thermal capillary modes (under the action of the electric field pressure and the tangential discontinuity of the velocity field at the interface). We take into account that, in accordance with (5) and (6), the velocity field potentials in both media, being the solutions to Eq. (4), must vanish at a point infinitely distant from the interface. This means [12, 13] that $\psi_1(\mathbf{r}, t) \sim \exp(kz)$ and $\psi_2(\mathbf{r}, t) \sim \exp(-kz)$. This leads us to the relationships

$$\frac{\partial\psi_1}{\partial z} = k\psi_1; \quad \frac{\partial\psi_2}{\partial z} = -k\psi_2. \tag{9}$$

Note that a perturbation of the interface induced by the capillary wave motion must be periodic:

$$\xi \sim \exp(ikx). \tag{10}$$

Substituting (9) and (10) into boundary conditions (6), we find

$$\psi_1(\mathbf{r}, t) = \frac{1}{k}\left(U(t)\frac{\partial\xi}{\partial x} + \frac{\partial\xi}{\partial t}\right) = \frac{1}{k}\left(ikU(t)\xi + \frac{\partial\xi}{\partial t}\right); \tag{11}$$

$$\psi_2(\mathbf{r}, t) = -\frac{1}{k}\frac{\partial\xi}{\partial t}, \tag{12}$$

where i is the imaginary unity.

Now, we substitute (11) and (12) into (8) to obtain the desired differential equation, describing the time variation of the capillary wave amplitudes in the linear approximation with respect to ξ:

$$\frac{\partial^2\xi}{\partial t^2} + 2ikU(t)\frac{\rho_1}{(\rho_1+\rho_2)}\frac{\partial\xi}{\partial t} + ik\rho\frac{\partial U(t)}{\partial t}\xi + \frac{k}{(\rho_1+\rho_2)}[g(\rho_1-\rho_2) - k\rho_1 U^2(t) - 4\pi\sigma^2 k + \alpha k^2]\xi = 0. \tag{13}$$

At $U(t) = U_0 = \text{const}$, Eq. (13) becomes an ordinary differential equation with constant coefficients that describes the time evolution of the capillary wave amplitudes under conditions of Kelvin–Helmholtz and Tonks–Frenkel instabilities and has the solution

$$\xi = \xi_0 \text{Re}\left\{\exp\left[ikU_0 t\frac{\rho_1}{(\rho_1+\rho_2)} + it\omega^2(k)\right]\right\};$$

$$\omega^2(k) \equiv \frac{k}{(\rho_1+\rho_2)}\left[g(\rho_2-\rho_1) + \alpha k^2 - k\frac{\rho_1\rho_2}{(\rho_1+\rho_2)}U_0^2 - 4\pi\sigma^2 k\right]. \tag{14}$$

The conditions for Kelvin–Helmholtz and Tonks–Frenkel instabilities require that the function $\omega^2(k)$ pass through zero toward negative values. In this case, the exponent in (14) has $\pm i$ before $\omega^2(k)$ and (14) is subdivided into two periodic solutions: exponentially damped and exponentially growing. The latter corresponds to Kelvin–Helmholtz instability at $\sigma = 0$. When $U_0 = 0$, (14) yields two aperiodic solutions, one of which exponentially grows and describes the behavior of the liquid surface under Tonks–Frenkel instability. In the general case $U_0 \neq 0$ and $\sigma \neq 0$, (14) describes the interface behavior under conditions of combined instability, when both the electric field pressure and the aerodynamic pressure above the perturbed interface are significant.

(3) Next, let the time dependence of the upper medium velocity $\mathbf{U} = \mathbf{U}(t)$ have the form of a truncated Fourier series:

$$\mathbf{U}(t) = \mathbf{U}_0 + \sum_{j=1}^{m}\mathbf{U}_{*j}\cos(j\omega_0 t); \quad (\mathbf{U}_0 \parallel \mathbf{U}_{*j}) \quad (j = 1, 2, 3, \ldots, m). \tag{15}$$

Substituting (15) into (13) yields a differential equation with variable coefficients for the temporal evolution of the capillary wave amplitudes in the system under consideration. To find its solutions, we change the desired variable as follows:

$$\xi = \zeta\exp(i\Phi(t));$$

$$\Phi = -k\rho\left[U_0 t + \sum_{j=1}^{m}\frac{1}{j\omega_0}U_{*j}\sin(j\omega_0 t)\right];$$

$$\xi = \zeta\exp\left\{ik\rho\left[U_0 t + \sum_{j=1}^{m}\frac{1}{j\omega_0}U_{*j}\sin(j\omega_0 t)\right]\right\}.$$

Ultimately, (8) is reduced to the form

$$\frac{\partial^2\zeta}{\partial t^2} + \zeta\left[\omega^2(k) - 2U_0\chi\sum_{j=1}^{m}U_{*j}\cos(j\omega_0 t) + \chi\sum_{j=1}^{m}U_{*j}^2\cos^2(j\omega_0 t)\right] = 0; \quad (16)$$

$$\chi \equiv k^2\frac{\rho_1\rho_2}{(\rho_1+\rho_2)^2}.$$

(4) For preliminary qualitative analysis, we assume that $m = 1$ and rewrite (16) in dimensionless variables such that the frequency ω_0 of the variable component of the upper medium velocity, the density ρ_2 of the lower medium, and the capillary constant a of the lower liquid are equal to unity ($\alpha = \omega_0 = \rho_2 = 1$). Then, we obtain

$$\frac{\partial^2\zeta}{\partial t^2} + \zeta[\delta - 2\varepsilon\lambda\cos t + 2\varepsilon\cos 2t] = 0;$$

$$\omega_*^2 \equiv \frac{k\alpha}{(1+\rho_1)}\left[(1-\rho_1) + k^2 - k\frac{\rho_1}{(1+\rho_1)}\mathrm{We} - k\omega\right];$$

$$\delta \equiv \omega_*^2 + 2\varepsilon; \quad (17)$$

$$\varepsilon = \frac{1}{4}U_*^2\chi; \quad \lambda \equiv U_0/U_*; \quad \chi \equiv k^2\frac{\rho_1}{(1+\rho_1)^2};$$

$$\mathrm{We} \equiv U_0^2/\alpha, \quad \omega \equiv 4\pi\sigma^2/\alpha.$$

The parameter δ characterizes the frequency of capillary oscillations. All the physical parameters are denoted by the same symbols as before.

With these designations, the domain of coexistence of Kelvin–Helmholtz and Tonks–Frenkel instabilities (without parametric oscillations) is given by the condition $\delta < 2\varepsilon$.

(4a) In the particular case $U_0 = 0$, Eq. (17) transforms into the conventional Mathieu equation. The boundaries of several first instability domains for solutions to this equation are shown in Fig. 1. These boundaries were calculated with the method of extended parameters [15] by expanding into a series in the dimensionless parameter ε, which can be assumed to be small in asymptotic calculations (although in the general case, ε may take arbitrarily large values [16]). It is seen that the apices of the instability zones are the integer values of the parameter $\delta^{1/2}$ on the abscissa axis; that is, $\delta^{1/2} = n$, where $n = 0, 1, 2, \ldots$.

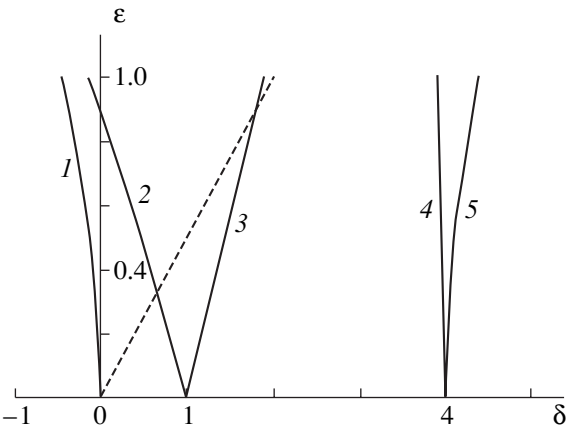


Fig. 1. Boundaries of instability domains in the plane of the parameters δ and ε for the Mathieu equation (Eq. (17) at $\lambda = 0$).

States that are unstable against parametric buildup on the plane (δ; ε) occupy the domain on the left of curve *1*, as well as the domains between curves *2* and *3* and *4* and *5*. The states with the parameters δ and ε between curves *1* and *2*, as well as *3* and *4*, are stable. As follows from direct calculation with Eq. (17) at $\lambda = 0$ (that is, $U_0 = 0$), the states corresponding to δ and ε in the domain between curve *1*, which, in the fourth-order approximation with respect to ε, is given by

$$n = 0: \delta = -\frac{1}{2}\varepsilon^2 + \frac{7}{128}\varepsilon^4 - \ldots,$$

and the dashed line $\delta = 2\varepsilon$, are also stable. On the left of the dashed line, Tonks–Frenkel instability is also expected. Thus, this instability is suppressed by parametric surface oscillations, since Kelvin–Helmholtz instability is impossible at $U_0 = 0$.

The boundaries of the next two instability domains for the Mathieu equation (at $n = 1$ and $n = 2$) are defined by the second-order approximation with respect to ε [15, 17]:

$$n = 1:$$

$$1 - \frac{1}{2}\left(\frac{2\varepsilon}{\delta}\right) + \frac{7}{32}\left(\frac{2\varepsilon}{\delta}\right)^2 \ll \delta \ll 1 - \frac{1}{2}\left(\frac{2\varepsilon}{\delta}\right) + \frac{7}{32}\left(\frac{2\varepsilon}{\delta}\right)^2; \quad (18)$$

$$n = 2:$$

$$4 + \frac{2}{3}\left(\frac{2\varepsilon}{\delta}\right)^2 - \left(\frac{2\varepsilon}{\delta}\right)^2 \ll \delta \ll 4 + \frac{2}{3}\left(\frac{2\varepsilon}{\delta}\right)^2 + \left(\frac{2\varepsilon}{\delta}\right)^2.$$

Here, $\delta^{1/2}$ has the meaning of the capillary wave frequency. If the lower liquid is assumed to be not ideal but low-viscous with the kinematic viscosity coefficient ν and the damping decrement of capillary waves in this liquid given by the well-known expression $\gamma \equiv 2\nu k^2$ [14] (the condition of low viscosity has the form $\gamma \ll 1$), the

inclusion of viscosity in (17) is reduced to adding a term $\gamma(\partial\zeta/\partial t)$ proportional to the first time derivative [17]. This gives rise to a threshold in the variable velocity component amplitude U_* or, in the notation used in (17), in ε, starting from which parametric instability may develop. The height of such a threshold grows with increasing instability domain number n. For domains with $n = 1$ and $n = 2$, the critical values of ε (in the same dimensionless variables) are defined by the relationships [17]

$$\begin{aligned} n &= 1 : \varepsilon/\delta > 2\gamma; \\ n &= 2 : \varepsilon/\delta > (2\gamma)^{1/2}. \end{aligned} \tag{19}$$

Let us estimate the possibility of simultaneously fulfilling conditions (18) and (19) for the water–air system with $\rho_1/\rho_2 \approx 10^{-3}$, $\nu = 10^{-2}$ cm^2/s, and $\alpha = 72$ dyn/cm.

For the widest zone in terms of ε (at $n = 1$), according to (18), the condition $\delta \approx \omega_0^2$ must be fulfilled. In the dimensional form, this condition is written as

$$\omega_*^2 + 2\varepsilon \approx \omega_0^2. \tag{20}$$

Similarly, from (19) in view of (20), it follows in the dimensional form for $n = 1$

$$\varepsilon > 2(\omega_*^2 + 2\varepsilon)\gamma/\omega_0 \cong 2\gamma\omega_0$$

or, taking into account (20) once again,

$$\varepsilon^2 > 4\gamma^2(\omega_*^2 + 2\varepsilon).$$

Considering this relation as an algebraic equation in ε, one can easily find its single positive root:

$$\varepsilon > \gamma[\gamma + (\gamma^2 + \omega_*^2)^{1/2}]. \tag{21}$$

Let us analyze condition (21) keeping in mind that, for the water–air system, the damping decrement for the wave number $k = a^{-1}$ (where a is the capillary constant for water) is defined by the relationship $\gamma \approx 0.2$ s^{-1}. As for the frequency ω_* of capillary oscillations of the liquid, its magnitude at We ≈ 0, according to (17), depends on the surface density σ of the charge induced by the external electric field (that is, on the parameter ω). When ω is close to its critical value in terms of the onset of Tonks–Frenkel instability, ω_*^2 is close to zero. This is because, according to the aforesaid, the critical conditions for Tonks–Frenkel instability are $\omega_*^2 \le 0$. In this case, γ^2 and ω_*^2 in (21) are of the same order of magnitude. Assume for definiteness that $\omega_*^2 = 3\gamma^2$; then, from (21), we find $\varepsilon = 3\gamma^2$. Going from ε to the velocity U_*, according to (17), we can rewrite the above condition as $U_* \approx 7$ cm/s. Thus, the parametric instability of a charged liquid surface can be caused, in view of (20), by the wind component varying with the frequency ω_0, which is found from the relationship $\omega_0^2 = 3\gamma^2 + 6\gamma^2$ or $\omega_0 = 3\gamma \approx 0.6$ s^{-1}. This value is close to the wind pulsation frequency under developed convection [18]. The pulsation amplitude may be relatively small in this case: $U_* \approx 10$ cm/s.

Such an evaluation is valid for low frequencies of capillary waves, i.e., only at very high external electrostatic fields approaching the limiting value for Tonks–Frenkel instability initiation (several tens of kilovolts per centimeter). So, high fields are not observed near the earth surface in stormy weather. Even inside thunderous clouds such fields are very seldom encountered [8]. In other words, Saint Elmo fire could hardly arise in this way.

Consider another limiting situation when the electric field strength near the earth surface is weak. Let, for definiteness, it be one order of magnitude lower than that necessary for the initiation of Tonks–Frenkel instability. Then, the frequency ω_* of gravitational capillary waves on the liquid surface will depend on the electric field strength only slightly and will be governed by the free-fall acceleration and the surface tension coefficient. At $k = a^{-1}$, the capillary wave frequency is on the order of hundreds of Hertz ($\omega_* \approx 70$ s^{-1}). This means that the inequality $\omega_*^2 \gg \gamma^2$ is fulfilled *a fortiori*. Then, it is easy to see that $\varepsilon \approx \gamma\omega_*$ from (21). In the dimensional variables, this condition yields the critical (in terms of stability loss) value of the variable velocity component amplitude for the upper medium, $U_* \approx 1$ m/s, at the pulsation frequency $\omega_0 \approx 70$ s^{-1}. Such a wind component may be associated with small-scale atmosphere turbulization caused, for instance, by rain.

It should be pointed out that the instability under discussion may be realized in different ways, depending on specific values of the parameters δ and ε. It is seen in Fig. 2, where the results of direct calculation with the use of (17) at different points of the plane (δ, ε) are presented.

(4b) Now let the constant velocity component U_0 be other than zero [that is, $\lambda \neq 0$ in (17)]. Assuming that the dimensionless parameter ε is small and using the method of extended parameters (for details, see [15]), one can find the boundaries of several first instability domains for solutions to complete equation (17). The boundaries for the first three instability domains are given by

$$Ce_0 = \left(-\frac{1}{2} - 2\lambda^2\right)\varepsilon^2 - 3\lambda^2\varepsilon^3 + \left(\frac{7}{128} - \frac{20\lambda^2}{9} + \frac{7\lambda^2}{2}\right)\varepsilon^4;$$

$$Ce_1 = 1 - \varepsilon + \left(-\frac{1}{8} + \frac{5\lambda^2}{3}\right)\varepsilon^2 + \left(\frac{1}{64} + \frac{121\lambda^2}{36}\right)\varepsilon^3 + \left(-\frac{1}{1536} + \frac{6233\lambda^2}{2880} - \frac{763\lambda^4}{216}\right)\varepsilon^4;$$

$$Se_1 = 1 + \varepsilon + \left(-\frac{1}{8} - \frac{\lambda^2}{3}\right)\varepsilon^2 + \left(-\frac{1}{64} - \frac{7\lambda^2}{36}\right)\varepsilon^3$$
$$+ \left(-\frac{1}{1536} - \frac{217\lambda^2}{2880} + \frac{5\lambda^4}{216}\right)\varepsilon^4; \qquad (22)$$
$$Ce_2 = 4 + \left(\frac{5}{12} + \frac{2\lambda^2}{15}\right)\varepsilon^2 - \frac{31}{90}\lambda^2\varepsilon^3$$
$$+ \left(-\frac{763}{13\,824} + \frac{58\lambda^2}{1575} + \frac{433\lambda^4}{13\,500}\right)\varepsilon^4;$$
$$Se_2 = 4 + \left(-\frac{1}{12} + \frac{2\lambda^2}{15}\right)\varepsilon^2 + \frac{19}{90}\lambda^2\varepsilon^3$$
$$+ \left(\frac{5}{13\,824} + \frac{142\lambda^2}{1575} - \frac{317\lambda^4}{13\,500}\right)\varepsilon^4.$$

Here, the function Ce_n defines the right-hand boundary of the instability domain with a number n and the function Se_n, the left-hand one. Constraints on the applicability of relationships (22) are imposed by the condition $\lambda\varepsilon \ll 1$ according to the theory of approximation.

The shapes of the instability domains calculated from relationships (22) at various values of the parameter λ are given in Fig. 3. The presence of two cosines in (17) deforms the boundaries of domains where the solutions are unstable up to their intersection. The deformation of the domains depends on the parameter λ. Direct calculations with the use of Eq. (17) show that its solutions become stable when they correspond to δ and ε located above the point where the boundaries of one instability domain intersect. The domains of the parameters δ and ε corresponding to unstable solutions become closed (certainly except for the zeroth domain bounded on one side). In other words, the effect of parametric instability becomes stable because of the interaction between the coefficients of Eq. (17), which periodically vary with time. Figure 3 shows that, as the parameter λ (the constant velocity component U_0 of the upper medium) increases, the domains of instability parametric stabilization shrink. Thus, the presence of the constant and variable components of the upper medium velocity destabilizes the interface, extending the domain of the parameters δ and ε where the interface is unstable.

It is worth noting that, as U_0 increases, not only the left- and right-hand boundaries of the parametric stability domains but also the boundaries of adjacent domains intersect, causing an extension of the domain of the parameters δ and ε where the interface is unstable (Figs. 3b, 3c). This is especially true for the boundary of the zeroth domain, which deflects to the right and intersects both the first and second domains. Unstable solutions take place for δ and ε lying on the left of and above the boundary of the zeroth zone in Figs. 3b and 3c (naturally, within the applicability of the calculations, that is, at $\varepsilon\lambda < 1$).

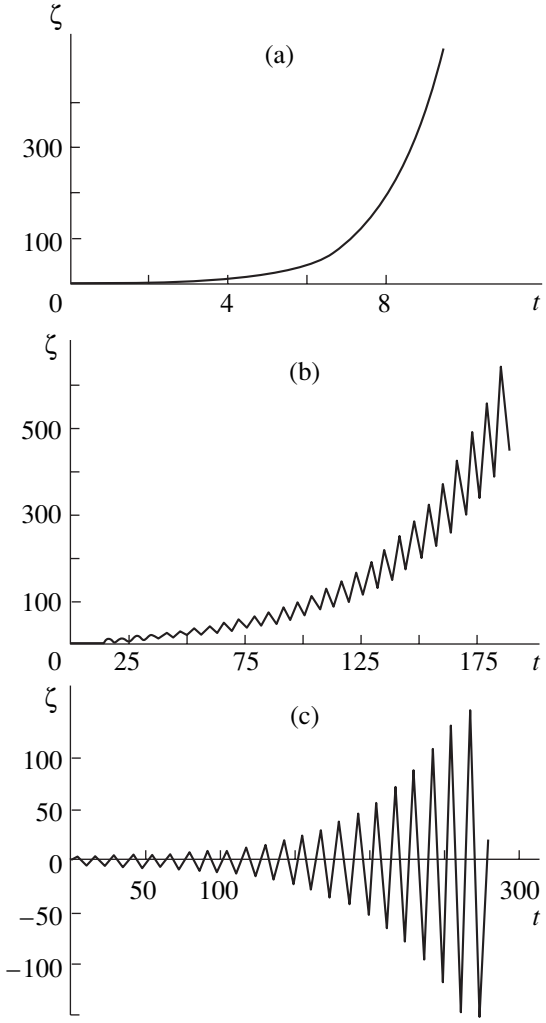


Fig. 2. Possible temporal variation of unstable solutions to Eq. (17) at $\lambda = 1$: (a) $\delta = -0.5$ and $\varepsilon = 0.1$, (b) $\delta = -0.03$ and $\varepsilon = 0.1$, and (c) $\delta = 1$ and $\varepsilon = 0.2$.

In addition, the deflection of the zeroth domain boundary to the right sets the upper limit on ε for the domains of existence of parametric stabilization at $\delta > 0$, since the calculations carried out are valid only at $\varepsilon\lambda < 1$ (for clarity, the boundary curves in the figures are plotted in a wider range of ε). For example, it is seen that, at $\lambda \geq 7$, the condition $\varepsilon\lambda < 1$ practically eliminates the stabilization domain.

(5) Now, assume that $m = 2$ in (16) to see how the boundaries of the instability domains for a solution will change. Let, for simplicity, $U_{*1} = U_{*2} = U_*$. Then, Eq. (16) in terms of the dimensionless variables can be rewritten in the form

$$\frac{\partial^2\zeta}{\partial t^2} + \zeta[\delta - 2\varepsilon\lambda(\cos t + \cos 2t) + 0.5\varepsilon(\cos 2t + \cos 4t)] = 0, \qquad (23)$$

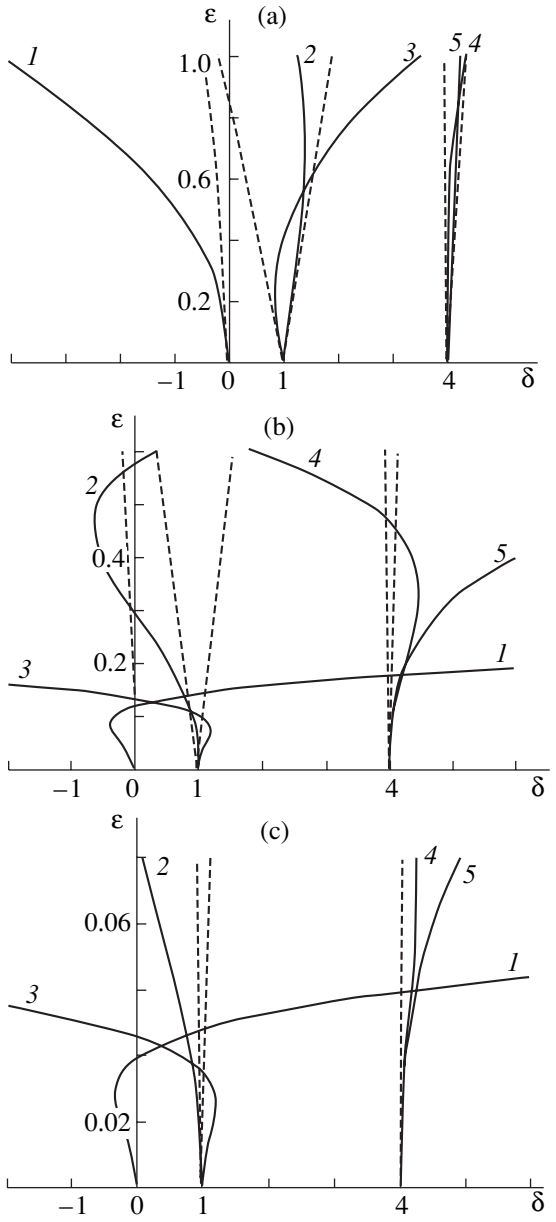


Fig. 3. The same as in Fig. 1 for Eq. (17) at λ = 1 (a), 7 (b), and 20 (c).

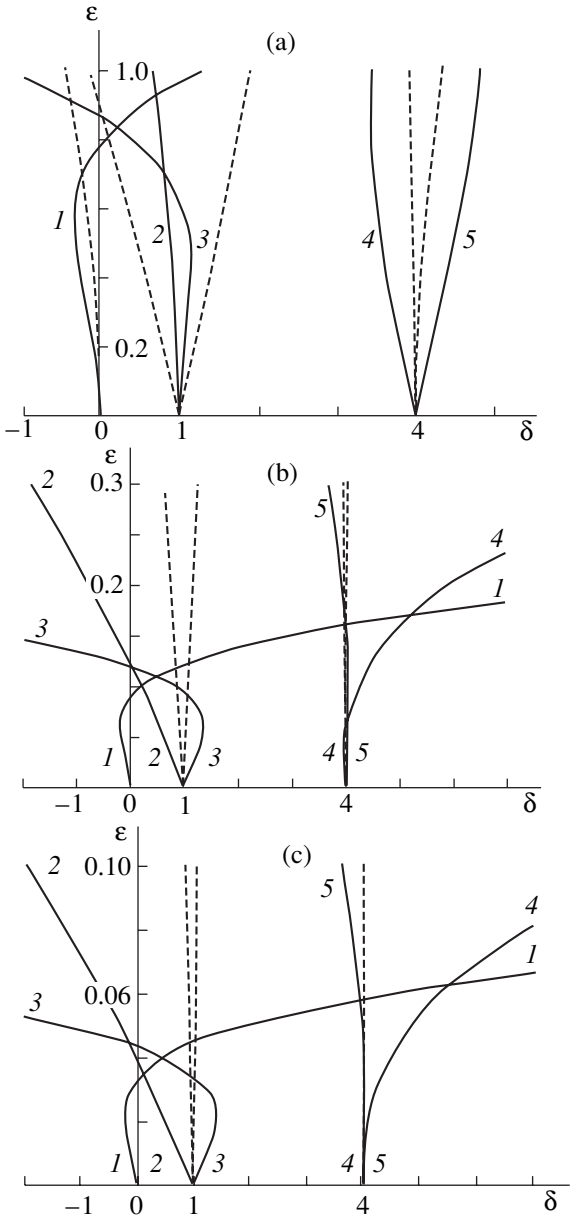


Fig. 4. The same as in Fig. 1 for Eq. (19) at $\lambda_1 = \lambda_2 = 1$ (a), 7 (b), and 20 (c).

where the parameter δ and the boundary of the domain where Tonks–Frenkel and Kelvin–Helmholtz instabilities take place simultaneously change slightly: $\delta \equiv \omega_*^2 + \varepsilon$.

Figure 4 displays the instability domain boundaries calculated with (19) at various λ. Comparing with Fig. 2 shows that, if the number of harmonics taken into account in (15) grows, the deformation of the boundaries of the parametric instability domains increases and the locus of points (in the plane δ, ε) where the instability is observed expands. This is also confirmed in Fig. 5, where the instability domain boundaries calculated for the first four harmonics ($m = 4$) in (15) are shown. In Figs. 3–5, the deformation of the boundaries of the parametric instability domains increases with the number of harmonics in relationship (15). The dimensions of the parametric instability domains grow with increasing ratio of the constant velocity component of the upper medium to the variable one.

(6) The above consideration has been concerned with the initiation of Saint Elmo fire. Nevertheless, our reasoning is applicable to another inadequately explored natural phenomenon, namely, the emergence of waves on the surface of a basin under the influence of wind. Conventionally, the phenomenon of wave buildup by a wind is associated with the development of

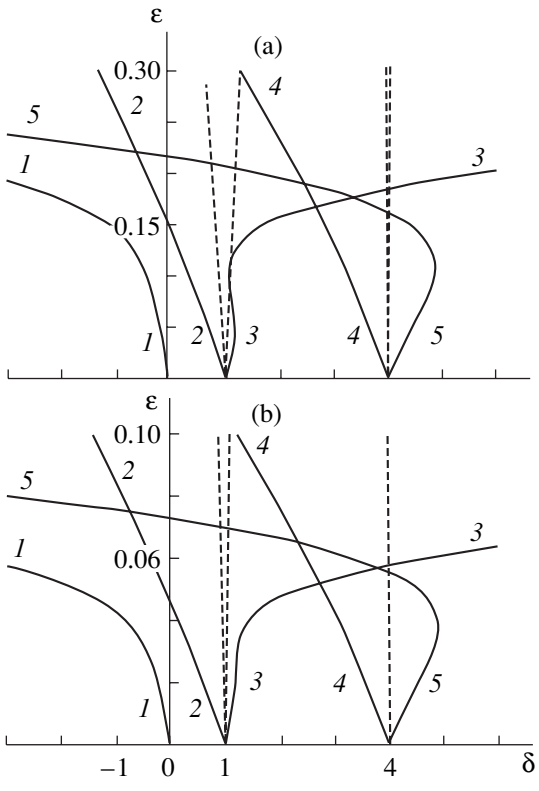


Fig. 5. The same as in Fig. 1 for Eq. (16) at $\lambda_1 = \lambda_2 =$ 7 (a) and 20 (b).

Kelvin–Helmholtz and Miles instabilities [19, 20]. Both instabilities are related to the uniform motion of the upper medium parallel to the interface in a two-layer system of immiscible liquids in the gravity field. The influence of the wind variable component and the possibility of the parametric buildup of waves are not taken into account, although, according to our calculations, such an influence may be appreciable.

CONCLUSION

The Mathieu–Hill equation, describing the temporal evolution of capillary waves at the charged interface between two media when the upper one moves relative to the denser lower medium with a time-dependent velocity, has unstable solutions due to Tonks–Frenkel and Kelvin–Helmholtz instabilities, as well as due to parametric instability typical of the Mathieu–Hill equation. The motion of the upper medium parallel to the interface with a time-dependent velocity leads to the destabilization of the liquid surface and reduces the critical strength of an external electrostatic field down to values observed near the Earth's surface during thunderstorms, when Saint Elmo fire is initiated.

ACKNOWLEDGMENTS

This work was supported by grant no. 00-15-9925 of the President of the Russian Federation.

REFERENCES

1. A. I. Grigor'ev, I. D. Grigor'eva, and S. O. Shiryaeva, J. Sci. Exploration **5** (2), 1 (1991).
2. A. I. Grigor'ev, Sorovs. Obrazov. Zh. **6** (6), 37 (2000).
3. B. V. Voĭtsekhovskiĭ and B. B. Voĭtsekhovskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **23**, 37 (1976) [JETP Lett. **23**, 32 (1976)].
4. B. B. Voĭtsekhovskiĭ, Dokl. Akad. Nauk SSSR **262**, 84 (1982) [Sov. Phys. Dokl. **27**, 54 (1982)].
5. A. I. Grigor'ev and O. A. Sinkevich, Pis'ma Zh. Tekh. Fiz. **11**, 182 (1985) [Sov. Tech. Phys. Lett. **11**, 74 (1985)].
6. A. I. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **11**, 1004 (1985) [Sov. Tech. Phys. Lett. **11**, 415 (1985)].
7. A. I. Grigor'ev and O. A. Sinkevich, Zh. Tekh. Fiz. **54**, 1276 (1984) [Sov. Phys. Tech. Phys. **29**, 735 (1984)].
8. I. P. Mazina, A. Kh. Khrgian, and I. M. Imyanitov, *Clouds and Cloudy Atmosphere: A Handbook* (Gidrometeoizdat, Leningrad, 1989).
9. O. A. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **66** (2), 23 (1996) [Tech. Phys. **41**, 124 (1996)].
10. S. O. Shiryaeva, Zh. Tekh. Fiz. **71** (3), 9 (2001) [Tech. Phys. **46**, 280 (2001)].
11. A. I. Grigor'ev, Zh. Tekh. Fiz. **70** (1), 24 (2000) [Tech. Phys. **45**, 22 (2000)].
12. A. I. Grigor'ev and A. S. Golovanov, Pis'ma Zh. Tekh. Fiz. **25** (20), 13 (1999) [Tech. Phys. Lett. **25**, 806 (1999)].
13. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
14. V. G. Levich, *Physicochemical Hydrodynamics* (Fizmatgiz, Moscow, 1959).
15. A. H. Nayfeh, *Introduction to Perturbation Techniques* (Wiley, New York, 1981; Mir, Moscow, 1984).
16. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
17. N. N. Bogolyubov and Yu. A. Mitropol'skii, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Nauka, Moscow, 1974; Gordon and Breach, New York, 1962).
18. P. N. Tverskoĭ, *Course of Meteorology* (Gidrometeoizdat, Leningrad, 1962).
19. J. W. Miles, J. Fluid Mech. **256**, 427 (1993).
20. E. A. Kuznetsov and P. M. Lushnikov, Zh. Éksp. Teor. Fiz. **108**, 614 (1995) [JETP **81**, 332 (1995)].

Translated by N. Mende

Technical Physics, Vol. 47, No. 11, 2002, pp. 1380–1388. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 35–43.

GASES AND LIQUIDS

On the Thermophoretic Motion of a Heated Spherical Drop in a Viscous Liquid

N. V. Malaĭ

Belgorod State University, Belgorod, 308015 Russia

e-mail: malay@bsu.edu.ru

Received October 12, 2001; in final form, April 22, 2002

Abstract—Expressions for the force and velocity of the thermophoretic motion of a spherical drop in a viscous liquid are derived for arbitrary temperature differences between the surface of the drop and regions away from it. The temperature dependence of the viscosity is taken into account in the form of an exponential–power series. © *2002 MAIK "Nauka/Interperiodica".*

STATEMENT OF THE PROBLEM

The thermophoretic motion of a drop arises when an immobile liquid is subjected to an external temperature gradient. Under the action of the thermocapillary and viscous forces, the drop acquires a constant velocity, the so-called rate of thermophoresis. In general, this motion is related to shear forces appearing on the surface of the drop because of the temperature variation of the surface tension coefficient σ (the Marangoni effect [1, 2]) and the thermal creep of the environment over the surface. In [3], the thermal creep coefficient K_{tc}^{e} was estimated for a number of liquids. For example, for mercury drops moving in water and glycerol, it was found to be $K_{tc}^{e} = 0.13$ and 2.5×10^{-5}, respectively. Hereafter, the indices *e* and *i* refer to a viscous liquid and drop, and respectively; the subscript ∞, to liquid parameters away from the drop; and the subscript *s*, to physical quantities taken at the mean temperature of the surface of the drop T_s. Note that the thermal creep coefficient has been so far evaluated with a reasonable accuracy only for gases [4]. This is because a rigorous mathematical theory of inhomogeneous liquids is absent.

Consider a heated drop of a viscous incompressible liquid steadily moving in another viscous incompressible liquid occupying the entire space. The liquids are immiscible with each other. By a heated drop, we mean a drop whose mean surface temperature far exceeds the ambient temperature. The drop may be heated, for example, by a chemical reaction proceeding in its bulk, radioactive decay of the drop material, absorption of electromagnetic radiation, etc. The heated surface of the drop has a significant effect on the thermal physical properties of the environment and thereby affects greatly the velocity and pressure fields in its vicinity. At infinity, the liquid is at rest and is subjected to a given temperature gradient. It is assumed that the densities, thermal conductivities, and specific heats inside and outside the liquids are constant; the viscosity coefficient of the drop exceeds that of the environmental liquid; and the surface tension coefficient is an arbitrary function of temperature. Also, the drop is assumed to move slowly (small Peclet and Reynolds numbers) and retain its spherical shape (the distortion of sphericity will be discussed later).

Unlike the case studied previously [1–8], we here consider the thermophoretic motion of a spherical drop for an arbitrary temperature difference between the surface of the drop and areas away from it. In the equation of hydrodynamics, the temperature dependence of the viscosity has the form of an exponential–power series; in the equation of heat conduction, only convective terms are taken into account.

For the hydrodynamic problem [9, 10] and for the thermal problem [11], it was shown that the inertial and convective terms away from the sphere become comparable to those responsible for molecular transfer by one order of magnitude. Therefore, the standard method of expansion in a small parameter introduces a noticeable error in this case, because it cannot satisfy rigorously boundary conditions at infinity and find an exact unified solution equally valid in the entire flow domain.

It was demonstrated [12] that the heating of the surface of the drop and taking into account the motion of the liquid significantly affect the drag force of the medium. In this work, we study the effect of the motion of the medium on the force and rate of thermophoresis for a heated drop with an arbitrary temperature difference in its vicinity in the presence of a given external temperature gradient.

Of all transport parameters of a liquid, only the viscosity coefficient strongly depends on temperature [13]. To include this dependence, we take advantage of the formula

$$\mu_e = \mu_\infty \left[1 + \sum_{n=1}^{\infty} F_n \left(\frac{T_e}{T_\infty} - 1\right)^n\right] \exp\left\{-A\left(\frac{T_e}{T_\infty} - 1\right)\right\}, \quad (1)$$

1063-7842/02/4711-1380$22.00

which describes the variation of the viscosity over a wide temperature range with any desired accuracy. (At $F_n = 0$, this formula is reduced to the well-known Reynolds relationship [13].)

It is known that the viscosity of a liquid decreases with increasing temperature by the exponential law [13]. Analysis of the available semiempirical formulas showed that expression (1) describes the viscosity variation over a wide temperature range most adequately with any desired accuracy. With the coefficients F_n not taken into consideration, the error can be as high as 40%. For illustration, Tables 1 and 2 list the values of F_n for glycerol and water. The relative error does not exceed 3%. The coefficients F_n were calculated with the Maple V software suite.

Let us place the origin of a fixed coordinate system at the instantaneous center of a spherical drop of radius R. We assume that the drop moves with a constant velocity $\mathbf{U}$ in the negative OZ direction. The velocity and pressure distributions must be symmetric about the line passing through the center of the drop parallel to the velocity vector $\mathbf{U}$. In terms of our assumptions, the equations and boundary conditions for the velocity and temperature in the spherical coordinate system are written as [14, 15]

$$\nabla P_e = \mu_e \Delta \mathbf{U}_e + 2(\nabla \mu_e \nabla)\mathbf{U}_e + [\nabla \mu_e \mathrm{curl}\mathbf{U}_e], \quad \mathrm{div}\mathbf{U}_e = 0, \tag{2}$$

$$\mu_i \Delta \mathbf{U}_i = \nabla P_i, \quad \mathrm{div}\mathbf{U}_i = 0, \tag{3}$$

$$\rho_e c_e (\mathbf{U}_e \nabla) T_e = \lambda_e \Delta T_e, \tag{4}$$

$$\rho_i c_i (\mathbf{U}_i \nabla) T_i = \lambda_i \Delta T_i + q_i, \tag{5}$$

$$r = R, \quad T_e = T_i, \quad \lambda_e \frac{\partial T_e}{\partial r} = \lambda_i \frac{\partial T_i}{\partial r},$$

$$U_r^e = U_r^i = -U\cos\Theta, \tag{6}$$

$$U_\Theta^e - U_\Theta^i = K_{\mathrm{tc}}^e \frac{\nu_e}{RT_e}\frac{\partial T_e}{\partial \Theta} - K_{\mathrm{tc}}^i \frac{\nu_i}{RT_i}\frac{\partial T_i}{\partial \Theta} \quad (\nu = \mu/\rho),$$

$$\mu_e\left[r\frac{\partial}{\partial r}\left(\frac{U_\Theta^e}{r}\right) + \frac{1}{r}\frac{\partial U_r^e}{\partial \Theta}\right] - \frac{1}{r}\frac{\partial \sigma}{\partial T_i}\frac{\partial T_i}{\partial \Theta}$$

$$= \mu_i\left[r\frac{\partial}{\partial r}\left(\frac{U_\Theta^i}{r}\right) + \frac{1}{r}\frac{\partial U_r^i}{\partial \Theta}\right], \tag{7}$$

$$r \longrightarrow \infty, \quad \mathbf{U}_e \longrightarrow 0, \quad P_e \longrightarrow P_\infty,$$

$$T_e \longrightarrow T_\infty + |\nabla T| r\cos\Theta,$$

$$r \longrightarrow 0, \quad |\mathbf{U}_i| \neq \infty, \quad P_i \neq \infty, \quad T_i \neq \infty. \tag{8}$$

Here, $q_i(r, \Theta)$ is the density of heat sources nonuniformly distributed in the drop. Specifically, if the drop is heated by absorbing electromagnetic radiation, the nonuniformity depends on the optical constants of the drop ($m_i = n_i + ia_i$, where m_i, n_i, and a_i are the complex refractive index, refractive index, and the absorption coefficient of the drop, respectively) and its diffraction parameter $x_d = 2\pi R/\lambda$ (λ is the wavelength). The expression for the radiation energy density converted to heat in the drop can be represented in the form [16]

$$q_i = \frac{4\pi R n_i a_i}{n_e} I B_i,$$

where I is the incident radiation intensity and B_k ($k = e, i$) is a coordinate-dependent function calculated by the Mie theory.

The results of numerical calculation for the B_k distribution in the case of water drops are given in [16]. They show that the nonuniformity of the absorbed energy distribution in a water drop increases with its radius. The nonuniformity is the greatest in the direction of radiation propagation.

Table 1

Glycerol: $A = 17.29$, $F_1 = -1.228$, $F_2 = 7.022$, $T_\infty = 303$ K

T, °C	μ_{cal}, Pa s	μ_{exp}, Pa s	$\frac{\lvert\mu_{\mathrm{cal}} - \mu_{\mathrm{exp}}\rvert}{\mu_{\mathrm{exp}}}100\%$
30	0.600000	0.600	0
40	0.327979	0.330	0.61
50	0.182001	0.180	1.11
60	0.102619	0.102	0.60
70	0.058797	0.059	0.34
80	0.034212	0.035	2.25
90	0.020189	0.021	3.86

Note: μ_{cal}, dynamic viscosity calculated by formula (1); μ_{exp}, experimental value.

Table 2

Water: $A = 5.779$, $F_1 = -2.318$, $F_2 = 9.118$, $T_\infty = 273$ K

T, °C	μ_{cal}, Pa s	μ_{exp}, Pa s	$\frac{\lvert\mu_{\mathrm{cal}} - \mu_{\mathrm{exp}}\rvert}{\mu_{\mathrm{exp}}}100\%$
0	0.0017525	0.0017525	0
10	0.0013151	0.0012992	1.22
20	0.0010089	0.0010015	0.74
30	0.0007943	0.0007971	0.35
40	0.0006433	0.0006513	1.22
50	0.0005359	0.0005441	1.51
60	0.0004581	0.0004630	1.06
70	0.0004002	0.0004005	0.07
80	0.0003556	0.0003509	1.35
90	0.0003199	0.0003113	2.76

Boundary conditions (6) on the surface of the drop ($r = R$) include the impermeability condition for the normal components of velocity, equality of the temperatures, heat flux continuity, equality of the shear velocities at the inner and outer surfaces of the drop, and continuity of the shear components of the stress tensor.

Away from the drop ($r \longrightarrow \infty$), boundary conditions (7) are valid, while the finiteness of the physical quantities characterizing the drop at $r \longrightarrow 0$ is included in (8).

The decisive parameters of the problem are the material constants ρ_e, μ_∞, λ_e, and c_e, as well as R, $|\nabla T|$, T_∞, and U, which remain constant during the motion of the spherical drop. With these parameters, one can compose a three-dimensional combination including $\varepsilon = R|\nabla T|/T_\infty \ll 1$ (this parameter characterizes a temperature difference within the drop), as well as the Peclet and Reynolds numbers.

We will make Eqs. (2)–(5) and boundary conditions (6)–(8) dimensionless by introducing the dimensionless velocity, temperature, and pressure: $\mathbf{V}_k = \mathbf{U}_k/U$, $t_k = T_k/T_\infty$, and $p_k = P_k/P_\infty$ ($P_\infty = (\mu_\infty U)/R$). Here, the radius of the drop R, temperature T_∞, pressure P_∞, and velocity U are used as units of measure of distance, temperature, pressure, and velocity, respectively ($U \sim |\mu_\infty|\nabla T\|/(\rho_e T_\infty)$).

At $\varepsilon \ll 1$, a solution to the equations of hydrodynamics should be sought in the form

$$\begin{aligned} \mathbf{V} &= \mathbf{V}^{(0)} + \varepsilon\mathbf{V}^{(1)} + \ldots, \\ p &= p^{(0)} + \varepsilon p^{(1)} + \ldots, \quad t = t^{(0)} + \varepsilon t^{(1)} + \ldots. \end{aligned} \tag{9}$$

The form of boundary conditions (6)–(8) indicates that an expression for the velocity components V_r and V_Θ is sought as the expansion in Legendre and Gegenbauer polynomials. The force acting on the drop is found by integrating the stress tensor over its surface. Taking into account the properties of Legendre and Gegenbauer polynomials, we can assume that this force depends largely on the first terms of the expansions; therefore,

$$V_r = G(y)\cos\Theta, \quad V_\Theta = -g(y)\sin\Theta, \tag{10}$$

where $G(y)$ and $g(y)$ are arbitrary functions depending on the dimensionless radial coordinate $y = r/R$.

When studying the motion of nonuniformly heated drops under an external temperature gradient in a viscous medium, one should take into account the temperature dependences of both the dynamic viscosity coefficient and the surface tension coefficient. This is because the density q_i of heat sources in the drop is nonuniform. In this work, the surface tension coefficient is taken to be an arbitrary function of temperature. Moreover, for the first time, an attempt is made to take into consideration the effect of the motion of the medium on the force and velocity of the thermocapillary drift of a heated drop in a viscous liquid. Therefore, the final formulas are of general character and are valid for any temperature difference between the surface of the drop and regions away from it.

TEMPERATURE FIELDS INSIDE AND OUTSIDE A HEATED DROP

When finding the force acting on a nonuniformly heated drop and its velocity, we will consider only first-order corrections. To find them, it is necessary to know temperature fields inside and outside the drop, i.e., to solve Eqs. (4) and (5) with appropriate boundary conditions. The dimensionless equations of heat conduction have the form

$$\varepsilon \mathrm{Pr}_\infty(\mathbf{V}_e\nabla)t_e = \Delta t_e, \tag{11}$$

$$\varepsilon \mathrm{Pr}_\infty\beta(\mathbf{V}_i\nabla)t_i + Q = \Delta t_i, \tag{12}$$

$$\begin{gathered} y = 1, \quad t_i = t_e, \quad \lambda_e\frac{\partial t_e}{\partial y} = \lambda_i\frac{\partial t_i}{\partial y}, \\ y \longrightarrow \infty, \quad t_e \longrightarrow 1 + \varepsilon y\cos\Theta, \\ y \longrightarrow 0, \quad t_i < \infty. \end{gathered} \tag{13}$$

Here, $\mathrm{Pr}_\infty = \mu_\infty c_e/\lambda_e$ is the Prandtl number, $\beta = \chi_e/\chi_i$, χ is the thermal diffusivity, and $Q = -q_i R^2/(\lambda_i T_\infty)$. Substituting (9) into Eqs. (11) and (12), we arrive at the following set of equations:

zeroth-order approximation ($\varepsilon = 0$),

$$\Delta t_e^{(0)} = 0, \tag{14}$$

$$\Delta t_i^{(0)} = Q_0, \tag{15}$$

$$\begin{gathered} y = 1, \quad t_e^{(0)} = t_i^{(0)}, \quad \lambda_e\frac{\partial t_e^{(0)}}{\partial y} = \lambda_i\frac{\partial t_i^{(0)}}{\partial y}, \\ y \longrightarrow \infty, \quad t_e^{(0)} \longrightarrow 1, \\ y \longrightarrow 0, \quad t_i^{(0)} < \infty; \end{gathered}$$

first-order approximation ($\sim\varepsilon$),

$$\mathrm{Pr}_\infty\left(V_r^e\frac{\partial t_e^{(0)}}{\partial y} + \frac{V_\Theta^e}{y}\frac{\partial t_e^{(0)}}{\partial\Theta}\right) = \Delta t_e^{(1)}, \tag{16}$$

$$\mathrm{Pr}_\infty\beta\left(V_r^e\frac{\partial t_e^{(0)}}{\partial y} + \frac{V_\Theta^e}{y}\frac{\partial t_e^{(0)}}{\partial\Theta}\right) + Q_1 = \Delta t_i^{(1)}, \tag{17}$$

$$\begin{gathered} y = 1, \quad t_i^{(1)} = t_e^{(1)}, \quad \lambda_i\frac{\partial t_i^{(1)}}{\partial y} = \lambda_e\frac{\partial t_e^{(1)}}{\partial y}, \\ y \longrightarrow \infty, \quad t_e^{(1)} \longrightarrow y\cos\Theta, \\ y \longrightarrow 0, \quad t_i^{(1)} < \infty. \end{gathered}$$

When deriving the equation for the temperature distribution inside the drop, we assumed that

$$-\sum_{n=0}^{\infty} q_i(r,\Theta)\frac{R^2}{\lambda_i T_\infty} = \sum_{n=0}^{\infty} \varepsilon^n Q_n P_n(x),$$

where

$$Q_n = -\frac{R^2}{T_\infty}\frac{2n+1}{2\lambda_i}\int_{-1}^{+1} q_i(r,\Theta)P_n(x)dx, \quad x = \cos\Theta.$$

Let us find the zeroth-order approximations. The general solutions to (14) and (15) have the form

$$t_e^{(0)}(y) = 1 + \frac{\gamma}{y} + \sum_{n=1}^{\infty}\frac{\Gamma_n}{y^{n+1}}P_n(x), \tag{18}$$

$$t_i^{(0)}(y) = B_0 + \frac{1}{4\pi R T_\infty \lambda_i y}\int_V q_i dV + \int_1^y \frac{\psi_0}{y}dy - \frac{1}{y}\int_1^y \psi_0 dy, \tag{19}$$

where

$$\psi_0 = -\frac{R^2}{2\lambda_i T_\infty}y^2\int_{-1}^{+1} q_i(r,\Theta)dx.$$

In (19), integration is over the entire volume of the drop. The constants of integration γ, Γ_n, and B_0 are determined by substituting (18) and (19) into the related boundary condition. After the substitution, we find that

$$\gamma = t_s - 1, \quad B_0 = 1 + \left(1 - \frac{\lambda_e}{\lambda_i}\right)\gamma,$$

$\Gamma_n = 0$ at $n \geq 1$, $t_s = T_s/T_\infty$, and T_s is the mean temperature on the surface of the drop given by

$$\frac{T_s}{T_\infty} = 1 + \frac{1}{4\pi R\lambda_e T_\infty}\int_V q_i dV. \tag{20}$$

In (20), integration is also over the entire volume of the drop.

For $\lambda_e < \lambda_i$, we can neglect the Θ dependence of the dynamic viscosity coefficient in the drop–liquid medium system and assume that the viscosity depends only on the temperature $t_e^{(0)}(y)$; i.e., $\mu_e(t_e) = \mu_e(t_e^{(0)})$. With this in mind, expression (1) takes the form

$$\mu_e = \mu_\infty \exp\left\{-A\frac{\gamma}{y}\right\}\left[1 + \sum_{n=1}^{\infty} F_n\frac{\gamma^n}{y^n}\right]. \tag{21}$$

Formula (21) will subsequently be used for finding the velocity and pressure fields in the vicinity of the heated drop.

Now, let us find the first-order approximations. Substituting (18) and (19) into (16) and (17), we come to

$$-\omega\frac{V_r^e}{y^2} = \Delta t_e^{(1)}, \tag{22}$$

$$\mathrm{Pr}_\infty\beta\frac{V_r^i}{y^2}\left(\int_1^y \psi_0 dy - \gamma\frac{\lambda_e}{\lambda_i}\right) + Q_1\cos\Theta = \Delta t_i^{(1)}, \tag{23}$$

where $\omega = \mathrm{Pr}_\infty\gamma$.

From (22) and (23), it follows that one must first solve the hydrodynamic problem, i.e., determine the velocity fields inside and outside the drop, in order to find $t_e^{(1)}$ and $t_i^{(1)}$.

DETERMINATION OF THE DRAG FORCE

Substituting (22) into the equations of hydrodynamics, taking into account (10), and separating the variables, we come to an equation similar to that obtained in [17]. Eventually, we have the following expressions for the components of the mass velocity and pressure that satisfy boundary conditions (7) and (8):

$$V_r^e(y,\Theta) = \cos\Theta(A_1G_1 + A_2G_2),$$

$$V_\Theta^e(y,\Theta) = -\sin\Theta(A_1G_3 + A_2G_4), \tag{24}$$

$$p_e(y,\Theta) = 1 + \eta_e\cos\Theta(A_1G_5 + A_2G_6),$$

$$V_r^i(y,\Theta) = \cos\Theta(A_3 + A_4y^2),$$

$$V_\Theta^i(y,\Theta) = -\sin\Theta(A_3 + 2A_4y^2), \tag{25}$$

$$p_i(y,\Theta) = p_{0i} + 10\eta_i\cos\Theta y^2 A_4, \quad \eta = \mu/\mu_\infty,$$

where

$$G_1 = -\frac{1}{y^3}\sum_{n=0}^{\infty}\frac{\Delta_n^{(1)}}{(n+3)y^n}, \quad G_2 = -\frac{1}{y}\sum_{n=0}^{\infty}\frac{\Delta_n^{(2)}}{(n+1)y^n} - \frac{\alpha}{y^3}\sum_{n=0}^{\infty}\left[(n+3)\ln\frac{1}{y} - 1\right]\frac{\Delta_n^{(1)}}{(n+3)^2y^n},$$

$$G_3 = G_1 + \frac{y}{2}G_1^{\mathrm{I}},$$

$$G_5 = \frac{y^2}{2}G_1^{\mathrm{III}} + y\left(3 + \frac{1}{2}\sum_{n=0}^{\infty} s_n\frac{\gamma^n}{y^n}\right)G_1^{\mathrm{II}} \tag{26}$$

$$+\left(2+\sum_{n=0}^{\infty}s_n\frac{\gamma^n}{y^n}\right)G_1^{\mathrm{I}},\quad G_4=G_2+\frac{y}{2}G_2^{\mathrm{I}},$$

$$G_6=\frac{y^2}{2}G_2^{\mathrm{III}}+y\left(3+\frac{1}{2}\sum_{n=0}^{\infty}s_n\frac{\gamma^n}{y^n}\right)G_2^{\mathrm{II}}+\left(2+\sum_{n=0}^{\infty}s_n\frac{\gamma^n}{y^n}\right)G_2^{\mathrm{I}},$$

$s_n=AF_{n-1}-nF_n-\sum_{k=1}^{n}s_{n-k}F_k$, $F_0=1$, and $F_n=0$ at $n<0$. In (26), G_k^{I}, G_k^{II}, and G_k^{III} are the first-, second-, and third-order derivatives of the associated functions with respect to y ($k=1, 2$).

The coefficients $\Delta_n^{(1)}$ and $\Delta_n^{(2)}$ are found from the recurrent relations

$$\Delta_n^{(1)}=-\frac{1}{n(n+5)}\sum_{k=1}^{n}[(n+4-k)\times\{\alpha_k^{(1)}(n+5-k)-\alpha_k^{(2)}\}+\alpha_k^{(3)}]\gamma^k\Delta_{n-k}^{(1)}\ (n\geq 1),\tag{27}$$

$$\Delta_n^{(2)}=-\frac{1}{(n+3)(n-2)}\left[-6\alpha_n^{(4)}\gamma^n+\sum_{k=1}^{n}\{(n+2-k)\times[(n+3-k)\alpha_k^{(1)}-\alpha_k^{(2)}]+\alpha_k^{(3)}\}\gamma^k\Delta_{n-k}^{(2)}+\alpha\sum_{k=0}^{n}\{(2n+5-2k)\alpha_k^{(1)}-\alpha_k^{(2)}\}\gamma^k\Delta_{n-k-2}^{(1)}\right]\ (n\geq 3).\tag{28}$$

Upon calculating the coefficients $\Delta_n^{(1)}$ and $\Delta_n^{(2)}$ by formulas (27) and (28), it should be taken into account that

$$\Delta_0^{(1)}=-3,\quad \Delta_0^{(2)}=-1,\quad \Delta_2^{(2)}=1,\quad \alpha_0^{(1)}=\alpha_0^{(4)}=1,$$

$$\alpha_0^{(2)}=4,\quad \alpha_0^{(3)}=-4,\quad \alpha_n^{(1)}=F_n,$$

$$\alpha_n^{(2)}=(4-n)F_n+AF_{n-1},\quad \alpha_n^{(4)}=A^n/n!,$$

$$\alpha_n^{(3)}=2AF_{n-1}-2(2+n)F_n,$$

$$\Delta_1^{(2)}=-\gamma[6\alpha_1^{(4)}+2(3\alpha_1^{(1)}-\alpha_1^{(2)})+\alpha_1^{(3)}],$$

$$\alpha=\frac{\gamma}{15}\{-6\gamma\alpha_2^{(4)}+[3(4\alpha_1^{(1)}-\alpha_1^{(2)})+\alpha_1^{(3)}]\Delta_1^{(2)}-[2(3\alpha_2^{(1)}-\alpha_2^{(2)})+\alpha_2^{(3)}]\gamma\}.$$

Substituting (24) into (22), we have for $t_e^{(1)}$

$$\Delta t_e^{(1)}=-\frac{\omega}{y^2}G(y)\cos\Theta.\tag{29}$$

Here, $G(y)=A_1G_1+A_2G_2$. A solution for $t_e^{(1)}$ is sought in the form

$$t_e^{(1)}=\tau_e(y)\cos\Theta.\tag{30}$$

Substituting (30) into (29), we check that the variables are separable; then, for $t_e^{(1)}$, we have

$$\frac{d^2\tau_e}{dy^2}+\frac{2}{y}\frac{d\tau_e}{dy}-\frac{2}{y^2}\tau_e=-\frac{\omega}{y^2}G.\tag{31}$$

The general solution to Eq. (31) that satisfies the boundary conditions for $y\longrightarrow\infty$ has the form

$$t_e^{(1)}(y,\Theta)=\left\{\frac{\Gamma}{y^2}+y+\omega\sum_{k=1}^{2}A_k\tau_k\right\}\cos\Theta,\tag{32}$$

where

$$\tau_1(y)=\frac{1}{y^3}\sum_{n=0}^{\infty}\frac{\Delta_n^{(1)}}{(n+1)(n+3)(n+4)y^n},$$

$$\tau_2(y)=-\frac{1}{y}\left\{-\frac{1}{2}+\frac{\Delta_1^{(2)}}{6y}\ln y-\sum_{n=2}^{\infty}\frac{\Delta_n^{(2)}}{(n^2-1)(n+2)y^n}-\frac{\alpha}{y^2}\sum_{n=0}^{\infty}\left[(n+1)(n+3)(n+4)\ln\frac{1}{y}-(3n^2+16n+19)\right]\times\frac{\Delta_n^{(1)}}{(n+1)^2(n+3)^2(n+4)^2y^n}\right\}.$$

A solution for $t_i^{(1)}$ is sought in the form

$$t_i^{(1)}=\tau_i(y)\cos\Theta.\tag{33}$$

Substituting (33) into (23), we come to the equation for $t_i^{(1)}$:

$$\mathrm{Pr}_\infty\beta\frac{V_r^i}{y^2}\left(\int_1^y\psi_0dy-\gamma\frac{\lambda_e}{\lambda_i}\right)+Q_1\cos\Theta=\frac{d^2\tau_i}{dy^2}+\frac{2}{y}\frac{d\tau_i}{dy}-\frac{2}{y^2}\tau_i.\tag{34}$$

The general solution to Eq. (34) that satisfies the boundary condition for $y\longrightarrow 0$ is

$$t_i^{(1)}=\cos\Theta\left[By+\frac{C}{y^2}+\omega\frac{\beta\lambda_e}{2\lambda_i}\left(A_3-\frac{A_4}{2}y^2\right)+\frac{1}{3}\left\{y\int_1^y\frac{\psi_1}{y^2}dy-\frac{1}{y^2}\int_1^y\psi_1ydy\right\}\right.\tag{35}$$

$$+\frac{1}{3}\left\{y\int_1^y\Omega\left(A_4+\frac{A_3}{y^2}\right)dy-\frac{1}{y^2}\int_1^y\Omega(A_3y+A_4y^3)dy\right\}\Bigg],$$

where

$$\psi_1 = -\frac{3R^2}{2\lambda_i T_\infty}y^2\int_{-1}^{+1}q_i x dx,$$

$$C = \frac{RJ}{3T_\infty\lambda_i}-\frac{\omega}{6}\frac{\beta\lambda_e}{\lambda_i}\left(A_3+\frac{A_4}{2}\right),\quad V=\frac{4}{3}\pi R^3,$$

$$J=\frac{1}{V}\int_V q_i z dV,\quad \Omega=\mathrm{Pr}_\infty\beta\int_1^y\psi_0 dy,$$

$$z=r\cos\Theta,\quad \int_V q_i z dV$$

is the dipole moment of the heat source density.

The constants of integration Γ and B are found from the boundary conditions on the surface (the equality of the temperatures and fluxes). Specifically,

$$\Gamma = -\frac{1}{\delta}\left(1-\frac{\lambda_e}{\lambda_i}\right)+\frac{RJ}{\lambda_i\delta T_\infty}$$
$$-\frac{\omega}{\delta}\left[A_1\left(\tau_1-\frac{\lambda_e}{\lambda_i}\tau_1^{\mathrm{I}}\right)+A_2\left(\tau_2-\frac{\lambda_e}{\lambda_i}\tau_2^{\mathrm{I}}\right)\right],$$

where $\delta = 1 + 2(\lambda_e/\lambda_i)$ and τ_1^{I} and τ_2^{I} are the first derivatives of the functions τ_1 and τ_2 with respect to y.

Thus, we determined the temperature fields inside and outside the drop in the first approximation in ε. Now, the constants of integration A_1, A_2, A_3, and A_4 entering into expressions (24) and (25) can be found from the boundary conditions for the velocity components on the surface of the drop.

Below, we give the coefficient A_2 in explicit form, since the total force acting on the drop is expressed through it:

$$A_2 = -\frac{1}{\Delta}\left(N_3+\frac{\mu_e}{3\mu_i}N_4\right)$$
$$-\varepsilon\frac{2G_1}{3U\Delta}\left[3\frac{\lambda_e}{\delta\lambda_i}+\frac{RJ}{\delta T_\infty\lambda_i}-\frac{\omega}{\delta}\frac{\lambda_e}{\lambda_i}\frac{2\tau_1+\tau_1^{\mathrm{I}}}{G_1}\right]$$
$$\times\left(K_{\mathrm{tc}}^e\frac{\nu_e^{\mathrm{s}}}{t_{\mathrm{s}}}-K_{\mathrm{tc}}^i\frac{\nu_i^{\mathrm{s}}}{t_{\mathrm{s}}}+\frac{R}{3\mu_i^{\mathrm{s}}}\frac{\partial\sigma}{\partial t_i}\right),$$

where

$$\Delta = N_1+\frac{\mu_e}{3\mu_i}N_2+\frac{2}{3\mu_\infty}\rho_e R\frac{\omega}{\delta}\frac{\lambda_e}{\lambda_i}[G_1\Phi_2-G_2\Phi_1]$$
$$\times\left(K_{\mathrm{tc}}^e\frac{\nu_e^{\mathrm{s}}}{t_{\mathrm{s}}R}-K_{\mathrm{tc}}^i\frac{\nu_i^{\mathrm{s}}}{t_{\mathrm{s}}R}+\frac{1}{3\mu_i^{\mathrm{s}}}\frac{\partial\sigma}{\partial t_i}\right),$$

$$\Phi_k = 2\tau_k+\tau_k^{\mathrm{I}}\ (k=1,2).$$

The force acting on the drop is found by integrating the stress tensor over its surface [18]:

$$F = \oint_S\{-P_e\cos\Theta+\sigma_{rr}^e\cos\Theta$$
$$-\sigma_{r\Theta}^e\sin\Theta\}r^2\sin\Theta d\Theta d\varphi|_{y=1},\qquad(36)$$

where

$$\sigma_{rr}^e = 2\mu_e\frac{\partial U_r^e}{\partial r},\quad \sigma_{r\Theta}^e=\mu_e\left(\frac{\partial U_\Theta^e}{\partial r}+\frac{1}{r}\frac{\partial U_r^e}{\partial\Theta}-\frac{U_\Theta^e}{r}\right).$$

Substituting expressions (24) into (36) yields after integration

$$F = 4\pi R\mu_\infty U A_2\exp\{-A\gamma\}.\qquad(37)$$

In view of the explicit form of A_2, the total force acting on a heated drop subjected to an external temperature gradient is the additive sum of the viscous force $\mathbf{F}_\mu$ of the medium, thermophoretic force $\mathbf{F}_{\mathrm{th}}$, force $\mathbf{F}_q$ proportional to the dipole moment of the density of heat sources nonuniformly distributed in the bulk of the drop, and force F_{m} due to the motion of the medium (i.e., the force including convective terms in the heat conduction equation):

$$\mathbf{F} = \mathbf{F}_\mu+\varepsilon\mathbf{F}^{(1)},\quad \mathbf{F}^{(1)}=\mathbf{F}_{\mathrm{th}}+\mathbf{F}_q+\mathbf{F}_{\mathrm{m}},\qquad(38)$$

where $\mathbf{F}_\mu = -6\pi R\mu_\infty U f_\mu\mathbf{n}_z$, $\mathbf{F}_{\mathrm{th}} = -6\pi R\mu_\infty f_{\mathrm{th}}\mathbf{n}_z$, $\mathbf{F}_q = -6\pi R\mu_\infty f_q J\mathbf{n}_z$, and $\mathbf{F}_{\mathrm{m}} = -6\pi R\mu_\infty f_{\mathrm{m}}\mathbf{n}_z$.

The coefficients f_μ, t_{th}, f_q, and f_{m} can be estimated from the expressions

$$f_\mu = \frac{2}{3\Delta}\left[N_3+\frac{\mu_e^{\mathrm{s}}}{3\mu_i^{\mathrm{s}}}N_4\right]\exp\{-A\gamma\},$$

$$f_{\mathrm{th}} = \frac{4G_1}{3\Delta R}\frac{\lambda_e^{\mathrm{s}}}{\lambda_i^{\mathrm{s}}\delta}\left(K_{\mathrm{tc}}^e\frac{\nu_e^{\mathrm{s}}}{t_{\mathrm{s}}}-K_{\mathrm{tc}}^i\frac{\nu_i^{\mathrm{s}}}{t_{\mathrm{s}}}+\frac{R}{3\mu_i^{\mathrm{s}}}\frac{\partial\sigma}{\partial t_i}\right)\exp\{-A\gamma\},$$

$$f_q = \frac{4}{9\Delta}\frac{G_1}{\lambda_i^{\mathrm{s}}T_\infty\delta}\left(K_{\mathrm{tc}}^e\frac{\nu_e^{\mathrm{s}}}{t_s}-K_{\mathrm{tc}}^i\frac{\nu_i^{\mathrm{s}}}{t_{\mathrm{s}}}+\frac{R}{3\mu_i^{\mathrm{s}}}\frac{\partial\sigma}{\partial t_i}\right)\exp\{-A\gamma\},$$

$$f_{\mathrm{m}} = -\frac{4}{9\Delta R}\frac{\omega}{\delta}\frac{\lambda_e^{\mathrm{s}}}{\lambda_i^{\mathrm{s}}}[2\tau_1+\tau_1^{\mathrm{I}}]$$
$$\times\left(K_{\mathrm{tc}}^e\frac{\nu_e^{\mathrm{s}}}{t_{\mathrm{s}}}-K_{\mathrm{tc}}^i\frac{\nu_i^{\mathrm{s}}}{t_{\mathrm{s}}}+\frac{R}{3\mu_i^{\mathrm{s}}}\frac{\partial\sigma}{\partial t_i}\right)\exp\{-A\gamma\}.$$

In the expressions for the coefficients f_μ, f_{th}, f_q, and f_m, the index s refers to the quantities taken at the mean surface temperature T_s of the drop, which is determined from (20), and the parameters N_1–N_4, τ_1, τ_2, τ_1^I, τ_2^I, G_1, and G_2 are taken at $y = 1$ ($N_4 = 2G_1^I + G_1^{II}$, $N_3 = -G_1^I$, $N_2 = [G_2(2G_1^I + G_1^{II}) - G_1(2G_2^I + G_2^{II})]$, and $N_1 = G_1G_2^I - G_2G_1^I$).

Equating the total force to zero, we obtain an expression for the directed velocity of the drop in a given external temperature gradient:

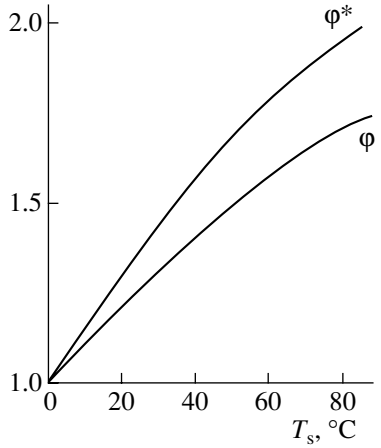
$$\mathbf{U} = -\varepsilon U^{(1)}\mathbf{n}_z, \quad U^{(1)} = U_{th} + U_q + U_m, \qquad (39)$$

where $U_{th} = h_{th}$, $U_q = h_q J$, and $U_m = h_m$ ($h_{th} = f_{th}/f_\mu$, $h_q = f_q/f_\mu$, and $h_m = f_m/f_\mu$).

If the drop is heated only slightly, that is, its mean surface temperature is close to the ambient temperature at infinity ($\gamma \longrightarrow 0$), the temperature dependence of the viscosity coefficient can be neglected; then, $G_1 = 1$, $G_1^I = -3$, $G_1^{II} = 12$, $G_2 = 1$, $G_2^I = -1$, $G_2^{II} = 2$, $N_1 = 2$, $N_2 = 6$, $N_3 = 3$, $N_4 = 6$, $\tau_1 = -1/4$, $\tau_1^I = 3/4$, $\tau_2 = 1/2$, and $\tau_2^I = -1/2$, and formulas (38) and (39) pass to expressions that are well-known from the literature [1–6].

If the distribution of heat sources over the volume of the heated drop is known, formulas (38) and (39) allow one to take into account (1) the effect of the motion of the medium on the drag force acting on the drop and (2) the effect of drop surface heating on the thermocapillary force and velocity for arbitrary temperature differences between the surface and regions away from the drop. These formulas also account for the exponential temperature dependence of the viscosity under an external temperature gradient. We emphasize once again that they are of general character.

These formulas imply that the magnitude and the direction of the force $\mathbf{F}^{(1)}$ and velocity $\mathbf{U}^{(1)}$ are also affected by the direction of the dipole moment of the heat source density $\int_V q_i z dV$. If, for example, the drop heats up, absorbing electromagnetic energy, the dipole moment can be both negative (a major part of the thermal energy is released in that side of the drop facing the radiation) and positive (a major part of the thermal energy is released in the opposite side), depending on the optical properties of the drop. Note that the surface tension decreases with temperature for most liquids, i.e., $\partial\sigma/\partial t_i < 0$, and that the value of $\int_V q_i z dV$ may be both positive and negative. Accordingly, the magnitude and the direction of the force $\mathbf{F}^{(1)}$ and velocity $\mathbf{U}^{(1)}$ vary.



Functions φ and φ* vs. mean surface temperature T_s.

In addition, these formulas suggest that the forced $\mathbf{F}^{(1)}$ and the velocity $\mathbf{U}^{(1)}$ also depend significantly on the thermal conductivity of the drop. At λ_i approaching infinity, $\mathbf{F}^{(1)}$ and $\mathbf{U}^{(1)}$ tend to zero with the dipole moment of the heat source density fixed.

For $\mu_i \longrightarrow \infty$, the above formulas can be used for estimating the force and the velocity of a hard nonuniformly heated particle in a viscous liquid where a constant temperature gradient is maintained.

The figure illustrates the effect of heating the surface of the drop on its drift velocity. The curves relate the function $\varphi = h_{th}/h_{th}|_{T_s = 273\ \mathrm{K}}$ to the mean surface temperature T_s of the drop. Estimates were obtained for mercury drops suspended in water at $T_\infty = 273$ K and $R = 15$ μm ($\partial\sigma/\partial T = -2 \times 10^{-4}$ N/mK, $\mathrm{Pr}_\infty = 12.99$). The value of the function φ* was estimated by formula (39) for a small temperature difference ($\gamma \longrightarrow 0$), but the molecular transfer coefficients were taken at the mean surface temperature T_s.

To estimate the contribution of the motion of the medium to the thermocapillary drift of the drop, it is necessary to clarify the nature of the heat sources. Knowing their nature, one can find an expression for the dipole moment of the heat source density. To analyze the situation qualitatively, let us consider the simplest case. We assume that the drop heats up, absorbing electromagnetic radiation as a black body. Under these conditions, absorption takes place in a thin layer of thickness $\delta R \ll R$ adjacent to the heated part of the drop. In this case, the heat source density within a layer of thickness δR is given by

$$q_i(r, \Theta) = \begin{cases} -\dfrac{I}{\delta R}\cos\Theta, & \dfrac{\pi}{2} \le \Theta \le \pi, \quad R - \delta R \le r \le R \\ 0, & 0 \le \Theta \le \dfrac{\pi}{2}, \end{cases} \qquad (40)$$

where I is the incident radiation intensity. It is related to the mean relative surface temperature of the drop as

$$t_s = 1 + \frac{R}{4\lambda_i T_\infty} I.$$

In view of (40), the integrals entering into the expression for the force acting on the heated drop and the drift velocity are calculated directly:

$$\int_V q_i(r,\Theta)z\,dV = -\frac{2}{3}\pi R^3 I, \quad \int_V q_i(r,\Theta)\,dV = \pi R^2 I.$$

Accordingly, formulas (38) and (39) can be represented in a more compact form:

$$\mathbf{F} = -6\pi R\mu_\infty U\varphi_\mu \mathbf{n}_z - \varepsilon 6\pi R\mu_\infty f_{\rm ph}\mathbf{n}_z, \qquad \mathbf{U}_{\rm ph} = -\varepsilon h_{\rm ph}\mathbf{n}_z, \tag{41}$$

where

$$\varphi_\mu = \frac{2}{3\Delta_h}\left[N_3 + \frac{\mu_e^{\rm s}}{3\mu_i^{\rm s}}N_4\right]\exp\{-A\gamma\}, \quad h_{\rm ph} = f_{\rm ph}/\varphi_\mu,$$

$$f_{\rm ph} = \frac{4}{3\mu_i^{\rm s}}\frac{G_1}{\Delta_h\delta}\frac{\lambda_e^{\rm s}}{\lambda_i^{\rm s}} \times\left[1 - \frac{R}{6\lambda_e^{\rm s}T_\infty}I\left(\frac{\mathrm{Pr}_\infty}{2}\frac{2\tau_1+\tau_1^{\rm I}}{G_1}+1\right)\right]\frac{\partial\sigma}{\partial t_i}\exp\{-A\gamma\},$$

$$\Delta_h = N_1 + \frac{\mu_e^{\rm s}}{3\mu_i^{\rm s}}N_2 + \frac{R^2}{6\mu_\infty}\frac{\rho_e}{\mu_i^{\rm s}}\frac{\mathrm{Pr}_\infty}{\delta\lambda_i^{\rm s}}I\frac{\partial\sigma}{\partial t_i}[G_1\Phi_2 - G_2\Phi_1].$$

When deriving formulas (41), we took into consideration that the thermal creep coefficient is very small [3]; therefore, it was omitted in numerical calculations. If the motion of the environment is not taken into account ($\omega = 0$),

$$\mathbf{F} = -6\pi R\mu_\infty U\varphi_\mu^*\mathbf{n}_z - \varepsilon 6\pi R\mu_\infty f_{\rm ph}^*\mathbf{n}_z, \qquad \mathbf{U}_{\rm ph}^* = -\varepsilon h_{\rm ph}^*\mathbf{n}_z, \tag{42}$$

where

$$\varphi_\mu^* = \frac{2}{3\Delta_h^*}\left[N_3 + \frac{\mu_e^{\rm s}}{3\mu_i^{\rm s}}N_4\right]\exp\{-A\gamma\}, \quad h_{\rm ph}^* = f_{\rm ph}^*/\varphi_\mu^*,$$

$$f_{\rm ph}^* = \frac{4}{3\mu_i^{\rm s}}\frac{G_1}{\Delta_h^*\delta}\frac{\lambda_e^{\rm s}}{\lambda_i^{\rm s}}\left[1 - \frac{R}{6\lambda_e^{\rm s}T_\infty}I\right]\frac{\partial\sigma}{\partial t_i}\exp\{-A\gamma\},$$

$$\Delta_h^* = N_1 + \frac{\mu_e^{\rm s}}{3\mu_i^{\rm s}}N_2.$$

Formula (42) lacks the terms responsible for the motion of the medium (convective terms in the heat conduction equation), unlike (41). The convective term depends on the mean surface temperature and the Prandtl number (it is proportional to the Prandtl number times the relative temperature difference between the surface of the drop and regions away from it). Therefore, if the temperature difference is large and the Prandtl number is high (which is a possibility in liquids), convective terms in the heat conduction equation may considerably change the qualitative picture of the thermocapillary drift of the drop. In gases, consideration of the motion of the environment cannot affect the drift of the drop significantly, because the Prandtl number in gases is about unity. If the temperature difference is small ($\gamma \longrightarrow 0$), the motion of the surrounding medium should be taken into account when the drop drifts in viscous liquids with a high Prandtl number.

Formulas for the motion of a hard spherical particle can be obtained in a similar way.

To illustrate the contribution of the motion of the medium to the force and velocity of the thermocapillary drift of the heated drop, Tables 3 and 4 summarize the values of the functions $f_{\rm ph} = f_{\rm ph}/f_{\rm ph}|_{T_{\rm s}=273\ {\rm K}}$ and $h_{\rm ph} = h_{\rm ph}/h_{\rm ph}|_{T_{\rm s}=273\ {\rm K}}$, respectively. The estimates were made for mercury drops suspended in water at $T_\infty = 273$ K. The values of $f_{\rm ph}^{\rm av}$ and $h_{\rm ph}^{\rm av}$ were estimated from formulas (41) for a small temperature difference ($\gamma \longrightarrow 0$), but the molecular transport coefficients were taken at the mean surface temperature $T_{\rm s}$ of the drop.

Table 3

$T_{\rm s}$, °C	$f_{\rm ph}$	$f_{\rm ph}^{\rm av}$
0	1.00000000	1.00000000
10	0.75679462	1.13172965
20	0.55865389	1.25217591
30	0.40612812	1.35890535
40	0.29192483	1.44815402
50	0.20865530	1.52363269
60	0.14791875	1.58086862
70	0.10410262	1.62387314
80	0.07247662	1.65233041
90	0.04977532	1.66842426

Table 4

$T_{\rm s}$, °C	$h_{\rm ph}$	$h_{\rm ph}^{\rm av}$
0	1.00000000	1.00000000
10	1.04610555	1.1018381
20	1.06591031	1.1918477
30	1.06609551	1.2690664
40	1.04899473	1.3307655
50	1.01989975	1.3807801
60	0.97702898	1.4155379
70	0.92304912	1.4387383
80	0.85740504	1.4502438
90	0.78130662	1.4518747

DISTORTION OF THE SURFACE SHAPE

The shape of the drop is unknown and should be found from the solution; therefore, boundary conditions (5)–(7) are set for a boundary of unknown shape. Since we restrict our analysis to first-order corrections, then

$$\sigma = \sigma_0 + \varepsilon\sigma^{(1)}, \tag{43}$$

where σ_0 is the zeroth-order term in the expansion of the function $\sigma(x)$ in Legendre polynomials $P_n(x)$ ($x = \cos\Theta$).

The shape of the drop is sought in the form [14]

$$r = R[1 + \varepsilon\xi]. \tag{44}$$

Let us expand the quantities $\sigma(\Theta)$ and $\xi(\Theta)$ in Legendre polynomials:

$$\sigma = \sum_{n=0}^{\infty}\sigma_n P_n(\cos\Theta), \quad \xi = \sum_{n=0}^{\infty}\xi_n P_n(\cos\Theta). \tag{45}$$

From the constancy condition for the volume of the drop, it follows that $\xi_0 = 0$. Bearing in mind that the origin of the coordinate system is placed at the center of mass of the heated particle, we have

$$\int_0^{\pi}\xi\sin^2\Theta d\Theta = 0, \tag{46}$$

hence,

$$\xi_1 \equiv 0. \tag{47}$$

When solving the problem, we did not consider the boundary condition for the normal components of the stress tensor. Up to terms proportional to ε, the boundary condition for the normal stresses on the surface takes the form [17]

$$\sigma_n^{e(1)} - \sigma_n^{i(1)} = \sigma_0 H^{(1)} + 2\frac{\sigma^{(1)}}{R}. \tag{48}$$

Here,

$$2H = \frac{1}{R_1} + \frac{1}{R_2} = \frac{2}{R} + \varepsilon H^{(1)},$$

R_1 and R_2 are the principal radii of curvature of the drop; and H is the mean curvature of the surface, which in the axisymmetric case is given by [17]

$$H^{(1)} = -\frac{2}{R}\xi - \frac{1}{R\sin\Theta}\frac{\partial}{\partial\Theta}\left(\sin\Theta\frac{\partial\xi}{\partial\Theta}\right). \tag{49}$$

In view of (45) and (47), we obtain expression (49) in the form

$$H^{(1)} = \sum_{n=2}^{\infty}\frac{(n+2)(n-1)}{R}\xi_n P_n(\cos\Theta). \tag{50}$$

Thus, as follows from (48) with regard for (50), a nonuniformly heated drop, when moving, retains its spherical shape within the approximation adopted in this work.

REFERENCES

1. Yu. I. Gupalo, A. E. Rednikov, and Yu. S. Ryazantsev, Prikl. Mat. Mekh. **53**, 433 (1989).
2. Yu. I. Yalamov and A. S. Sanasaryan, Inzh.-Fiz. Zh. **28**, 1061 (1975).
3. G. S. Mollob and A. Meison, J. Colloid Interface Sci. **44**, 339 (1973).
4. V. A. Gorodtsov, Prikl. Mekh. Tekh. Fiz., No. 6, 32 (1975).
5. V. I. Naĭdenov, Prikl. Mat. Mekh. **38**, 162 (1971).
6. Yu. I. Yalamov and V. S. Sanasaryan, Zh. Tekh. Fiz. **45**, 2052 (1975) [Sov. Phys. Tech. Phys. **20**, 1351 (1975)].
7. A. M. Golovin and V. V. Fominykh, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 1, 38 (1983).
8. O. A. Popov, Author's Abstract of Candidate's Dissertation (Moscow, 1986).
9. C. W. Oseen, Ark. Mat., Astron., Fys. **6** (29), 1 (1910).
10. L. Praudman and J. B. Pearson, J. Fluid Mech. **2**, 237 (1957).
11. A. Acrivos and T. D. Taylor, J. Phys. Fluids **5**, 387 (1962).
12. N. V. Malaĭ, Inzh.-Fiz. Zh. **73**, 728 (2000).
13. St. Bretsznajder, *Properties of Gases and Liquids. Engineering Methods of Calculation* (Wydawnictwo Naukowo-Techniczne, Warsaw, 1962; Khimiya, Moscow, 1966), translated from Polish.
14. J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics* (Prentice-Hall, Englewood Cliffs, 1965; Mir, Moscow, 1976).
15. G. Z. Gershuni, E. M. Zhukhovitskiĭ, and A. A. Nepomnyashchii, *Stability of Convective Flows* (Nauka, Moscow, 1989).
16. O. A. Volkovitskiĭ, Yu. S. Sedunov, and L. P. Semenov, *Laser Radiation Propagation in Clouds* (Gidrometeoizdat, Leningrad, 1982).
17. N. V. Malaĭ, M. A. Amatov, and E. R. Shchukin, in *Differential Equations (Qualitative Theory): Interuniversity Collection of Scientific Works* (Ryazan', 1995), pp. 102–112.
18. L. D. Landau and E. M. Lifshitz, *Mechanics of Continuous Media* (Gostekhizdat, Moscow, 1954).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1389–1396. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 44–51.

GAS DISCHARGES, PLASMA

Penetration of a Low-Pressure Reflex-Discharge Plasma into a Hollow Electrode

V. Ya. Martens

North Caucasian State Technical University, Stavropol, 355038 Russia

e-mail: VMartens@yandex.ru

Received December 24, 2001

Abstract—It is shown experimentally that the plasma of a hollow-cathode reflex discharge is characterized by a nonequilibrium electron velocity distribution. The parameters of the electron distribution, which is approximated by a superposition of two Maxwellian distributions with different temperatures, are estimated. The penetration of the discharge plasma into the hollow cathode at various cathode potentials and a gas pressure of ~10^{-2} Pa is studied. It is shown that the plasma parameters in the hollow electrode depend not only on the parameters of the reflex-discharge plasma, but also on the magnitude and configuration of the magnetic and electric fields in the plasma expansion region. It is shown that the plasma penetration can be accompanied by quasineutrality violation and the formation of space-charge double layers. Experiments confirm that the ion current from the nonequilibrium plasma exceeds the Bohm current. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

One of the methods for producing large plasma emitting surfaces in ion and electron sources is the expansion of a dense plasma penetrating from a discharge chamber through a small-size outlet into a special hollow electrode, whose transverse dimensions correspond to the required dimensions of the emitting surface [1, 2]. It is believed that such a plasma possesses some characteristic features, and Gabovich [1] even introduced a special term "penetrating plasma." When the potential of the positive electrode is negative with respect to the potential of the discharge plasma, the penetrating-plasma model assumes that the plasma is quasineutral throughout the expansion region. The space charge of ions, whose density decreases with distance away from the outlet of the discharge chamber as the ion flow expands and accelerates, is neutralized by the space charge of electrons, whose density falls according to the Boltzmann law [1]. The potential distribution in the penetrating plasma is determined from the quasineutrality condition. A characteristic feature of this distribution is that the potential decreases sharply with distance away from the outlet.

By analyzing experimental results of [1, 3, 4], as well as my own experimental results, we can conclude that, in many cases, the quasineutrality of the penetrating plasma is violated. For example, this occurs when the reflex-discharge plasma penetrates into the low-pressure region through a hole made in one of the cathodes. The reason why the plasma quasineutrality is violated is that the penetrating plasma undergoes the perturbing action of the electric field produced in the expansion region by neighboring bodies, in particular, by the electrode with a hole through which the plasma penetrates into the expansion region.

To study this effect, we consider the penetration of the reflex-discharge plasma into a low-pressure region at various potentials of the hollow electrode enclosing this region.

1. EXPERIMENTAL TECHNIQUE

Experiments were carried out in the test model of a large-size emitter of charged particles (Fig. 1). The plasma is produced by a reflex-discharge with a cold hollow cathode in a discharge unit formed by hollow cathode *1*; cylindrical anode *2* with an inner diameter of

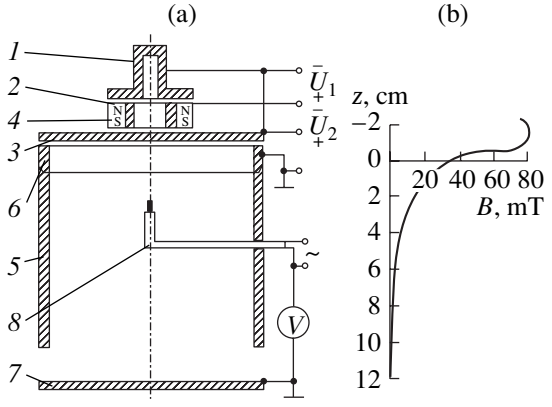


Fig. 1. (a) Schematic of the experimental device and (b) the axial profile of the magnetic induction *B*.

1063-7842/02/4711-1389$22.00 © 2002 MAIK "Nauka/Interperiodica"

10 mm; and reflex cathode *3*, which has a hole 3.5–4.0 mm in diameter. A magnetic field of ~80 mT in the discharge chamber is produced by permanent annular magnet *4*. The divergent field of this magnet (Fig. 1b) forms a magnetic nozzle. The plasma penetrates into the hollow cathode along the diverging magnetic field lines of this nozzle through the hole in the reflex cathode. Below, the hollow cathode will be referred to as an emitting-plasma former. The former is a hollow cylinder *5* that has an inner diameter of 110 mm and is attached to ring *6*. When employing this emitter in ion or electron sources, the grid emission electrode is placed on the bottom of cylinder *5*. When the accelerating voltage is applied, the emitting surface of the plasma is formed near this electrode. In order to bound the plasma region in the absence of an accelerating voltage, which makes the interpretation of experimental data more difficult, the grid electrode is replaced with solid electrode *7*, which is placed at a distance of 20 mm from cylinder *5* and is at the same potential. The former height counted from reflex cathode *3* to end electrode *7* is 130 mm.

The residual gas pressure in the discharge chamber is $\sim 9.3 \times 10^{-3}$ Pa. The working gas (air or helium) is admitted into the hollow cathode and is pumped out through an annular slot between cylinder *5* and electrode *7*. At an air flow rate of $Q = 2$–3 m^3 mPa/s, the pressure in the discharge chamber is 0.6–0.7 Pa and the pressure in the former is 1.3×10^{-2} Pa. At a helium flow rate of $Q = 5$–7 m^3 mPa/s, the partial helium pressure is 0.6–0.7 Pa in the discharge chamber and 2.4×10^{-2} Pa in the former. The discharge chamber is supplied with power from a controlled current stabilizer. The former potential U_2 is maintained with an accuracy of $<\pm 1\%$ by an individual voltage source connected between the cathodes of the discharge chamber and the former. The discharge (cathode) current I_d varied within the range 0.1–0.8 A.

The electron energy spectrum and the reflex-discharge plasma potential in the hole of the reflex cathode were measured with a multigrid probe analyzer. The probe analyzer was located 1 mm from the reflex cathode (the former was removed in this case). To reduce perturbations introduced in the plasma by the probe, it is desirable to reduce the hole diameter in the reflex cathode. On the other hand, as the hole diameter decreases, the measurements of the electron energy spectrum become more complicated, because the thickness of the ion sheath near the hole wall is as low as 0.5 mm under our experimental conditions and the dimensions of the plasma under study become rather small. For this reason, the measurements of the plasma parameters in the reflex discharge were conducted for a hole diameter of 2.6 mm. The first electrode of the probe analyzer was a 1-mm-thick copper disk; its axial hole 1.5 mm in diameter was covered with a molybdenum grid with 0.15-mm spacing.

To exclude heat effects, the measurements of the plasma parameters in the hole of reflex cathode were conducted in repetitive discharges with a pulse duration of 8 ms and a repetition rate of 10 Hz. When measuring the energy spectrum, the first electrode of the probe analyzer was at a potential that was 10–15 V higher than the plasma potential measured preliminarily with a probe analyzer by the method described in [5]. A certain excess of the potential of the first probe electrode over the measured plasma potential is required for all of the electrons arriving from the plasma can be analyzed. Otherwise, because of measurement errors and variations in the plasma potential, a fraction of the slow electrons can be cut off and excluded from analysis. The electron distribution over longitudinal velocities was determined by numerically differentiating the current at the probe collector with respect to the retarding potential. The plasma potential was taken as the value of the retarding potential at which the derivative of the collector current with respect to the retarding potential was maximum.

The measurements of the plasma parameters in the former were conducted in a discharge operating in the continuous mode. The potential distribution was measured by "floating" thermoprobe *8* (Fig. 1), which could be displaced in the axial and radial directions. The preliminary measurements showed that the potential near the former wall can differ by 20–40 V from the former voltage U_2. This voltage drop occurs across a film produced on the former surface due to vapor condensation, in particular, the vapor of the vacuum oil (the working chamber was evacuated with the help of an oil-vapor pump). For $U_2 < U_{p0}$ (where U_{p0} is the plasma potential in the hole of the reflex cathode), the film surface is charged by an ion flux to a potential higher than U_2 (Fig. 2b, curve *1*). For $U_2 > U_{p0}$, the film surface is charged by electrons and its potential falls below U_2 (Fig. 2b, curve *5*). In a preliminarily cleaned former operating at a temperature of ~250°C, a potential drop of this kind is absent. Taking into account that such films, which can also be present in devices used in practice, do not change the character of the spatial potential distribution, further measurements were performed without specially heating the former.

To determine the distribution of the electron current density over the surface of end electrode *7*, the latter was replaced with a grid electrode with 2×2-mm meshes and a transmittance of ~70%. A voltage of 1 kV, which accelerated the electrons and cut off the ions, was applied between the grid and a plane accelerating electrode (which is not shown in Fig. 1), located 10 mm from the grid. The current density distribution was measured in the plane of the accelerating electrode with the help of a movable Faraday cup and was recorded by an S8-1 storage oscillograph. To measure the ion cur-

rent distribution, the polarity of the accelerating voltage was reversed. The experimental technique is described in more detail in [6].

2. PARAMETERS OF THE REFLEX-DISCHARGE PLASMA

A typical electron distribution function over longitudinal (with respect to the magnetic field) velocities v_z in a hollow-cathode reflex discharge is plotted in Fig. 3 on the linear (Fig. 3a) and logarithmic (Fig. 3b) scales. The retarding potential U is plotted on the abscissa (here and below, the potential is counted from the cathode of the discharge chamber). The linear segment of curve *2* in the range 352–270 V indicates that, for this group of electrons (which are referred to as thermal electrons), the distribution is close to Maxwellian. Some deviation from linearity at low energies may be explained by the fact that a fraction of electrons intercepted by the grids of the probe analyzer increases as the electron energy decreases [7]. For air, the temperature of thermal electrons T_{et} increases from 7 to 13 eV as the discharge current increases from 0.2 to 0.6 A; for helium, this temperature increases from 14 to 20 eV as the discharge current increases from 0.5 to 0.8 A. In this current range, the plasma potential U_{p0} in the hole of the reflex cathode is 305–352 V for air and 305–335 V for helium, which is lower than the anode potential by 25–50 V.

Besides the thermal electrons, there is a group of high-energy electrons, whose maximum energy can exceed the energy corresponding to the cathode potential drop. Electrons with abnormally high energies were observed also in a reflex discharge with a thermal cathode [3, 8]. The measured distribution function of high-energy electrons is rather complicated, and, in further considerations, it is roughly approximated by a Maxwellian with a temperature of high-energy (hot) electrons T_{eh} that is much greater than T_{et}. A similar approximation was used, e.g., in [9]. A comparison of the measured distribution function with the distribution function calculated in the given approximation shows that the best fit (Fig. 3) is obtained when the value of T_{eh} (in eV) is determined from the relationship $T_{eh} = 0.5eU_{p0}$.

The fraction of hot electrons $\alpha = n_{eh}/(n_{et} + n_{eh}) = n_{eh}/n_i$ (where n_i, n_{et}, and n_{eh} are the densities of ions and thermal and hot electrons, respectively) was estimated for two characteristic values of the current at the probe-analyzer collector: (i) when the retarding potential was equal to the plasma potential (i.e., when the field decelerating electrons was absent in the analyzer) and (ii) when the retarding potential was equal to a potential starting from which the logarithmic dependence $f(v_z)$ on U (Fig. 3, curve *2*) became nonlinear. It was assumed that, in the former case, the current at the probe-analyzer collector was determined by both groups of electrons and, in the latter case, this current was only determined by hot electrons. In the above range of discharge currents, the value of α for air varies from 0.4×10^{-2} to 1.6×10^{-2} and is equal to 1.6×10^{-2} for helium.

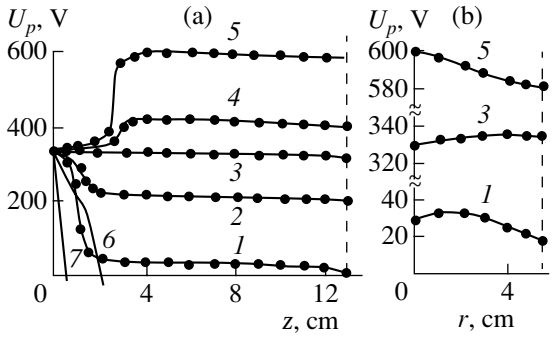


Fig. 2. (a) Axial and (b) radial (z = 8 cm) potential profiles in the former at U_2 = (*1*) 0, (*2*) 200, (*3*) 340, (*4*) 420, and (*5*) 600 V for the same discharge conditions as in Fig. 4. Curves *6* and *7* show calculations by Eq. (1) for $\beta = 0.12$ (T_{et} = 20 eV and T_{eh} = 167 eV) and α = (*6*) 0.01 and (*7*) 0.1.

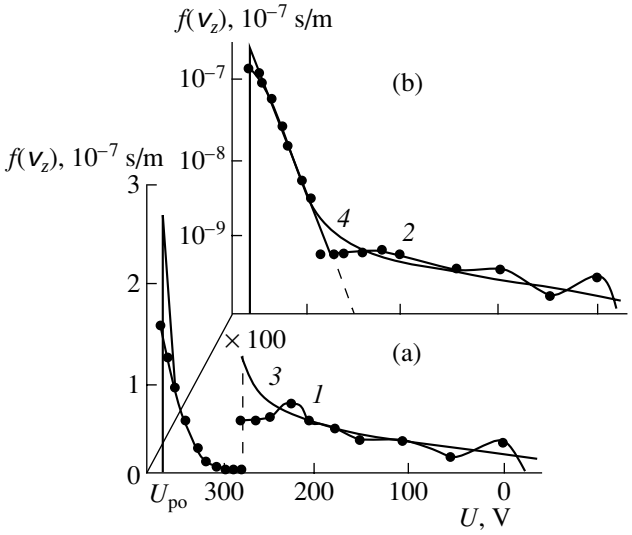


Fig. 3. Comparison of (*1*, *2*) the measured and (*3*, *4*) calculated electron distributions over longitudinal velocities plotted on (a) the linear and (b) semilogarithmic scales. The experimental curves are obtained at Q = 2 m^3 mPa/s, I_d = 0.6 A, and U_1 = 400 V; the working gas is air. In calculations, the electron distribution is assumed to be a superposition of two Maxwellian distributions with T_{et} = 13 eV, T_{eh} = 176 eV, and $\alpha = 1.6 \times 10^{-2}$.

Let us comment on the results obtained taking into account the adopted procedure of measuring α. A number of experimental observations, in particular, the higher brightness of the penetrating plasma in the axial region (Fig. 4), allow us to assume that the radial density profile of high-energy electrons in the reflex discharge is peaked at the axis more strongly than the density profile of thermal electrons. This means that the value of α at the axis exceeds the values determined from the experimental data, because the latter were

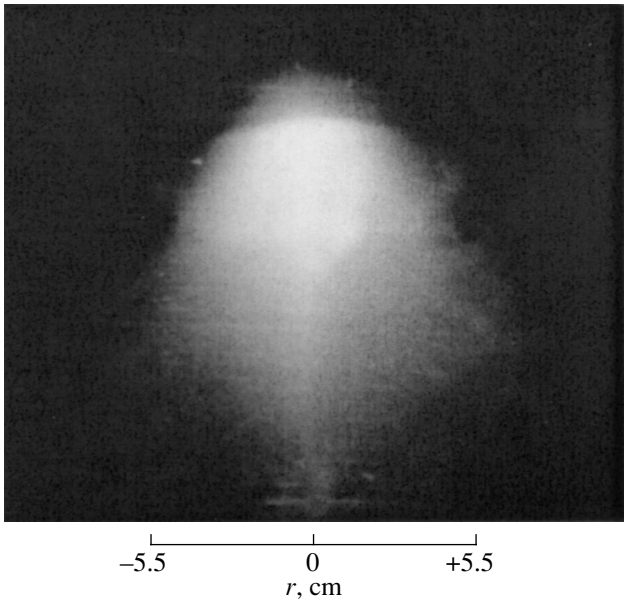


Fig. 4. Penetrating plasma at $Q = 5$ m^3 mPa/s and $I_d = 0.8$ A. The working gas is helium.

averaged over a circle 1.5 mm in diameter (i.e., over the hole in the first electrode of the probe analyzer). In addition, the value of α depends on the way electrons leave the discharge. If a fraction of electrons leave the discharge through the axial hole in the reflex cathode, then, as is shown in [10], their density at the axis is smaller as compared to the case where all electrons escape from the discharge onto the anode. The higher the efficiency of electron extraction through the hole, the more pronounced the decrease in the electron density. Taking into account the above difference in the radial density profiles of high-energy and thermal electrons, it may be argued that the extraction efficiency of high-energy electrons is higher than that of thermal electrons. It follows from here that, if the former is at the cathode potential (i.e., almost all of the thermal and high-energy electrons return to the discharge), all of the expected values of α must be much higher than the measured α values, which correspond to the case of efficient extraction of electrons from the discharge at $U_2 > U_{p0}$.

3. PENETRATION OF THE PLASMA INTO THE FORMER

Assuming that the quasineutrality holds when the plasma with the above parameters penetrates into the former, the potential distribution in the penetrating plasma can be calculated by using the model of a quasineutral plasma flowing out through a magnetic nozzle [11]. The model assumes that the plasma electron component consists of thermal and hot electrons with the temperatures T_{et} and T_{eh}, respectively. In this case, the quasineutrality equation for the penetrating plasma can be written in the form

$$(1 + 2W)^{-1/2} S_0/S - (1 - \alpha)\exp(-W) - \alpha\exp(-\beta W) = 0, \quad (1)$$

where $W = e(U_{p0} - U_p)/kT_{et}$; $\beta = T_{et}/T_{eh}$; U_{p0} and S_0 are the plasma potential and the area of the nozzle in the hole of the reflex cathode, respectively; U_p and S are the corresponding parameters of the penetrating plasma in some arbitrary cross section of the nozzle; and k is the Boltzmann constant.

The values of the area S of the nozzle cross section, which are required for calculations, were determined from the luminosity of the penetrating plasma (see Fig. 4; the photograph is made in the absence of cylinder *5* and electrode *7*).

In Fig. 2a, the axial profiles of the potential measured at various former potentials U_2 (curves *1–5*) are compared with the theoretical profiles (curves *6*, *7*) calculated by using Eq. (1) for the experimentally determined values of S_0, S, U_{p0}, T_{et}, T_{eh}, and α. It should be noted that the measured profiles depend on the former potential, which is not taken into account in the quasineutrality model [11]. Calculations show that, at distances longer than ~0.5–0.2 cm from the hole in the reflex cathode, the potential of the quasineutral plasma should become lower than the potential of the former even if the latter is at the cathode potential; this difference increases sharply with distance from the hole, which is not observed in the experiment. When comparing the calculated and measured profiles for $U_2 = 0$ (curve *1*), we can see that, near the hole, the measured potential is higher than the calculated one. This means that, for $U_2 = 0$, the penetrating plasma is not quasineutral and the ion space charge dominates over the electron charge near the reflex cathode. We can arrive at the same conclusion if we doubly differentiate experimental curve *1* (as was done in [11]). In this case, we find that, as the distance from the hole in the reflex cathode increases, the sign of the predominant space charge changes from positive to negative; i.e., a space-charge double layer is formed near the reflex cathode.

To find out why the quasineutrality of the penetrating plasma is violated and to explain the potential distributions observed in the experiment, we consider the "vacuum" potential distribution, i.e., the potential distribution in the former in the absence of a penetrating plasma.

The vacuum potential distribution in the region of interest is affected not only by the geometry and potentials of the electrodes surrounding this region, but also by the plasma potential U_{p0} in the hole of the reflex cathode. The potential U_{p0} depends on the physical conditions in the reflex discharge where the plasma is generated and, in the first approximation, is independent of the conditions in the former into which it penetrates.

The axial profile of the vacuum potential was calculated in the model in which the plasma in the hole of the reflex cathode was replaced with an equipotential spherical electrode of radius r_1 at potential U_{p0}. The z-coordinate origin lies on the surface of the spherical electrode, in the plane coinciding with the outer surface of the reflex cathode (i.e., with the surface facing the former). The vacuum electric field can be represented as a superposition of two fields—the field between the spherical electrode and the reflex cathode and the field between the reflex cathode and the former. The axial potential profile of the first field was assumed to correspond to the profile of the potential of a charged sphere, providing that the potential falls to zero (to the cathode potential) at the distance $z = r_2$, where r_2 is the former radius. The axial potential profile of the second field was determined in the same way as the axial vacuum potential profile in a Penning cell [1]. We note that, in contrast to [1], the potential reached the U_2 value (i.e., the former potential) at the above distance from the reflex cathode and then remained constant. Under these assumptions, the axial profile of the vacuum potential U_v can be represented as

$$U_v = U_{p0}\left[1 - \frac{z(r_2 + r_1)}{r_2(z + r_1)}\right] + U_2\left[1 - \left(\frac{z - r_2}{r_2}\right)^2\right]. \quad (2)$$

First, we consider the case where the former is at the cathode potential ($U_2 = 0$). Calculations of the vacuum potential distribution in the former by formula (2) show (Fig. 5) that, characteristically, there are high potential gradients near the hole in the reflex cathode. This field perturbs the plasma penetrating from the discharge; as a result, a space-charge double layer forms in the axial region near the reflex cathode. The reason why the double layer (rather than a positive space-charge layer) forms in this region is that the discharge plasma contains high-energy electrons. The possibility of such a situation occurring was predicted theoretically in [12].

One should note another factor that was not taken into consideration in [12] but also contributes to the change of the sign of the dominant charge in the axial region of the layer from positive to negative as the distance from the hole in the reflex cathode increases. Under the experimental conditions, the divergence of the ion flow penetrating from the discharge into the former is significantly greater than the divergence of the flow of high-energy electrons (compare curves *1* and *2* in Fig. 6). This is a consequence of the motion of the charged particles of the penetrating plasma through the magnetic nozzle (the Larmor radius of ions is much greater than the electron Larmor radius). For this reason, when the former is at the cathode potential, the current at its lateral surface is carried predominantly by the ions (Fig. 7), whereas the electron and ion currents at the end electrode are comparable to each other and the electron-current density is higher than the ion-current density in the center of the electrode (Fig. 6). Such a distribution of the charged particles of the penetrating plasma is consistent with the radial potential profile in the former (Fig. 2b, curve *1*).

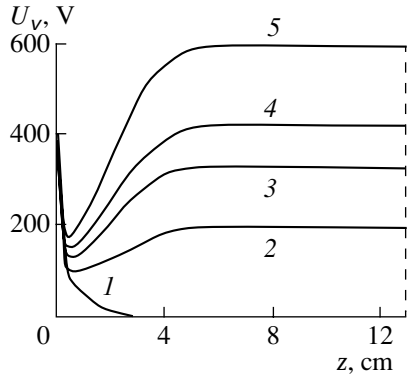


Fig. 5. Results of calculations of the "vacuum" potential distribution at U_{p0} = 335 V, r_1 = 1.7 mm, and r_2 = 55 mm for U_2 = (*1*) 0, (*2*) 200, (*3*) 340, (*4*) 420, and (*5*) 600 V.

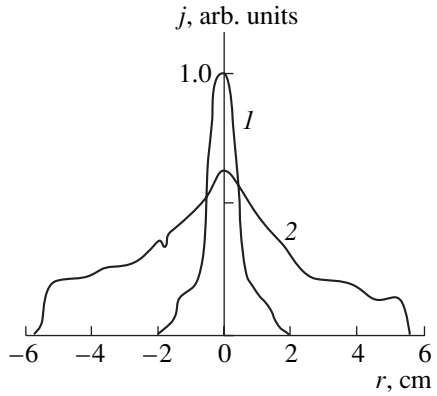


Fig. 6. Distributions of (*1*) the electron and (*2*) ion currents over the surface of the end electrode at U_2 = 0. The discharge conditions are the same as in Fig. 4.

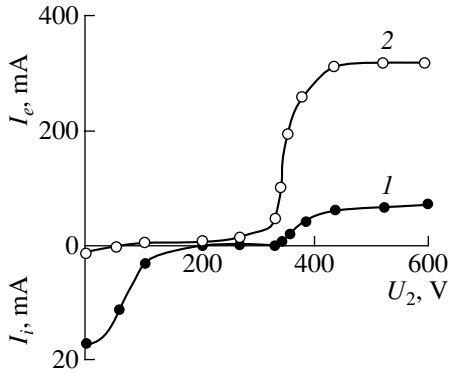


Fig. 7. Currents at (*1*) the lateral and (*2*) end surfaces of the former as functions of its potential. The discharge conditions are the same as in Fig. 4.

As the former potential increases, the axial potential gradient near the hole decreases and, at the same time, the electric field favoring the motion of ions onto the

outer surface of the reflex cathode increases. As a result, the fraction of ions arriving at the former decreases and the fraction of ions arriving at the reflex cathode increases, which is evidenced by a decrease in the ion current onto the lateral surface of the former as its potential increases (Fig. 7, curve *1*) and by traces of the intense ion cleaning and etching of the outer surface of the reflex cathode at $U_2 > 0$. As U_2 increases and the number of ions arriving at the outer surface of the reflex cathode also increases, the effect of the reflex cathode on the potential distribution in the former becomes weaker because of the ion screening of the reflex cathode. For this reason, the minimum of the potential, which is characteristic of the vacuum field (Fig. 5, curves *2*–*5*), decreases to several volts in the self-consistent electric field of the penetrating plasma even at small discharge currents of 0.1–0.2 A, and it disappears completely when the current increases to 0.8 A (Fig. 2a, curves *2*–*5*). At $U_2 \sim U_{p0}$, the axial potential of the penetrating plasma changes only slightly and is close to U_{p0}. A similar potential distribution of the plasma penetrating from the reflex discharge into a hollow glass cylinder was observed in [4] for two cases: when the reflex cathode was shielded with a special electrode that was at the anode potential, and when the gas pressure in the expansion region was increased to ~1 Pa. Seemingly, in the latter case, the reflex cathode was screened by an ion layer produced via the ionization of the gas by the charged particles of the penetrating plasma.

At $U_2 > U_{p0}$, the plasma arriving from the discharge falls into the electric field accelerating electrons toward the end electrode; as a result, the electron current at the former electrodes (particularly, at the end electrode; see Fig. 7) increases. The measured distributions of the spatial potential in the former (Fig. 2a; curves *4*, *5*) indicate that the field accelerating the electrons is located in the space-charge double layer that is formed at a certain distance from the reflex cathode. The reason why the layer is formed exactly in this spatial region becomes evident if we take into account the vacuum potential distributions (Fig. 5; curves *4*, *5*) and the fact that the reflex cathode is screened by the incoming ions of the penetrating plasma. The potential drop across the layer depends on the former potential, and the distance from the layer to the reflex cathode depends on the shape, dimensions, and potential of the former. In [6], the layer was produced in the immediate vicinity of the reflex cathode, e.g., when a disk with an axial hole was attached to the upper face of the former.

The space-charge double layer separates the penetrating plasma and the plasma produced in the former by the flow of electrons accelerated in the layer. The beam plasma is characterized by the increase in the potential toward the former axis (Fig. 2b, curve *5*). The mechanism for the production of this plasma is considered in [13].

4. EXPERIMENTAL VERIFICATION OF THE RESULTS OF [12]

The experimental results obtained allow us to verify some conclusions of [12], where a model of the intermediate region between the nonequilibrium plasma and the negative electrode was considered. The electron energy spectra (Fig. 3) indicate that the reflex-discharge plasma is nonequilibrium and the former at $U_2 < U_{p0}$ is at a negative potential with respect to this plasma.

As was mentioned above, the results obtained confirm the possibility of the existence of a space-charge double layer that forms near the reflex cathode in the axial region of the former. Another conclusion of [12] is that the ion current onto the negative electrode from the nonequilibrium plasma should exceed the Bohm current (i.e., the ion current from a plasma with an equilibrium electron velocity distribution) if the electron velocity distribution in a nonequilibrium plasma is enriched by high-energy electrons as compared to an equilibrium (Maxwellian) distribution, all other factors being the same.

To verify this conclusion, we consider the ratio of the ion current I_i to the electron current I_e extracted from the reflex-discharge plasma through the hole in the reflex cathode at a constant discharge current and at the former potentials $U_2 = 0$ and $U_2 > U_{p0}$ (Fig. 7). If the electron velocity distribution in this plasma were equilibrium at any values of the former potential, then this ratio would be equal to

$$I_i/I_e = 1.5(m_e/m_i)^{1/2}/\mu, \qquad (3)$$

where m_e and m_i are the electron and ion masses, respectively, and μ is a factor accounting for a decrease in the plasma density at the discharge axis due to the escape of electrons [10].

Estimates show that, when the electron extraction efficiency is $I_e/I_d \sim 0.5$, the factor μ is equal to 0.8. It is seen from Fig. 7 that, under our experimental conditions, this efficiency is attained at $U_2 > 450$ V.

The measured values of I_i/I_e were compared with the calculated ones. The use of helium as the working gas ensured that ions were of the same sign and mass. When deriving expression (3), all the ions were assumed to be singly charged. The assumption for helium under our experimental conditions seems to be quite reasonable, because the fraction of doubly charged helium ions is very small even in high-current reflex discharges [14].

The comparison shows that the measured value of I_i/I_e (Fig. 7) is no less than two times higher than the value calculated by expression (3). Actually, this ratio is even larger, because, when determining the current I_i at $U_2 = 0$, the net current of ions and high-energy electrons arriving at the former was measured.

The excess of the measured value of I_i/I_e over the calculated one may be explained as follows. In the case of electron extraction ($U_2 > U_{p0}$), the value of α, which characterizes the extent to which the electron velocity distribution is nonequilibrium, is small ($\sim 10^{-2}$). Consequently, the distribution can be considered to be equilibrium and the calculated and experimental values of I_e should coincide. In the case of ion extraction ($U_2 = 0$), the value of α is significantly larger (see Section 2); hence, the departure of the electron velocity distribution from equilibrium cannot be ignored. In this case, the measured ion current I_i should exceed the calculated Bohm current; this is observed experimentally as an excess of the measured value of I_i/I_e over the calculated value. Hence, the second conclusion of [12] is also confirmed experimentally. Estimates show that, for the ion current from a nonequilibrium plasma to be twice the Bohm current at the measured value $\beta = 0.12$, the value of α should be ~0.6. This value is quite reasonable and consistent with our notion of how α depends on the former potential in the reflex discharge.

CONCLUSION

The present study has shown the incorrectness of reasoning that "the penetrating plasma automatically fits the quasineutrality condition" [4] and that the electric field of the penetrating plasma "is related exclusively to the properties of the plasma leaving the source, rather than to the electrode potential" [3]. The behavior of the penetrating plasma and its parameters depend not only on the initial plasma parameters, but also on the physical conditions in the expansion region, namely, the gas pressure and the magnitudes and configurations of the magnetic and electric fields produced by the components of the discharge chamber and the former, as well as by electrodes that may be inserted into the former. The possibility of affecting the expansion of the penetrating plasma by varying these conditions allows us to predict the parameters of large-sized plasma emitters and to control them in the course of the emitter operation.

Based on the physical principles of the processes occurring in the former under different conditions, we can distinguish among the following operating modes of the large-sized emitter: the ion-beam mode ($U_2 < U_{p0}$), the penetrating-plasma mode ($U_2 \sim U_{p0}$), and electron-beam mode ($U_2 > U_{p0}$). Estimates and experiments [10] show that the penetrating-plasma mode is best suited to ion emission. In an ion emitter [15] operating in this mode, the lateral surface of the former was at the anode potential, whereas the end emission electrode was at the cathode potential. Such a potential distribution between the former electrodes makes it possible to minimize ion losses at the lateral surface of the former and to enable ion emission from the penetrating plasma through the end electrode. In the ion-beam and electron-beam modes, the efficiency of ion emission is lower than in the penetrating-plasma mode. In the former case, it is lower because of the substantial loss of the ions that are produced in the reflex discharge and reach the lateral surface of the former, and, in the latter case, it is lower because of the low efficiency of the production of ions in the former at a gas pressure of $p \sim 10^{-2}$ Pa (note that it is these ions that can provide the emission current at $U_2 > U_{p0}$).

The electron-beam mode is best suited to electron emission. The ion-beam and penetrating-plasma modes are less favorable because of both the electron deceleration in the former and the substantial loss of electrons on its lateral surface. In the electron-beam mode, the electrons leaving the reflex discharge are accelerated by the electric field in the double layer and form a divergent electron flow directed toward the emission electrode. Estimates show [13] that a fraction of these electrons in the total emission current exceeds 90%. By varying the former potential and, thereby, by changing the form of the double layer and, consequently, the flow divergence, it is possible to control the current distribution over the emitter surface [6].

It is possible that a fourth operating mode of large-sized emitter exists, namely, the volume discharge in the former. This mode can be realized either by increasing the pressure in the former to $p \geq 5 \times 10^{-2}$ Pa [6] or by creating in the former the electric and magnetic fields that ensure the existence of a volume reflex discharge [10, 16]. In both cases, the emitting plasma is generated in the volume discharge, whereas the penetrating plasma ensures the discharge ignition.

REFERENCES

1. M. D. Gabovich, *Physics and Technology of Ion Plasma Sources* (Atomizdat, Moscow, 1972), pp. 216, 258.
2. Yu. E. Kreĭndel', *Plasma Sources of Electrons* (Atomizdat, Moscow, 1977), p. 9.
3. M. D. Gabovich, L. I. Romanyuk, and E. A. Lozovaya, Zh. Tekh. Fiz. **34**, 488 (1964) [Sov. Phys. Tech. Phys. **9**, 380 (1964)].
4. M. D. Gabovich, L. I. Romanyuk, and E. A. Lozovaya, Zh. Tekh. Fiz. **35**, 94 (1965) [Sov. Phys. Tech. Phys. **10**, 70 (1965)].
5. B. I. Makov, Dokl. Akad. Nauk SSSR **198**, 327 (1971) [Sov. Phys. Dokl. **16**, 399 (1971)].
6. Yu. E. Kreĭndel', V. Ya. Martens, and V. Ya. S''edin, in *Electron Sources with Plasma Emitters*, Ed. by Yu. E. Kreĭndel' (Nauka, Novosibirsk, 1983), p. 25.
7. O. V. Kozlov, *Electrical Probe in Plasma* (Atomizdat, Moscow, 1969).

8. L. I. Romanyuk and N. E. Svavil'nyĭ, Ukr. Fiz. Zh. **21**, 981 (1976).
9. R. A. Demirkhanov, Yu. V. Kursanov, and L. P. Skripal', Zh. Tekh. Fiz. **39**, 666 (1969) [Sov. Phys. Tech. Phys. **14**, 500 (1969)].
10. *Charged Particle Sources with Plasma Emitters*, Ed. by P. M. Shchanin (Nauka, Yekaterinburg, 1993), pp. 12, 98.
11. V. V. Kuznetsov and N. N. Semashko, Zh. Tekh. Fiz. **42**, 2440 (1972) [Sov. Phys. Tech. Phys. **17**, 1890 (1972)]; Zh. Tekh. Fiz. **42**, 2609 (1972) [Sov. Phys. Tech. Phys. **17**, 2027 (1972)].
12. V. Ya. Martens, Zh. Tekh. Fiz. **72** (10), 45 (2002) [Tech. Phys. **47**, 1250 (2002)].
13. Yu. E. Kreĭndel' and V. Ya. Martens, in *Proceedings of the 5th All-Union Symposium on High-Current Electronics, Tomsk, 1984*, Part 2, p. 201.
14. Yu. D. Pigarov and P. M. Morozov, Zh. Tekh. Fiz. **31**, 476 (1961) [Sov. Phys. Tech. Phys. **6**, 342 (1961)].
15. M. G. Abraizov, V. A. Kaplin, and V. Ya. Martens, Élektron. Obrab. Mater., No. 2, 38 (1981).
16. V. Ya. Martens, S. I. Belyuk, and V. N. Posokhov, Prib. Tekh. Éksp., No. 2, 194 (1992).

Translated by N. Larionova

Technical Physics, Vol. 47, No. 11, 2002, pp. 1397–1407. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 52–63.

SOLIDS

Nanomachinery: A General Approach to Inducing Directed Motion at the Atomic Level

V. L. Popov

TU Berlin, Institute of Mechanics, Sekr. C8-4, 10623 Berlin, Germany

Institute of Strength Physics and Materials Science, Siberian Division, Russian Academy of Sciences, Akademicheskiĭ pr. 2/1, Tomsk, 634021 Russia

e-mail: v.popov@tu-berlin.de

Received September 18, 2001; in final form, May 17, 2002

Abstract—The motion of bodies in a periodic potential relief with weak attenuation is considered. When subjected to various periodic external effects, the bodies may spontaneously move with a velocity uniquely defined by the frequency of a periodic action and the space period of the potential. Principles of inducing directed motion with a strictly controllable velocity that are described in this paper can be used for (1) handling individual molecules or molecular clusters on crystal surfaces, (2) creating nanomachines—objects that are free to spontaneously move both in the absence of an external force and in the presence of a force opposite to the direction of motion (and thus capable of transporting other objects), (3) designing actuators providing a strictly controllable velocity of motion, and (4) designing controllable tribological systems by appropriately profiling tribosurfaces and applying ultrasonic actions. Under periodic external perturbations, the dependence of the mean velocity of a system on the mean applied force (which macroscopically appears as the "friction law" for the system) is shown to contain plateaus of constant velocity not only when the velocity of motion is zero but also when a set of discrete equidistant velocities is present. The problem of creating totally controllable nanomachines can be posed as the problem of controlling the width and position of these plateaus. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The modern tendency toward miniaturizing mechanical systems and the rapid advancement of nanotechnologies raise the question of theoretical limits of miniaturization. The fundamental problem in developing micromechanical systems of any level is the conversion of various kinds of energies to the energy of system directed motion. What are the minimal dimensions of mechanisms allowing for such a conversion? What are the physical concepts underlying the operation of these mechanisms? In answering these questions, we proceed from the fact that atomic-sized objects on a solid surface experience the action of the periodic potential of the crystalline "substrate." We will show that this periodic background, which is natural for the atomic world, in combination with at least two independent external periodic actions can be used for generating motion in a desired direction with a controllable velocity. The minimal dimensions of the system depend only on the possibility of implementing these two periodic actions. Theoretically, the smallest nanomachine may consist of a single body provided that periodical actions have an effect on it in both the tangential (relative to the solid substrate surface) and normal directions.[1]

Most approaches to generating the directed motion of molecular objects consider the interaction of a moving object with an inhomogeneous, usually periodically structured, substrate [1–4]. The substrate can be asymmetric or symmetric. In the former case, the direction of motion obeys (is fixed by) the ratchet-and-pawl principle (see, e.g., [2, 3]). In the latter, the direction of motion is not initially fixed and is controlled dynamically. An example of a dynamically controllable machine was given in [5]. In this work, we show that the nanomachine suggested in [5] is a specific case of a wider class of systems that have different designs but identical principles of control. It should be noted that approaches to inducing directed motion by periodical (undirected on average) effects have been long known and are widely used in vibroconveying, vibroseparation, vibroimmersion, and consolidation of free-flowing materials. Mathematical methods for solving problems of motion under vibration and the applications of this effect can be found elsewhere [6]. It should be noted that motion under vibration at the microlevel has a number of specific features (for example, constant-velocity plateaus mentioned above) that are directly related to the presence of the periodic atomic potential and not found in macrosystems.

The object of investigation in this work will be a body or an ensemble of bodies subjected to a spatially periodic potential. To be specific, one may consider a

[1] For an electrically charged object, such an action can be provided, for example, by means of electromagnetic radiation that has both the tangential and normal (relative to the substrate surface) components of the electric field.

molecule or a molecular cluster adsorbed on a crystal surface. The physical nature of the objects and periodic potential, as well as its period, are of minor significance for the theory elaborated; all the effects described above can take place in systems of any scale. It is assumed that the body (bodies) are subjected to external forces that have both constant and variable components. The major issue to be tackled is how the mean velocity of a body moving in a periodic potential depends on the mean force acting on it. We show that, in the presence of periodic external forces, this dependence may be unusual; that is, the object may move in a certain direction even in the absence of a macroscopic force or in the presence of a force opposite to the direction of motion. Thus, when subjected to periodic (undirected) actions, the body can be set in motion and transport a "load." Such unusual properties of the systems considered allow us to call them nanomachines.

A physical system can be referred to as a machine only if it is controllable, i.e., can be set in motion in a desired direction with a controllable velocity and stopped. Is such controllability of an object achievable with the aid of periodical actions? We answer in the affirmative.

Consider the 1D motion of a body in a periodic potential with weak attenuation. In this case, the equation of motion has the form

$$m\ddot{x} = F - \eta\dot{x} - N\sin(2\pi x/a), \qquad (1)$$

where x is the position (coordinate) of the body, m is its mass, F is the force acting on the body, η is the attenuation coefficient, N is the periodic force amplitude, and a is the wavelength of the periodic potential.

This model was suggested in 1929 by Tomlinson, who called it a dry friction model [7]. While simple, it reflects adequately many aspects of dry friction and has been extensively used (with various modifications) in recent publications concerning the physics of dry friction (see, e.g., [8–11]). We therefore will treat the systems under study as tribological and the dependence of the mean velocity of the body on the mean applied force as a friction law. Accordingly, the theory of nanomachinery can be viewed as a division of molecular tribology.

It is worth noting that Eq. (1), when rewritten as

$$\left(\frac{\hbar C}{2e}\right)\ddot{\varphi} = j - \left(\frac{\hbar}{2eR}\right)\dot{\varphi} - j_0\sin\varphi, \qquad (2)$$

is similar to the equation describing the dynamics of a single Josephson contact [12]. In (2), $\hbar$ is the Planck constant; φ is the phase difference between contacting superconductors; C and R are the capacitance and ohmic resistance of the contact, respectively; e is the unit charge; and j_0 is the maximal current through the contact.

The mathematical identity of the two problems means that tribological systems with models obeying Eq. (1) must exhibit the same effects as in Josephson contacts. One of them is the modification of the I–V characteristic of the Josephson contact exposed to an external periodic perturbation. This modification shows up as a plateau within which the voltage is strictly constant (this effect underlies the present-day quantum voltage standard). An analogue of this effect in terms of tribology is a plateau where the mean velocity of a body remains constant within a certain interval of acting forces.[2] These plateaus play a crucial role in the subsequent analysis of nanomachines.

The physical mechanism responsible for constant-velocity plateau formation is as follows. If a body moves with a macroscopically constant velocity in a space-periodic potential, its instantaneous velocity and the force acting on it are generally periodic functions of time with a frequency depending on the mean velocity. In this case, the time-averaged force and velocity are obviously independent of the "initial phase" of these periodic functions, and this phase can be arbitrary. The situation changes qualitatively if an additional external periodic action with a frequency equal to the above-mentioned "natural" oscillation frequency of the velocity and force is imposed on the periodic velocity (force) oscillations. Since the system is nonlinear, the superposition of the external periodic perturbation will cause the product of terms with the same frequency to appear in the equation of system dynamics. Being time-averaged, these terms do not disappear, and their mean depends on the initial oscillation frequency. The dependence on the initial phase means that the mean force can be varied by tuning the phase with the mean velocity remaining unchanged. This effect of phase tuning is embodied in the Josephson effect [12].

1. MOTION OF A BODY IN A PERIODIC POTENTIAL UNDER A STATIONARY EXTERNAL ACTION

The Tomlinson model reflects many basic features of dry friction. In fact, the application of a small force to a body causes only its insignificant displacement from the minimum of the potential energy, after which the motion ceases. The resultant counter force is macroscopically treated as the static friction force. Obviously, equilibrium in a periodic potential in the presence of a constant tangential force is impossible if this force exceeds some critical value that is macroscopically perceived as the maximal static friction force. When this critical force is exceeded, the body starts sliding macroscopically.[3] In model (1), the critical force is equal to N. To maintain a macroscopically uni-

[2] The presence of this plateau means that the variation of the force within a certain range does not affect the macroscopic velocity of the body.

[3] By the "macroscopic" behavior of a body, we mean here the behavior of a body on a space scale much larger than the period of the potential. The scale defined in terms of the period of the potential is considered as microscopic in the context of this paper.

form motion in the presence of dissipation, one must apply a force that is the weaker, the lower the dissipation. If the dissipation is small, an already moving body may continue moving under the action of a force smaller than critical. Macroscopically, this means that, in general, the sliding friction force may be less than the static friction force, which is also a specific feature of dry friction.

In our model, any macroscopic motion of a body, including uniform motion, is microscopically the superposition of uniform motion and periodic oscillations, as follows from Fig. 1a. It shows the results of the numerical integration of Eq. (1) at some specific parameters of the model. The tangential force F first slowly rose from zero and then decreased again. The instantaneous velocity of the body is plotted against the instantaneous force. First, the body was at rest. After the critical value has been exceeded, the motion starts with a finite macroscopic velocity, which grows in proportion with the applied force. At some critical value of the velocity, the macroscopic motion is however suddenly interrupted: the body executes several oscillations and comes to rest.

At the macroscopic level, microscopic oscillations are not perceived. From the macroscopic point of view, the above state of motion is the quasi-stationary process of friction. An outside observer perceives the dependence of the mean velocity vs. applied force (Fig. 1b) as a macroscopic law of friction.

A law of friction is found by setting a constant acting force and calculating the mean velocity of motion. The inverse statement of the problem is however possible: one assigns a constant velocity of motion to a body placed in a periodic potential and calculates a force necessary to maintain this motion. Once the force, which obviously represents the superposition of the constant and oscillating components in this case, has been time-averaged, we obtain a certain dependence of the mean sliding velocity on the mean force, i.e., again come to a law of friction. In general, this law of friction does not necessarily coincide with that derived for a constant force. Many calculations, however, indicate that both "laws of friction" are intimately related to each other, since both approaches can be used to the same advantage in many cases. Below, we will briefly discuss a correlation between the laws of friction at a constant force and a constant velocity. This correlation will be repeatedly referred to in subsequent analysis.

Consider a body moving in a periodic potential with a constant velocity and calculate the force acting on it. This force has two components: the dissipative force, which is proportional to the velocity, and the conservative force, induced by the periodic potential:

$$F = \eta v + N\sin x. \tag{3}$$

For motion with a constant velocity v, the coordinate of the body varies in proportion to time: $x = x_0 + vt$. Thus, the force necessary to sustain uniform motion

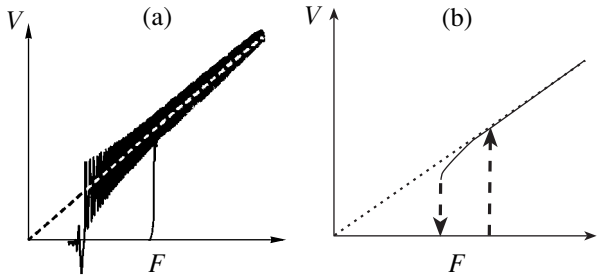


Fig. 1. (a) Instantaneous velocity V vs. instantaneous force F. The force slowly grows from zero to the maximal value exceeding the static friction force and then again goes to zero. (b) Macroscopic representation of this process: macroscopic friction law.

is given by

$$F = \eta v_0 + N\sin(x_0 + vt). \tag{4}$$

Let us determine the time-averaged value of this force. Two cases are possible.

(1) If the velocity is identically zero, $v \equiv 0$, the mean value of the first term in (4) equals zero and that of the second term, $N\sin x_0$. Thus, the mean force is

$$\langle F\rangle = N\sin x_0. \tag{5}$$

Depending on the initial coordinate, it can take any value in the range $-N < \langle F\rangle < +N$. In terms of the friction law, this corresponds to a plateau of constant, namely zero, velocity. This plateau describes nothing but the static friction force.

(2) If the velocity is other than zero, the mean value of the force due to the periodic potential is identically zero. Conversely, the dissipative force makes a nonzero contribution and grows with the velocity linearly:

$$\langle F\rangle = \eta v_0. \tag{6}$$

The friction law consists of two branches and is represented in Fig. 2.

Comparing the friction law at preset velocity with that at preset force, we see that they coincide to a great extent. The only essential difference is that a part of the linear branch near the middle of the plateau is absent in the case of preset force. In this range, the motion at preset force is unstable. Note that the critical velocity at which macroscopic motion ceases does not depend on the attenuation if the attenuation is weak [13] and is given by

$$v_{\min} = \sqrt{\frac{8}{\pi^3}\frac{Na}{m}}. \tag{7}$$

By order of magnitude, this critical velocity equals that gained by the body when it is free to roll down from the potential maximum to the potential minimum.

The form of the friction law at preset force suggests that macroscopic motion corresponding to the linear branch can be adequately approximated as free motion

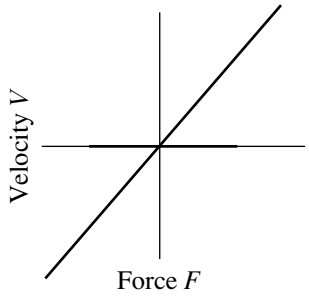


Fig. 2. Friction law (mean velocity vs. mean force) at a given constant velocity.

(with the periodic potential ignored) virtually up to the point where macroscopic motion loses stability.

Here, it is important that the basic properties of the friction law can be discovered by considering motion at preset velocity. Mathematically, such an approach is much simpler than the analysis of motion at preset force.

Thus, the macroscopic law of stationary friction is the averaged representation of the microscopically nonstationary process. Clearly, the friction law derived above is applicable only if external forces vary insignificantly during one macroscopic period of motion. If this condition is violated, the macroscopic law of friction becomes invalid and one should take a fresh look at the system dynamics at the microscopic level. In the next section, we will investigate the motion of a body in a periodic potential, assuming that an acting force (or a given velocity) also has a periodically oscillating component (along with a constant one).

2. MOTION IN A PERIODIC POTENTIAL UNDER THE ACTION OF PERIODIC FORCES

Consider the general case when both the tangential force F and the amplitude N of the force due to the periodic potential are periodic functions of time. Again two complementary approaches are possible: we preset either a periodic force and calculate the mean velocity of motion or a periodically varying velocity and determine the mean force sustaining the motion. Again, the basic properties of the friction laws are the same in both statements. We will begin with the determination of the force at a given periodically varying velocity, since this problem is mathematically much simpler.

Let a periodic perturbation with an amplitude v_1 be imposed on motion with constant velocity v_0:

$$v = v_0 + v_1 \cos \omega t. \tag{8}$$

The coordinate of the body as a function of time is given by

$$x = x_0 + v_0 t + \frac{v_1}{\omega} \sin \omega t, \tag{9}$$

and the force acting on it has the form

$$F = \eta(v_0 + v_1 \cos \omega t) + N \sin \frac{2\pi}{a}\left(x_0 + v_0 t + \frac{v_1}{\omega} \sin \omega t\right). \tag{10}$$

Hereafter, we omit the "inertial force" $m\ddot{x}$, because its mean is identically zero.

To highlight the basic properties of the friction law being derived in this way, we will proceed from the case when the dissipative force is absent and find the mean value of the conservative part of force (10), which is designated as $\tilde{F}$:

$$\tilde{F} = N \sin \frac{2\pi}{a}\left(x_0 + v_0 t + \frac{v_1}{\omega} \sin \omega t\right). \tag{11}$$

The concept of computing is the easiest to grasp when the velocity oscillation amplitude is much less than the mean sliding velocity: $v_1 \ll v_0$ (the case of an arbitrary oscillation amplitude will be discussed below). If the oscillation amplitude is small, function (11) can be expanded in the small parameter v_1/v_0:

$$\tilde{F} = N\left\{\sin \frac{2\pi}{a}(x_0 + v_0 t) + \frac{2\pi}{a} \cos \frac{2\pi}{a}(x_0 + v_0 t) \frac{v_1}{\omega} \sin \omega t\right\}. \tag{12}$$

What is the mean value of this force? Let us consider three cases.

(1) If the mean velocity of motion equals zero, $v_0 = 0$, the mean force is

$$\langle F \rangle = N \sin \frac{2\pi}{a} x_0. \tag{13}$$

It can take any value from the interval $-|N| < \langle F \rangle < |N|$ and apparently corresponds to the static friction force.

(2) If the mean velocity is other than zero, $v_0 \neq 0$, and the condition $v_0 \neq a\omega/2\pi$ is fulfilled, the mean force is identically zero:

$$\langle F \rangle \equiv 0. \tag{14}$$

(3) Finally, if the velocity is $v_0 = a\omega/2\pi$, the mean value of the first term in (12) equals zero and the second term gives a nonzero value of the mean force:

$$\langle F \rangle = N \frac{\pi v_1}{a\omega} \sin \frac{2\pi}{a} x_0. \tag{15}$$

The force depends on the initial coordinate and can take any value from the interval

$$-\left|N \frac{\pi v_1}{a\omega}\right| < \langle F \rangle < \left|N \frac{\pi v_1}{a\omega}\right|$$

with the mean velocity remaining unchanged.

Thus, schematically, the friction law appears as shown in Fig. 3a. Apparently, the dissipative force omitted here will uniformly deform the plot in proportion to the velocity (Fig. 3b). A distinctive feature of the friction law in the presence of an external periodic perturbation is the presence of a constant-velocity plateau not only for the zero velocity but also for finite velocities.

It is obvious that the plateau crosses the vertical axis (see Fig. 3b) if its half-width exceeds ηv_0. In this case, directed motion may occur at the zero mean velocity or even in the presence of a force opposite to the direction of motion. In other words, a body may produce a propulsive force and, for example, transport any load. This property is at the root of the design of nanomachines.

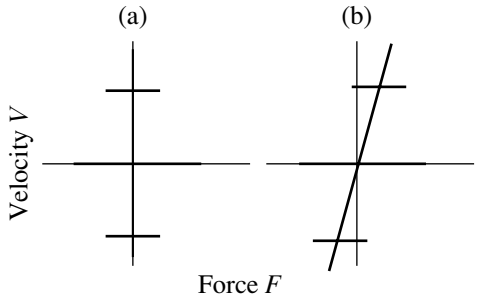


Fig. 3. Friction law (mean velocity vs. mean force) in the presence of velocity periodic oscillations: (a) without and (b) with attenuation. The characteristic property of the friction law in the presence of a periodic external perturbation is a constant-velocity plateau.

3. NANOMACHINES

By nanomachines, we mean tribosystems for which a friction law has constant-velocity plateaus whose widths and positions are controlled by external effects so that the system can be put in motion in any desired direction.

The system with such a friction law considered above (Fig. 3b) does not yet fall into this category, since its state of rest is stable. Although the body can move with a velocity corresponding to the first plateau in the absence of the mean force, initial impetus must be given to put it into motion.

For a controllable machine, the possibility of completely removing the zeroth-order plateau should be provided. Under such conditions, the state of rest would become unstable and the machine would be put in motion spontaneously. However, the direction of motion in our symmetric system would depend on initial fluctuations and would not be uniquely determinate.

For a completely controllable machine, there must be provided the possibility of independently varying the widths of the plateaus corresponding to counter motions so that the motion be stable in a desired direction. Below, we will show that such controllability can be realized by imposing periodic forces acting in mutually orthogonal directions.

First, we will show that the zeroth-order plateau (and hence the static friction force) can be eliminated by varying the amplitude of an external periodic effect. To do this, we consider again the effect described by Eq. (8) for the case when the perturbation is not small. Mean force (11) can be calculated analytically with the following expansions [12]:

$$\cos(\psi\sin\omega t) = J_0(\psi) + 2\sum_{k=1}^{\infty}(-1)^k J_{2k}(\psi)\cos(2k\omega t), \tag{16}$$

$$\sin(\psi\cos\omega t) = 2\sum_{k=0}^{\infty}(-1)^k J_{2k+1}(\psi)\cos((2k+1)\omega t), \tag{17}$$

where $J_n(\psi)$ is the nth-order Jacobian function and

$$\psi = \frac{2\pi v_1}{a\omega}. \tag{18}$$

This mean force is identically zero at any velocity except for those satisfying the condition

$$v_0 = \frac{\omega a}{2\pi} = V_0 n, \tag{19}$$

where V_0 is the velocity corresponding to the first plateau and n is an integer.

For velocities defined by Eq. (19), the mean force is given by

$$\langle F\rangle = N(-1)^n J_n(v_1/V_0)\sin\frac{2\pi}{a}x_0. \tag{20}$$

For each of discrete velocities (19), the force depends on the initial coordinate and thus can take any value in the interval specifying the width of the related plateau:

$$-|NJ_n(v_1/V_0)| \le \langle F\rangle \le |NJ_n(v_1/V_0)|. \tag{21}$$

The dependences of the plateau width on the amplitude of velocity periodic oscillations are presented in Fig. 4. As the oscillation period grows, the width of the zeroth-order plateau is seen to decrease and vanish at a certain amplitude. This means that the static friction force disappears and any force as small as desired may remove the system from rest, although the width of the first-order plateau is still close to the maximum.

The numerical simulation shows that setting an oscillating force (instead of an oscillating velocity) changes the properties of the friction law insignificantly. Consider a body subjected to a constant propul-

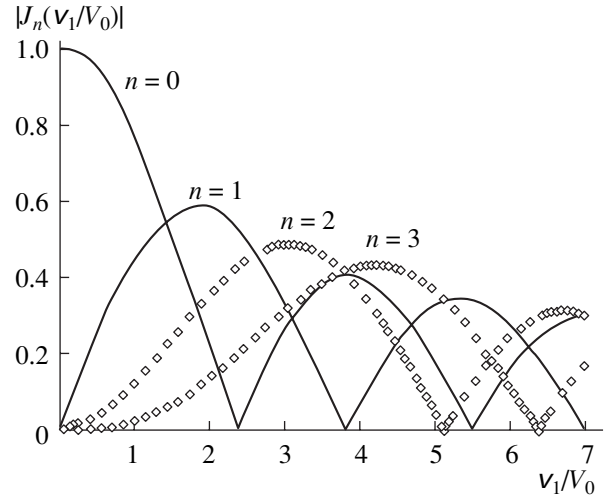


Fig. 4. Width of the zeroth-, first, second-, and third-order plateaus vs. velocity oscillation amplitude. As the oscillation amplitude grows, the zeroth-order plateau shrinks and disappears at $V_1/V_0 = 2.4$.

sive force F_0 on which a periodically oscillating force is imposed. The equation of motion has the form

$$m\ddot{x} = F_0 + F_1 \sin\omega t - \eta\dot{x} - N\sin(2\pi x/a). \qquad (22)$$

A typical dependence of the mean velocity of motion on the mean applied force is illustrated in Fig. 5a, which shows the result of numerical simulation of Eq. (22) for $m = 1$, $F_1 = 1.8$, $\omega = 1$, $\eta = 0.4$, and $2\pi/a = 1$. In this case, too, one can eliminate the zeroth-order plateau by varying the force oscillation amplitude. Figure 5b demonstrates the same as in Fig. 5a but for $F_1 = 2.3$. The amplitude of the oscillating component of the force was selected so as to remove the zeroth-order plateau. Here, the static friction force in the system is absent; however, the body can equiprobably move in the positive or negative direction.

It is easy to check that constant-velocity plateaus also appear if the amplitude of the periodic potential varies periodically. In actual systems, the amplitude of the periodic potential can be changed, e.g., by changing normal pressure. Therefore, a change in the oscillation amplitude will hereafter be referred to as "the effect of normal force" for the sake of clarity. The actual mechanism behind the change in the potential amplitude is of minor significance. Consider a body moving with a constant velocity $v = V_0$ in a potential field with the time-oscillating amplitude

$$N = N_0 + N_1\cos\omega t. \qquad (23)$$

In this case, the force acting on the body (with the dissipative force neglected) is given by

$$F = (N_0 + N_1\cos\omega t)\sin\frac{2\pi}{a}(x_0 + V_0 t). \qquad (24)$$

In calculating the time-averaged force, three cases should be considered.

(1) If the mean velocity of motion equals zero, $V_0 = 0$, the mean force is

$$F = N_0 \sin\frac{2\pi}{a}x_0. \qquad (25)$$

It can take any value from the interval $-|N_0| < \langle F\rangle < |N_0|$.

(2) If the mean velocity is other than zero, $V_0 \neq 0$, and the condition $V_0 \neq a\omega/2\pi$ is fulfilled, the mean force is identically zero:

$$\langle F\rangle = 0. \qquad (26)$$

(3) Finally, if the velocity is $V_0 = a\omega/2\pi$, the mean value of the first term in (24) equals zero and the second term gives a nonzero value of the mean force:

$$\langle F\rangle = \frac{N_1}{2}\sin\frac{2\pi}{a}x_0. \qquad (27)$$

The force depends on the initial coordinate and can take any value from the interval

$$-\left|\frac{N_1}{2}\right| < \langle F\rangle < \left|\frac{N_1}{2}\right|.$$

Below, it will be shown that one can provide the asymmetry of the friction law and thus the complete controllability of the system by simultaneously applying oscillating perturbations in the tangential and normal directions. We will again start with the simple situation, i.e., specify an oscillating velocity and calculate the time-averaged force. In this case, the force acting on the body is given by

$$F = \eta(V_0 + V_1\cos\omega t) + (N_0 + N_1\cos(\omega t + \varphi_0)) \times \sin\frac{2\pi}{a}\left(x_0 + V_0 t + \frac{V_1}{\omega}\sin\omega t\right). \qquad (28)$$

Its mean value can be calculated with identities (16) and (17). Analytical expressions can be derived for a plateau of any order. In particular, the half-width of the zeroth-order plateau is expressed as

$$\Delta F = \sqrt{N_0^2 J_0^2(\psi) + N_1^2 J_1^2(\psi)\sin^2\varphi_0}. \qquad (29)$$

It is easy to show that the zeroth-order plateau, together with the static friction force, disappears if

$$\psi = 2.4, \quad \varphi_0 = \{0 \quad \text{or} \quad \pi\}. \qquad (30)$$

The half-widths of the first- and minus-first-order plateaus (Fig. 6a) are given by

$$\Delta F_1 = |0.520N_0 \mp 0.215N_1|, \qquad (31)$$

$$\Delta F_2 = |0.520N_0 \pm 0.215N_1|, \qquad (32)$$

where the upper sign corresponds to $\varphi_0 = 0$ and the lower, to $\varphi_0 = \pi$.

With dissipation taken into account, the first- and minus-first-order plateaus shift in opposite directions by the same distance and the friction law takes the form schematically shown in Fig. 6b. An optimal machine has an attenuation at which the left edge of the first-order plateau lies exactly over the right edge of the minus-first-order plateau (or vice versa) (Fig. 6b). In this case, the body will move in the positive direction both in the absence of the mean external force and under the action of any oppositely directed force not exceeding

$$F_0 = \left|\frac{\Delta F_2 - \Delta F_1}{2}\right|.$$

This force is the maximal propulsive force this system can exert and, when measured in units of N_0, takes the form

$$\frac{F_0}{N_0} = \frac{1}{2}\left(\left|0.520 + 0.215\frac{N_1}{N_0}\right| - \left|0.520 - 0.215\frac{N_1}{N_0}\right|\right). \quad (33)$$

It depends only on the dimensionless amplitude N_1/N_0 of normal oscillations. The maximal possible propulsive force ($0.52N_0$) is achieved at an oscillation amplitude $N_1/N_0 \approx 2.4$ and remains unchanged when the oscillation amplitude grows further. However, the normal oscillation amplitude exceeding the mean normal force is in most cases unrealistic; therefore, the maximal propulsive force in this system can actually reach $0.22N_0$.

The attenuation coefficient of an optimal machine is given by

$$\eta = \frac{\Delta F_1 + \Delta F_2}{2V_0} = \pi\frac{|0.520N_0 + 0.215N_1| + |0.520N_0 - 0.215N_1|}{\omega a}. \quad (34)$$

The development of ways of controlling attenuation is an important problem in creating and optimizing controllable nanomachines.[4]

Finally, let us calculate the maximal efficiency of our machine. The efficiency is defined as the ratio of the useful power F_0V_0 to the total power dissipated in the system. The mean power of friction forces with parameters given by (30) is

$$\langle vF\rangle = -v_0[0.520N_0 + 0.215N_1]. \quad (35)$$

[4] The fundamental possibility of controlling dissipating forces follows from the following example. Imagine that a nanomachine is implemented on the surface of a semiconductor with an inversion layer. Applying a constant bias to the layer, one can vary (control) its conductivity and thus the dissipative forces acting on adatoms.

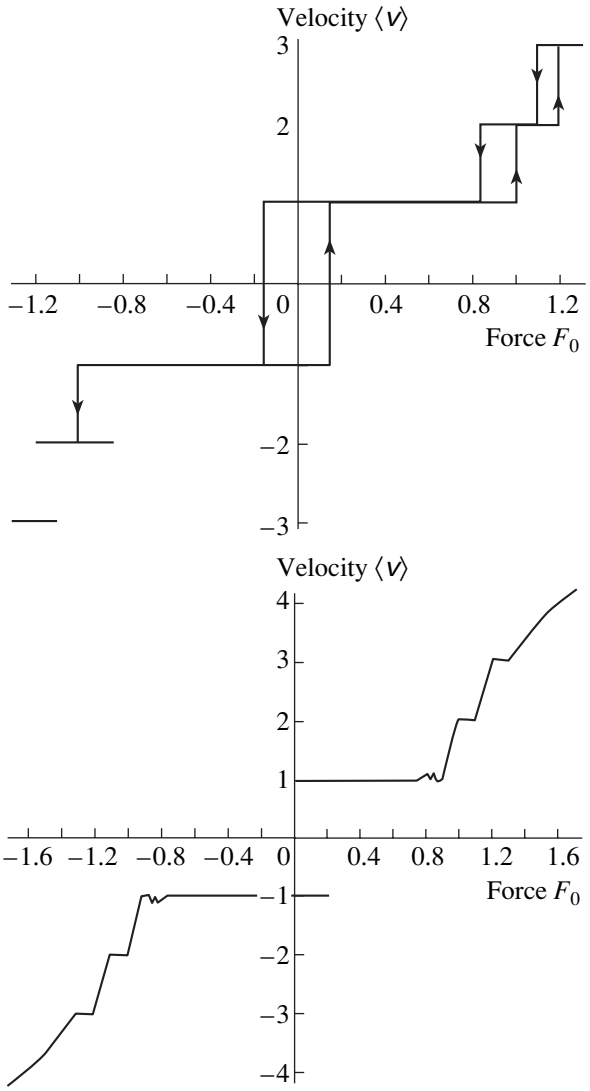


Fig. 5. Friction law in the presence of a periodically oscillating force derived by the numerical solution of Eq. (22) for $m = 1$, $\omega = 1$, $\eta = 0.4$, and $2\pi/a = 1$ and $F = 1.8$ (a) and 2.3 (b). (a) Plateaus of various orders partially overlap, giving rise to hysteresis when the force slowly rises and declines (the variation of the system state is displaced by lines with an arrow). (b) Oscillation amplitude is selected such that the zeroth-order plateau is absent. At the same time, the widths of other plateaus changed so that bistability is observed only near the first- and minus-first-order plateaus.

For the efficiency, we get

$$\kappa = \frac{\frac{1}{2}\left(\left|0.520 + 0.215\frac{N_1}{N_0}\right| - \left|0.520 - 0.215\frac{N_1}{N_0}\right|\right)}{\left[0.520 + 0.215\frac{N_1}{N_0}\right]}. \quad (36)$$

As the oscillation amplitude grows, the efficiency increases, reaches a maximum of 1/2 at $N_1/N_0 \approx 2.4$, and then drops. For a maximal realistic amplitude of normal oscillation (on the order of N_0), the efficiency is about 0.3.

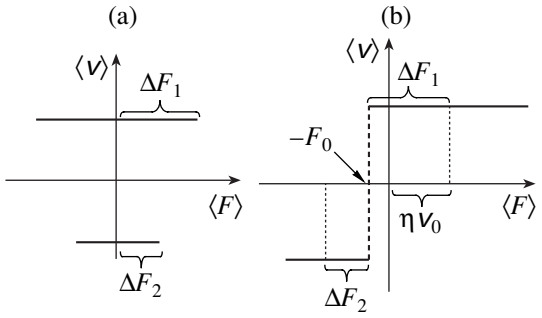


Fig. 6. Schematic representation of a constant-velocity plateau when the velocity and potential amplitude oscillate simultaneously (a) without attenuation and (b) with optimal attenuation. In the general case, the widths of the first- and minus-first-order plateaus (which correspond to opposing motions) differ.

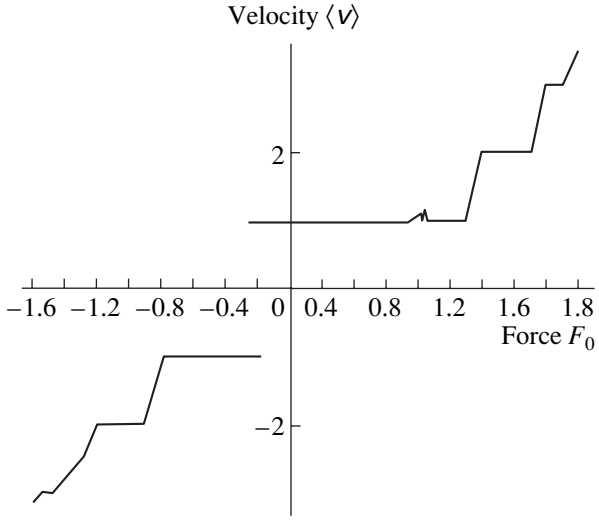


Fig. 7. Friction law for the simultaneous periodical oscillation of the force and potential amplitude derived by the numerical solution of Eq. (37) at $m = 1$, $F_1 = 2.3$, $\omega = 1$, $N_0 = 1$, $N_1 = 0.6$, $\eta = 0.6$, $\varphi_0 = 2.5$, and $2\pi/a = 1$. The parameters are selected such that the zeroth-order plateau is absent.

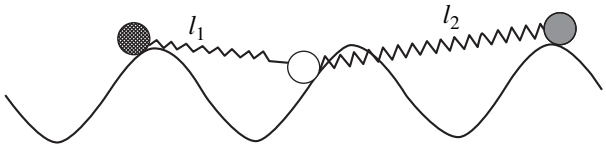


Fig. 8. Three-body nanomachine. The bodies are linked together by active components whose equilibrium length can be controlled by external perturbations.

The basic properties of the friction law are also retained if a periodic force is specified and the mean velocity is calculated. Figure 7 illustrates the results of numerical solution of the equation

$$m\ddot{x} = F_0 + F_1\sin\omega t - \eta\dot{x} - (N_0 + N_1\sin(\omega t + \varphi_0))\sin(2\pi x/a), \quad (37)$$

which describes the motion under the action of tangential and normal oscillating forces. The oscillation amplitude and the phase shift were selected so as to remove the zeroth-order plateau. Here, the first- and minus-first-order plateaus are asymmetric; as a result, in the absence of the mean applied force (or at weak negative forces), the body can execute a stable macroscopic motion only in the positive direction. By varying the phase shift with all other parameters being the same, one can "reverse the roles" of these two plateaus so that the spontaneous motion will take place only in the negative direction.

4. MANY-BODY NANOMACHINES

When applying the above ideas to designing nanomachines, one should bear in mind that the above-mentioned techniques of periodic actions on the system are not the only approach to provide a friction law with a controllable constant-velocity plateau. In essence, the oscillation of any parameter influencing the interaction with the periodic potential causes a constant-velocity plateau in the mean velocity vs. mean force dependence and can basically be used for creating systems developing a propulsive force. Let us illustrate the aforesaid by an example of the nanomachine described in [5].

Consider a system of three bodies in a periodic potential (Fig. 8). The bodies interact with each other via springs whose equilibrium length can be varied by any external perturbations.[5] The equations of motion for the bodies involved in the system have the form

$$m\ddot{x}_1 + \eta\dot{x}_1 + N_0\sin x_1 = k(x_2 - x_1 - l_1), \quad (38)$$
$$m\ddot{x}_2 + \eta\dot{x}_2 + N_0\sin x_2 = k(x_3 - x_2 - l_2) - k(x_2 - x_1 - l_1),$$

$$m\ddot{x}_3 + \eta\dot{x}_3 + N_0\sin x_3 = k(x_3 - x_2 - l_2), \quad (39)$$
$$l_1 = l_0 + \delta l\sin(\omega t), \quad l_2 = l_0 + \delta l\sin(\omega t + \varphi_0).$$

In this case, too, the mean velocity vs. mean force dependence shows well-defined constant-velocity plateaus. Figure 9 demonstrates the result of the numerical experiment where the mean force acting on the system slowly grew starting from zero. Here, only the time dependence for the coordinate of the middle body is presented. If the force is relatively weak, the group of the bodies remains at rest and, accordingly, the system is on the zeroth-order plateau. As soon as the propulsive force exceeds some threshold value, the system starts executing a macroscopic motion. However, it moves with a constant force-independent macroscopic velocity corresponding to the first-order plateau. At a still greater force, the system jumps to the second-order plateau and moves with a new (discrete) velocity for a time, and so on.

[5] In the case of microsystems, this can be achieved, for example, by using materials exhibiting the electro- or magnetostriction effect; for molecular objects, chromophore-containing molecules, which change their equilibrium length when exposed to a radiation of particular wavelength, can serve as active springs.

Figure 10 illustrates the time dependence for the coordinates of each of the three bodies. With a fairly good accuracy, the motion of each of the bodies can be represented as the superposition of uniform and periodic harmonic motions, the center of gravity of the system as a whole oscillating much more weakly than each of the bodies individually. Therefore, to treat the situation qualitatively, it appears to be reasonable to proceed from a model where the bodies move in such a way that their center of gravity executes a uniform translational motion. Such an assumption can be substantiated by the fact that, as we will see later, the optimal operation of the machine takes place at comparatively large amplitudes of molecule oscillations. Therefore, in the first approximation, the periodic potential can be viewed as a perturbation imposed on the motion of the system of the bodies. In this approximation, the motion of the bodies is described by the following equations:

$$x_1 = x_0 - l_0 + v_0 t - \frac{2}{3}\Delta l \sin\omega t - \frac{1}{3}\Delta l \sin(\omega t + \varphi_0),$$

$$x_2 = x_0 + v_0 t - \frac{1}{3}\Delta l \sin(\omega t + \varphi_0) + \frac{1}{3}\Delta l \sin\omega t, \quad (40)$$

$$x_3 = x_0 + l_0 + v_0 t + \frac{2}{3}\Delta l \sin(\omega t + \varphi_0) + \frac{1}{3}\Delta l \sin\omega t.$$

These equations can be recast as

$$x_1 = x_0 - l_0 + v_0 t - \frac{\Delta l}{3}\sqrt{5 + 4\cos\varphi_0}\sin(\omega t + \varphi_1),$$

$$\tan\varphi_1 = \frac{\sin\varphi_0}{2 + \cos\varphi_0},$$

$$x_2 = x_0 + v_0 t + \frac{\Delta l}{3}\sqrt{2(1 - \cos\varphi_0)}\sin(\omega t - \varphi_2), \quad (41)$$

$$\tan\varphi_2 = \frac{\sin\varphi_0}{1 - \cos\varphi_0},$$

$$x_3 = x_0 + l_0 + v_0 t + \frac{\Delta l}{3}\sqrt{5 + 4\cos\varphi_0}\sin(\omega t + \varphi_3),$$

$$\tan\varphi_3 = \frac{2\sin\varphi_0}{1 + 2\cos\varphi_0}.$$

For the mean value of the sum of forces acting on the bodies (on the zeroth-order plateau $n = 0$ and $v_0 = 0$), we have

$$\left\langle \frac{F}{N_0} \right\rangle = \sin k_0 x_0 \left[2\cos k_0 l_0 J_0\left(\frac{k_0 \Delta l}{3}\sqrt{5 + 4\cos\varphi_0}\right) + J_0\left(\frac{k_0 \Delta l}{3}\sqrt{2(1 - \cos\varphi_0)}\right)\right], \quad (42)$$

where $k_0 = 2\pi/a$.

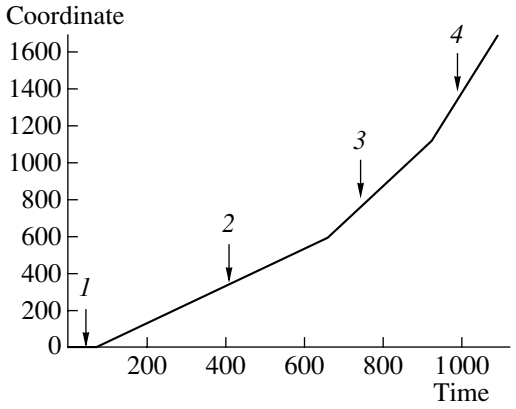


Fig. 9. Time dependence of the middle body coordinate for the three body-machine obtained by the numerical solution of Eqs. (38) and (39) for $m = 1$, $\omega = 1$, $F_0 = 0$, $\varphi_0 = \pi$, $l_0 = 3\pi$, $\delta l = \pi$, and $\eta = 0.4$. The force slowly grew from zero to the maximal value. (*1*) Zeroth-order ($n = 0$), (*2*) first-order ($n = 1$), (*3*) second-order ($n = 2$), and (*4*) third-order ($n = 3$) plateaus.

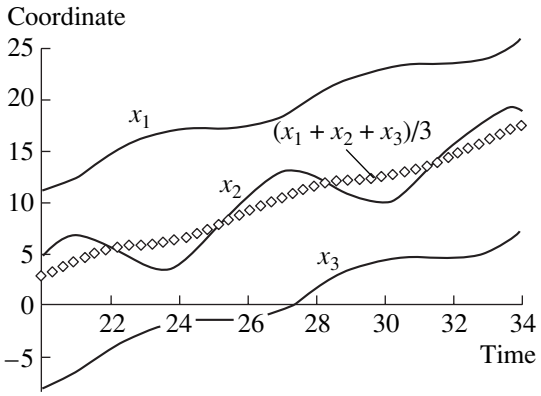


Fig. 10. Time dependence of the coordinates of the bodies and the center of gravity for the three-body nanomachine. In the first approximation, it can be assumed that the center of gravity moves with a constant velocity.

For the operation of the machine to be stable, it is necessary that the width of the zeroth-order plateau vanish. With the condition

$$\cos k_0 l_0 = 0, \quad l_0 = a(1/4 + n/2), \quad (43)$$

satisfied, the first term in (42) vanishes identically (at any Δl and φ_0) and the vanishing condition for the width of the zeroth-order plateau is reduced to vanishing the second term in (42):

$$J_0\left(\frac{k_0 \Delta l}{3}\sqrt{2(1 - \cos\varphi_0)}\right) = 0, \quad (44)$$

hence,

$$\frac{k_0 \Delta l}{3}\sqrt{2(1 - \cos\varphi_0)} = 2.4. \quad (45)$$

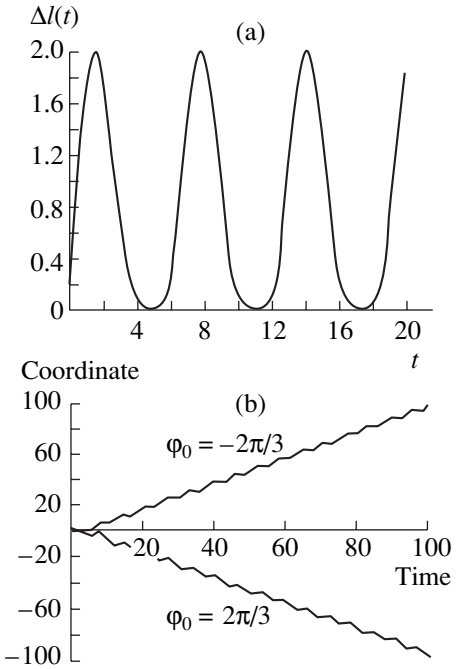


Fig. 11. (a) Time dependence of the link equilibrium lengths in the model stated by (38) and (39) and (b) the motion of the nanomachine (middle body coordinate vs. time) for two values of the phase shift between link length oscillations. The change in the phase shift reverses the direction of motion. The numerical parameters of the system are the same as in Fig. 9.

We thus see that the machine of the given design functions in a stable mode only if

$$\Delta l > 0.573a, \tag{46}$$

where a is the period of the potential.

The width of the zeroth-order plateau may also vanish if condition (43) is violated. In this case, however, the phase shift φ_0 and the change in the length of the links (springs) Δl turn out to be strictly fixed. Indeed, both Jacobian functions in (42) must vanish, which implies the simultaneous fulfillment of the conditions

$$\frac{k_0\Delta l}{3}\sqrt{5 + 4\cos\varphi_0} = 2.4, \tag{47}$$

$$\frac{k_0\Delta l}{3}\sqrt{2(1 - \cos\varphi_0)} = 2.4. \tag{48}$$

This is possible only if

$$\cos\varphi_0 = -0.5 \tag{49}$$

and

$$k_0\Delta l = 2.4\sqrt{3}, \quad \Delta l \approx 0.66a. \tag{50}$$

The equilibrium length of the links remains arbitrary and can be selected so as to provide the maximum propulsive force.

It should be noted that both the analytical calculation and a variety of numerical experiments indicate that a three-body machine retains the symmetry at any phase shift between the link length oscillations if a perturbation imposed on the system is strictly harmonic. Because of this, we cannot provide the stable motion of the machine in a desired direction. The symmetry breaks if one uses, for example, a pulsed, rather than harmonic, perturbation (see Fig. 11a). In this case, the direction of motion can be controlled by varying the phase shift (Fig. 11b).

The nanomachines considered in Sections 3 and 4 share the common property, namely, the presence of at least two independent periodic actions (perturbations). Clearly, this is the general requirement following from symmetry considerations. In fact, the imposition of a single periodic perturbation, while causing a constant-velocity plateau, does not break the system symmetry with respect to the motion in the positive and negative directions. The presence of the second perturbation with a fixed phase shift relative to the first one breaks this symmetry and makes motion in a particular direction possible.

5. PROSPECTS AND APPLICATIONS

An obvious domain of applicability of the concepts put forward above is handling the motion of molecules or a molecular ensemble on crystal surfaces or the motion of nanoobjects on profiled solid surfaces. A challenge to be tackled upon implementing nanomachines is the development of microactuators allowing for the generation of periodic perturbations of desired frequency and amplitude.

In the case of integrated circuits, these may be components offering electro- and magnetostrictive properties. For example, if one of the links in a three-body machine (see Section 4) is made of electrostrictive and the other of magnetostrictive materials, one can control either of the links independently by simultaneously applying electric and magnetic fields. This will suffice to put the machine in motion in a desired direction.

For molecular clusters, chromophore-containing molecules, which change their geometry when exposed to a radiation of specific wavelength, can serve as active components. The use of two different chromophores makes it possible to independently control the length of either of the links.

Another important application of the effect of constant-velocity plateau is the high-precision control of the rate of actuator response in various mechanotronic systems [14]. As follows from (19), “quantized” velocity values depend only on the external effect frequency and the potential relief wavelength. Therefore, these values can be set with a high accuracy and are easy to control by varying the frequency.

An experimental demonstration of this effect could be the contact between two artificially patterned plates

that have a periodic (relative to translations or rotations) microrelief. A periodic perturbation in this case can be generated by an ultrasonic technique.

The concepts discussed in this paper are also of scientific value, allowing a better understanding of the physics of tribological phenomena. Apparently, the conventional treatment of dry friction forces with the static and sliding friction coefficients is adequate to describe only quasi-stationary processes. A friction law derived from quasi-static measurements fails when applied to situations where high-frequency effects are dealt with (for example, to traveling-wave engines [15]). Measuring the force of friction between bodies subjected to the superposition of constant and oscillating forces as a function of the amplitude and frequency of the periodic perturbation would be much more informative. Such measurements would allow researchers to gain direct information on characteristic time–space scales "responsible" for friction in the system under consideration and could serve as a peculiar kind of the "spectroscopy" of tribosystems.

Obviously, further investigation into artificial (profiled) tribosystems with active components will lay the foundation for creating controllable tribosystems.

CONCLUSION

In this work, we put forth the general principles of inducing the directed motion of objects in a periodic potential field on which periodic external actions are imposed. These principles are scale-invariant, that is, can be implemented at any scale level. We showed that periodically oscillating applied forces or parameters of interaction between the body and the periodic external action can cause the spontaneous directed motion of objects with a velocity uniquely defined by the frequency of the periodic action and the space period of the potential field.

By way of example, we considered a body subjected to periodic forces in the tangential and normal directions, as well as a group of bodies linked by two active components whose length periodically oscillates. The basic principles of generating directed motion and the methods of mathematical analysis used are applicable to any systems where some parameters specifying the system dynamics oscillate in time. Identical properties are offered, for example, by a system of two bodies that have one active link with an oscillating length when an extra periodic (tangential or normal) force is applied. The number of possible designs of nanomachines can readily be increased. Their implementation is by no means limited by the designs considered above: it must fit currently available techniques of periodically varying the parameters of nanosystems.

ACKNOWLEDGMENTS

The author thanks K.-H. Antony and O.K. Dudko for valuable discussions, as well as the German Academic Exchange Service for financial support.

REFERENCES

1. K. E. Drexler, *Nanosystems: Molecular Machinery, Manufacturing and Computation* (Wiley, New York, 1992).
2. F. Jülicher, A. Ajdary, and J. Prost, Rev. Mod. Phys. **69**, 1269 (1997).
3. P. Reimann, Phys. Rep. **361**, 57 (2002).
4. R. D. Astumian, Science **276**, 917 (1997).
5. M. Porto, M. Urbakh, and J. Klafter, Phys. Rev. Lett. **84**, 6058 (2000).
6. I. I. Blekhman and G. Yu. Dzhanelidze, *Vibration Displacement* (Nauka, Moscow, 1964).
7. G. A. Tomlinson, Philos. Mag. **7**, 905 (1929).
8. B. N. J. Persson, *Sliding Friction. Physical Principles and Applications* (Springer-Verlag, New York, 2000, 2nd ed.).
9. V. Zaloj, M. Urbakh, and J. Klafter, Phys. Rev. Lett. **81** (6), 1227 (1998).
10. V. Zaloj, M. Urbakh, and J. Klafter, Phys. Rev. Lett. **82** (24), 4823 (1999).
11. V. L. Popov, Solid State Commun. **115**, 369 (2000).
12. A. Barone and G. Paterno, *Physics and Applications of the Josephson Effect* (Wiley, New York, 1982; Mir, Moscow, 1984).
13. H. Risken, *The Fokker Planck Equation* (Springer-Verlag, Berlin, 1989, 2nd ed.).
14. V. L. Popov, Pis'ma Zh. Tekh. Fiz. **27** (13), 50 (2001) [Tech. Phys. Lett. **27**, 551 (2001)].
15. J. Wallaschek, Smart Mater. Struct. **7**, 369 (1998).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1408–1412. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 64–68.

SOLID-STATE ELECTRONICS

Error Estimation in Recording *I–V* Characteristics of Josephson Junctions

S. I. Borovitskiĭ, V. D. Gelikonova, A. V. Komkov, Kh. A. Aĭnitdinov†, and A. M. Klushin

Kvarts Institute of Electronic Measurements, Nizhni Novgorod, 603009 Russia

e-mail: cel@sandy.ru

Received February 20, 2001; in final form, February 11, 2002

Abstract—A method for characterizing a current total step in the *I–V* characteristic of a Josephson junction array is considered. In this method, an appropriately selected approximating curve is statistically fitted to the experimentally found curve. A self-calibration algorithm for an array of junctions incorporated into programmable voltage standards is suggested. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In recent years, interest has arisen in programmable voltage standards based on Josephson junction arrays [1]. With such standards, it becomes possible to produce a set of quantized reference voltages from 0 to 10 V and simplify the calibration of precision ADCs and digital voltmeters. Accordingly, the calibration scheme for these devices radically changes, and the certification speed and accuracy are greatly improved. To produce a set of precision reference voltages, an array of a large number m of series-connected nonhysteresis Josephson junctions is used. One of the working algorithms for programmable voltage standards uses the quantized reference voltage across a junction array at the first step with index $n = \pm 1$. The total voltage is then $V_J = (m_1 - m_2)f/K_J$ [2]. Here, m_1 is the number of junctions at the step $n = 1$, m_2 is that at the step $n = -1$, f is the frequency of an applied harmonic signal, and $K_J = 483.5979$ GHz/V is the Josephson constant. It should be emphasized that this algorithm allows for the application of Josephson junctions on high-temperature superconductors with a great (≈100%) spread in Josephson critical currents $\delta I = I_{c,\,max}/I_{c,\,min}$, where $I_{c,\,max}$ and $I_{c,\,min}$ are, respectively, the maximal and minimal critical currents of the junctions incorporated into the array [3, 4]. At $f \geq K_J I_{c,\,max} R_N$ (R_N is the normal resistance of the junctions), the microwave signal power can be selected such that the amplitude of the first total current step ΔI_1 is on the order of $I_{c,\,min}$. Actually, ΔI_1 decreases because of a spread in R_N and the nonuniform distribution of the microwave current along the array. A recent investigation of a Josephson programmable standard based on niobium nonhysteresis junctions has shown that, after each cooling of the junctions, a new distribution of the microwave power along the array arises and the system must be tuned to the optimal frequency of irradiation [5]. At low $I_{c,\,min}$ and/or elevated temperatures (≈78 K), thermal noise may additionally decrease ΔI_1 and accordingly increase the inaccuracy of the voltage standard [4, 6]. Because of this, the control of the slope and other parameters of the current total step is of great importance.

In this work, we suggest a possible way of characterizing current total steps appearing in the *I–V* characteristic of a Josephson junction array and discuss a self-calibration algorithm for the array. The feasibility of the method is demonstrated with the critical current of a bicrystal Josephson junction made of a high-temperature semiconductor.

PRECISE METHOD FOR MEASURING THE CURRENT STEP PARAMETERS

In the conventional technique for finding the current step parameters (specifically, its slope), the voltage V_J of the step is successively measured at several points ($i \approx 10$) the number of which depends on the constant bias current passing through the array. To improve the measurement accuracy, the voltage to be measured first is shifted to the zero level by inserting a known voltage of opposite polarity and second is measured N times at each ith point [7]. The compensation of V_J allows one to perform measurements with a minimal standard deviation σ depending on the intrinsic noise of a nanovoltmeter. In the best digital nanovoltmeters, $\sigma \approx 10$ nV; in analog devices, $\sigma \leq 1$ nV. As the number of independent measurements grows, the standard deviation of the mean value decreases as $\sigma_m = \sigma/(N-1)^{-1/2}$. For the time of single-point measurement $\tau = 100$–1000 s, the value of σ_m can be reduced to ≤0.1 nV [8]. However, small-σ_m measurements at all i points take a considerable amount time (about 10τ), during which the stability of the thermoelectric voltage in the measuring circuit is hard to maintain. In this case, the drift of the thermal emf will make a major contribution to the total standard

† Deceased.

deviation of the measurements. There are a number of methods taking into account the thermal emf drift [7]. However, their use is associated with a further increase in the number of measurements and, hence, elongates the measuring cycle.

The essence of our method is as follows. First, for a time $\approx\tau$, the *I–V* characteristic $V_k = f(I_k)$ is recorded in the range of a step of the current (or critical current) being studied, where the subscript k runs from 1 to N. Then, the data array obtained is approximated by the mathematical expression $V = f(I, p_1, p_2, p_3, \ldots)$, where V is the voltage across the junction; I is the bias current; and p_1, p_2, p_3, … are parameters of the theoretical model. The mean values of the parameters and their standard deviations are found by the least squares method. Thus, the problem is reduced to finding an approximating curve that adequately fits experimental data.

As a curve approximating the *I–V* characteristic of a Josephson junction near a step of the current (or I_c), we take the expression

$$V = V_0 + r(I - i_0) + \frac{r_d}{i_1 + i_2}\left(i_1^2 e^{(I-i_0)/i_1} - i_2^2 e^{-(I-i_0)/i_2}\right). \quad (1)$$

Expression (1) involves six parameters. The parameter V_0 specifies the position of the current step on the voltage axis and tends toward V_J in the optimum case. The factor r reflects the presence of parasitics in the measuring circuit. The study of this factor can shed light on the value and nature of the parasitics and help in eliminating them or reducing to a reasonable level. Approximating curve (1) has an inflection where the current and the differential resistance equal, respectively, i_0 and r_d. The parameter i_0 defines the position of the current step midpoint on the current axis. The parameters i_1 and i_2 have the dimension of current and specify the bend (smooth or sharp) of the *I–V* curve at the edges of the steps.

When selecting the form of formula (1), we took into consideration that junctions with the nonhysteresis *I–V* curve may exhibit one-particle tunneling of magnetic quanta at the edges of I_c, causing the exponential current dependence of the voltage, and that ranges where this takes place are much smaller than the critical current. From the theory of the Josephson effect [9], it is known that the current–voltage characteristic near the step has a similar form. However, in analyzing data for current total steps, one should bear in mind that the parameters i_1 and i_2 involved in the exponential dependence depend largely on the spread in R_N of the junctions that are series-connected to form an array and on the nonuniformity of the microwave field along the array, rather than on the conduction mechanism. Formula (1) is a good fit to the characteristic only in the vicinity of the voltage V_0; at higher voltages, the *I–V* curve describes resistive regions, where expression (1) is invalid.

RESULTS AND DISCUSSION

The statistical fit of the approximating curve to experimental data implies the minimization of the total standard deviation

$$\chi = \sqrt{\frac{\sum_{k=1}^{N} (V_k - V(I_k))^2}{N - p}}, \quad (2)$$

where p is the number of desired parameters [10].

Let us consider sum (2) as a function of the number N of points and the range of fit. By way of example, we will look for the parameters of curve (1), which describes the *I–V* characteristic of a Josephson junction in the vicinity of the critical current.

We studied the *I–V* characteristic of a bicrystal Josephson junction based on a high-temperature superconductor [3, 6]. The characteristic was recorded according to the scheme shown in Fig. 1 with the program developed in [11]. As a voltage amplifier, a nanovoltmeter with an intrinsic noise of 1–2 nV and a gain of 5×10^4 was used. Such nanovoltmeters are incorporated into industrial Josephson voltage standards [12]. Figure 2 shows the record of the critical current and its approximation by function (1). The table summarizes the fitting parameters and their standard deviations corresponding to a confidence interval of 2.8χ. The probability of falling outside this interval is no more than 1%.

The parameter χ vs. number of points N. It is known that the minimizing procedure is aimed at finding the absolute, rather than relative, minimum of sum (2). To find the absolute minimum, one usually changes the

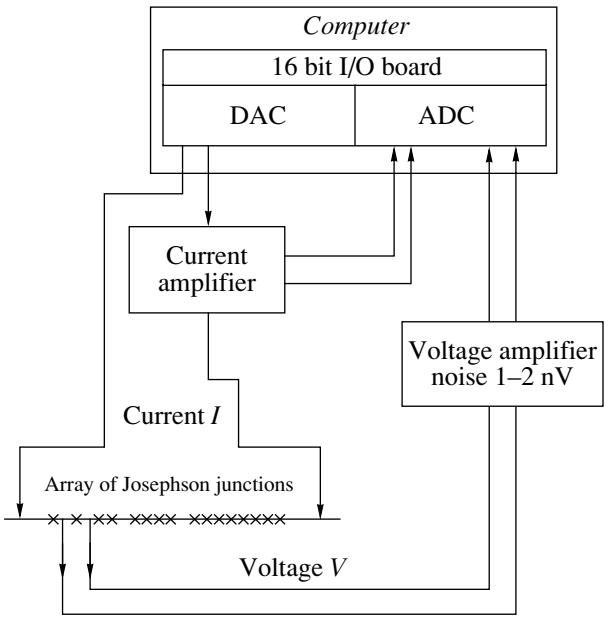


Fig. 1. Precision measurement of the current–voltage characteristic of Josephson junctions.

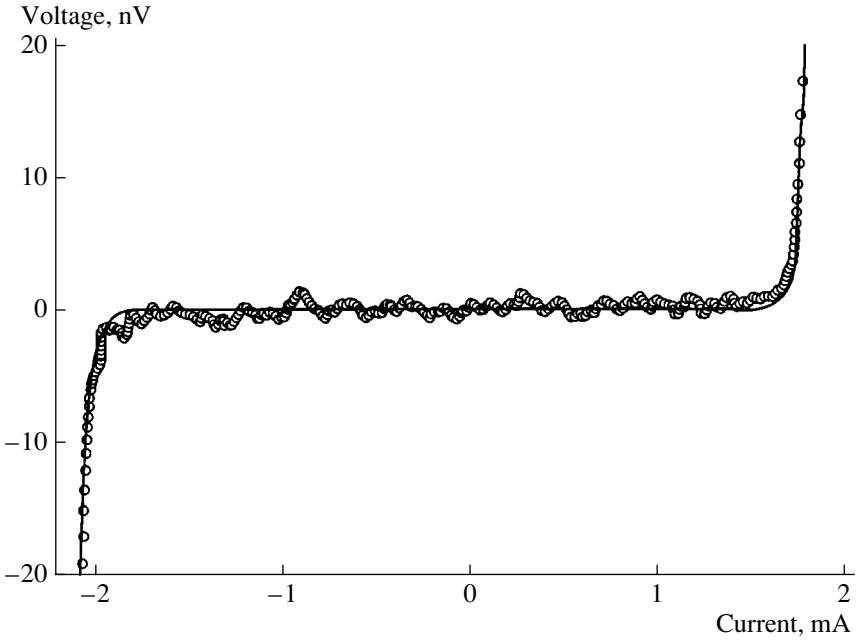


Fig. 2. *I–V* characteristic of a bicrystal Josephson junction near the critical current: ○, data points; solid line, fitting curve.

initial conditions of the process and monitors the value of χ and the parameters of the fitting curve. When the final result becomes independent of the initial conditions, it can be argued that the absolute minimum is found [10]. To check this statement, we studied the dependence of the approximation parameters and their standard deviations on the number of points used for the approximation. The number of points was varied from $N = 473$ to $N = 3786$. Figures 3 and 4 show the values of the parameters and their standard deviations normalized to the related means. As follows from Fig. 3, the approximation parameters depend on N only slightly. From Fig. 4, χ is seen to be independent of the number of points, whereas the normalized standard deviations of the parameters in expression (1) decrease as $N^{-1/2}$. These observations suggest that the absolute minimum of sum (2) is reached and the parameter means found provide the best fit to the experimental curve. Moreover, it can be argued that N-related systematic errors, if present, play an insignificant role.

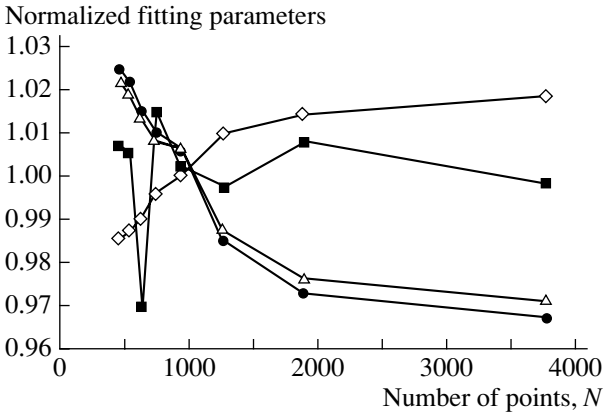


Fig. 3. (■), (◇), (●), and (△) are the normalized fitting parameters V_0, i_0, i_1, and i_2 vs. number of points N taking part in the approximation. Solid lines connect corresponding symbols.

The effect of the approximation voltage range on χ. Figure 5 shows the dependence of χ on the approximation voltage range. At voltages above 20–30 nV, the total standard deviation χ grows rapidly. Figure 6 demonstrates the differences $V_k - V(I_k)$ between the *I–V* curve recorded and the fitting curves with voltages of ±9, ±20, and ±160 nV, along with the experimental dependence $V_k(I_k)$. It is seen that at low voltages the approximation gives I_c higher than the values measured; at high voltages, the differences are due to the limited applicability domain of formula (1). Note also that the standard deviations of all the approximation parameters, except σ_{V_0}, grow near zero. From these dependences, one can conclude that the optimal range of approximation is from ±20 to ±30 nV in our case.

The contribution of the standard deviations of the parameters to the standard deviation of the voltage. The standard deviation of the voltage depends on the bias current as

$$\sigma_V^1(I) = \sqrt{\sum_{i=1}^{6}\left(\frac{\partial V(I)}{\partial p_i}\right)^2 \sigma p_i^2}. \qquad (3)$$

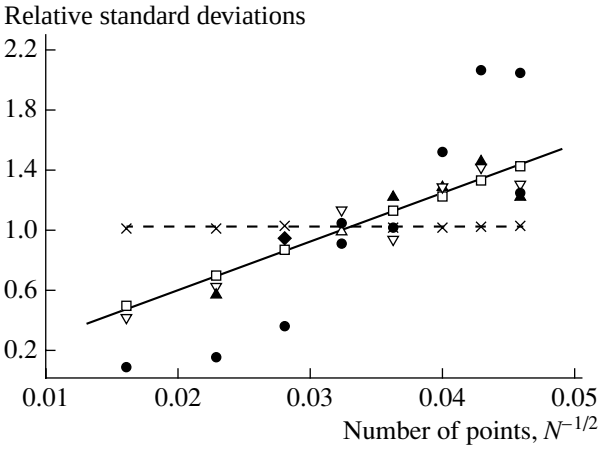


Fig. 4. Normalized standard deviations (□) σ_{V_0}, (◆) σ_{i_0}, (○) σ_{rd}, (▲) σ_{i_1}, (▽) σ_{i_2} and (×) normalized total standard deviation χ vs. number of points N participating in the approximation. The solid line represents the linear fit to the parameter standard deviations vs. N dependence.

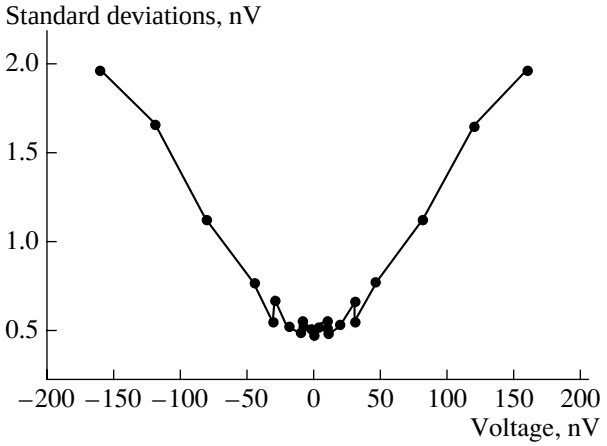


Fig. 5. Total standard deviation χ vs. fitting voltage range (●). The solid line connects data points.

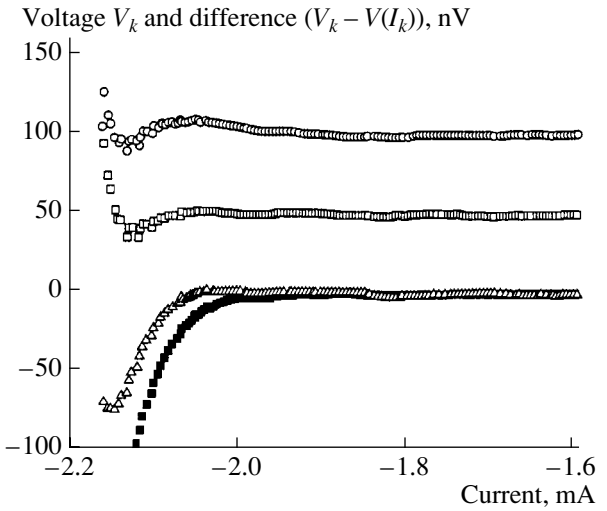


Fig. 6. Difference between the experimental I–V characteristic (■) and fitting curves for voltages in the range from (□) –9 to +9, (△) –20 to +20, and (○) –160 to +160 nV. For convenience, the differences calculated for the ranges ±9 and ±160 nV are shifted by 50 and 100 nV relative to the zero level.

The contributions of the standard deviations of the parameters to $\sigma_V^1(I)$ are different. Figure 7 shows the logarithm of the contribution of the parameter standard deviations to $\sigma_V^1(I)$ vs. bias current for a voltage range from –20 to 20 nV. From the curves in Fig. 7, it follows that, at currents equal to or higher than I_c, the major contribution to the total error is from $\sigma_{V_0}^1(I)$ and $\sigma_r^1(I) \approx \sigma_{V_0}^1(I)$. Thus, in this current range, $\sigma_r^1(I) \approx (2\sigma_{V_0}(I))^{1.2}$ and grows rapidly when the current exceeds I_c. The standard deviation σ_{V_0}, as well as all other standard deviations, decreases with increasing N. In our example (Fig. 2), it became 50 to 100 times lower than the intrinsic noise of the nanovoltmeter, not exceeding σ_{V_0} = 20 pV. The low mean value of the differential resistance r_d (see table) indicates the absence of the slope of critical current and demonstrates the potentialities of the method as applied to measuring low parasitic resistances.

SELF-CALIBRATION OF JUNCTION ARRAYS FOR A PROGRAMMABLE VOLTAGE STANDARD

The method suggested allows for the natural self-calibration of an array of 2^m junctions that is incorporated into a programmable voltage standard [2]. The array is subdivided into (m + 1) sections (bits). The number of junctions in the first m bits corresponds to the binary code (1, 2, 4, 8…). The last, (m + 1), bit has only one junction. During measurements, the total step must be made coincident with the zero voltage level to retain the desired accuracy. Therefore, the calibration of the entire array is carried out in several stages. First, the voltage of the step +1 of one junction is compared with that of the step –1 of the second junction. Then, these two junctions are connected in series and their

Fitting parameters and their standard deviations

Parameters	Units of measure	Parameter values	Standard deviations
V_0	nV	0.13	0.02
r	Ω	2.9×10^{-7}	2×10^{-8}
i_0	mA	0.08	0.05
r_d	Ω	7.6×10^{-21}	8.5×10^{-21}
i_1	mA	0.04	0.002
i_2	mA	0.05	0.002
χ	nV	0.41	

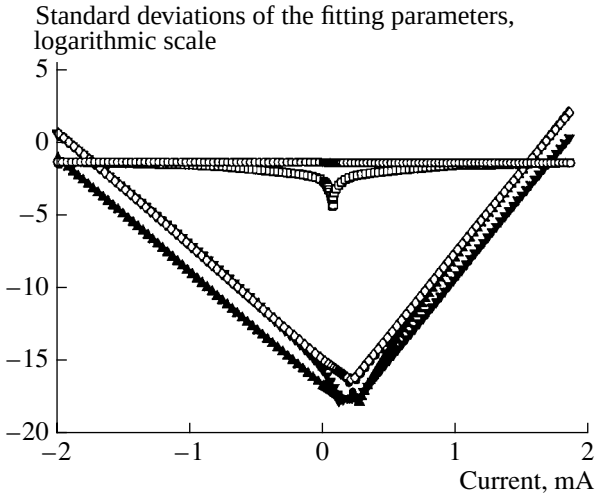


Fig. 7. Logarithmic contribution of the parameter standard deviations vs. bias current for voltages in the range ±20 nV. (○) $\log\sigma^1_{V_0}$, (□) $\log\sigma^1_r$, (◇) $\log\sigma^1_{i_0}$, (▼) $\log\sigma^1_{rd}$, (▲) $\log\sigma^1_{i_1}$, and (●) $\log\sigma^1_{i_2}$.

total voltage is compared with that of a section of two junctions that have opposite polarity, etc. Having established (with a high accuracy) the absence of the slope on the first steps of individual contacts, one can establish the absence of parasitic slopes on total steps in all the bits of the array. This algorithm allows one to automate the self-calibration of arrays of series-connected junctions.

CONCLUSION

We suggest a method for processing a body of experimental data for the *I–V* characteristics of Josephson junctions that can estimate the accuracy of measurements. For this purpose, it is necessary to select an analytical expression for the *I–V* characteristic that involves a small number of parameters. The choice can be based on both a theoretical model and intuitive concepts.

The desired mean values of the parameters and their standard deviations are found by the least squares method. This method combines the advantage of reducing the problem to finding a function depending on a small number of parameters and the advantages of a large number of measurements. A small value of χ found by fitting is the best corroboration of the potentialities of the method. Another verification of the method is the study of the dependence of the parameters on the number of points and the dependence of the results on the limitation of the approximation interval.

REFERENCES

1. C. A. Hamilton, Rev. Sci. Instrum. **71**, 3611 (2000).
2. A. M. Klushin, S. I. Borovitski, C. Weber, *et al.*, in *Proceedings of the 3rd European Conference on Applied Superconductivity (EUCAS'97), 1997*, Vol. 1, Inst. Phys. Conf. Ser., No. 158, p. 587.
3. A. M. Klushin, W. Prusseit, E. Sodtke, *et al.*, Appl. Phys. Lett. **69**, 1634 (1996).
4. S. I. Borovitskiĭ, A. M. Klushin, T. B. Korotina, *et al.*, Pis'ma Zh. Tekh. Fiz. **11**, 663 (1985) [Sov. Tech. Phys. Lett. **11**, 275 (1985)].
5. B. Jeanneret, A. Rüfenacht, and J. Burroughs, IEEE Trans. Instrum. Meas. **50**, 188 (2001).
6. A. M. Klushin, C. Weber, M. Darula, *et al.*, Supercond. Sci. Technol. **11**, 609 (1998).
7. R. Behr, H. Schulze, F. Müller, *et al.*, IEEE Trans. Instrum. Meas. **48**, 270 (1999).
8. T. J. Witt, IEEE Trans. Instrum. Meas. **50**, 445 (2001).
9. K. K. Likharev, *Introduction to the Dynamics of Josephson Junctions* (Nauka, Moscow, 1985), pp. 177–183.
10. *Guide for the Expression of Uncertainty in Measurement* (International Organization for Standardization, Switzerland, 1993), Annex H.
11. E. Goldobin, http://www.geocities.com/SiliconValley/Heights/7318/review.htm.
12. Kh. A. Aĭnitdinov, S. I. Borovitskiĭ, V. G. Bykov, *et al.*, Radioizmer. Élektron. KVARTS, No. 1, 5 (1994).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1413–1419. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 69–75.

OPTICS, QUANTUM ELECTRONICS

Possibility of Generating High-Energy Photons by Ultrarelativistic Electrons in the Field of a Terawatt Laser and in Crystals

A. Kh. Khokonov*, M. Kh. Khokonov*, and A. A. Kizdermishov**

** Kabardino-Balkar State University, ul. Chernyshevskogo 173, Nalchik, 360004 Russia*

*** Adygei State University, Maĭkop, 385000 Russia*

e-mail: Askhad_75@pochtamt.ru

Received November 27, 2001; in final form, April 8, 2002

Abstract—It is found that the spectral characteristics of relativistic electrons moving in the field of a terawatt laser and through static transverse fields differ significantly. It is shown that, as applied to studying the nonlinear generation of higher harmonics and quantum recoil and spin effects upon hard photon emission, the Baĭer–Katkov method has advantages over other methods. Numerical data for the efficiency of hard photon generation in the field of terawatt lasers and in oriented crystals are reported. *© 2002 MAIK "Nauka/Interperiodica".*

One of the topical problems in contemporary experimental and applied physics is producing sources of high-intensity gamma radiation [1, 2]. A promising source of intense gamma radiation is charged particle channeling in oriented crystals (OCs) (Kumakhov radiation [3]). The efficiency of this radiation increases with electron energy [4]. Recently, however, interest has been generated in the production of high-energy photons through the interaction of relativistic electrons with intense (~10^{12} W) laser fields. In this work, the electron energies E that we are interested in lie in the range of tens to hundreds of GeV, and the photon energies are comparable to E, $\hbar\omega \sim E$.

The interaction of relativistic electrons with a plane electromagnetic wave was studied in the mid-1960s [5]. For the lasers available at that time, the wave intensity was of minor importance for the nonlinear generation of higher harmonics. The quantum recoil and spin effects were taken into account with the Klein–Nishina formula for Compton scattering [6, 7]. In practice, polarized gamma radiation with a monochromatism of several percent can be generated by scattering laser photons by a relativistic electron beam [6, 8].

Recently developed powerful terawatt lasers [9–11], for which the Lorentz-invariant field parameter

$$\nu_0^2 = \frac{e^2\mathscr{E}_0^2}{2m^2\omega_0^2c^2} \quad (1)$$

may exceed unity, have opened up basically new opportunities for producing hard gamma quanta via the interaction of relativistic electrons with the laser field. In (1), $\mathscr{E}_0$ is the amplitude of the electric field of a laser wave with a frequency ω_0, e and m are the charge and mass of an electron, and c is the speed of light. Laser beams with $\nu_0 > 1$ will be referred to as superintense.

In this work, we compare the efficiency of generating hard gamma quanta upon electron emission in the field of a laser wave with that upon channeling. The emission of relativistic electrons in the field of a terawatt laser is considered based on the Baĭer–Katkov quasi-classical method [12]. The advantage of this method over recent approaches developed by other authors (see, e.g., [13]) is that it allows for the inclusion of the quantum recoil effects upon hard photon emission, the effect of the spin on the emission, and nonlinear effects in simple terms [14].

The appearance of higher harmonics in the electron emission spectra arising in the field of an intense plane wave (IPW) means that an electron absorbs several photons of the laser field and subsequently emits one photon whose energy is shifted, due to the Doppler effect, toward the "harder" range: $\omega \sim \omega_0\gamma^2$, where γ is the Lorentz factor. Theoretically, the cross sections of the associated processes were calculated using the Volkov solution to the Dirac equation [15, 16]. Experimentally, the emission of higher harmonics by electrons in the IPW field was studied in [9, 17]. In the former work, the first four harmonics were observed when electrons of energy 46.6 GeV interacted with a circularly polarized laser beam with a field parameter $\nu_0 = 0.6$.

Salamin and Faisal [13] theoretically investigated the generation of higher harmonics based on formulas of classical electrodynamics. They derived relatively simple analytical formulas for the case of a circularly polarized plane wave. Salamin and Faisal's classical calculations are consistent with earlier quantum electrodynamic results [15, 16] in the limit $\hbar\omega \ll E$.

1063-7842/02/4711-1413$22.00

When comparing the motion of electrons in the plane wave field and in external static transverse fields (ESTFs) (for example, in undulators or upon channeling), one can notice that the invariant called the nondipolity parameter, $\beta_\perp\gamma$, coincides with the field parameter given by (1) ($\beta_\perp$ is the velocity component transverse relative to the electron mean velocity) [18]. In what follows we will consider an electron moving toward a linearly polarized plane wave. In this case, the equality $\overline{\beta_\perp^2\gamma^2} = \nu_0^2$ is true, where the bar means averaging over the period of the transverse motion of the electron. This equality indicates that the emission in the IPW field differs greatly from that in ESTFs. In the former case, the nondipolity parameter does not depend on the electron energy; that is, at $\nu_0 \ll 1$, the dipole approximation is applicable even if the energy of photons emitted becomes comparable to the electron energy. Since, the transverse motion in the field of a plane wave is essentially harmonic, only a single harmonic with a sharp dipole peak at the center is emitted. In transverse static fields, where the emission is the result of virtual photon scattering by the electron, such a situation does not take place, since the nondipolity parameter $\beta_\perp\gamma$ grows with energy (for example, $\beta_\perp\gamma \sim \gamma^{1/2}$ upon channeling). When the angle of electron deflection by the external field exceeds the characteristic angle of radiation (i.e., when $\beta_\perp\gamma \gg 1$), the emission of higher harmonics becomes appreciable and the spectrum is described by formulas of synchrotron type [19].

For electron energies currently available with present-day accelerators ($E \approx 300$ GeV), the contribution of the spin to the emission upon channeling is insignificant [3] though tangible [4]. The reason is that the emission spectrum contains largely soft photons and the number of electrons with an energy $\hbar\omega \sim E$ is relatively small. When the electron moves in the IPW field, the energy of photons emitted grows with E faster than upon channeling. Then, if field parameter (1) is not too much greater than unity, the contribution of the spin to the hard region of the spectrum will play a decisive role, as will be shown subsequently. The detailed comparison of the emission upon channeling and that when the electron moves in the plane wave field for the case of the dipole emission spectrum lying in the X-ray range has been recently made in [18] with the generalized method of virtual photons [20].

Along with field parameter (1), we are interested in two more Lorentz invariants:

$$a = \frac{2\hbar k_0^\mu \rho_\mu}{m^2c^2} \approx \frac{2\hbar\Omega_0\gamma^2}{E}, \qquad \chi = \frac{e\hbar\left|F_{\mu\nu}p^\nu\right|}{m^3c^4} \approx \frac{e\mathscr{E}\hbar c}{m^2c^2}(1+\beta)\gamma, \tag{2}$$

where $\mathscr{E} = \mathscr{E}_0\cos\omega_0(z + z/c)$ is the electric field strength in the laser wave, k_0^μ is the four-wave vector of the incident laser wave, p^μ is the four-dimensional momentum of the electron, and $F^{\mu\nu}$ is the electromagnetic field tensor.

The quantity $\Omega_0 = \omega_0(1 + \beta)$ has the meaning of the transverse oscillation frequency for an electron in the plane wave field. The approximate equalities in (2) are valid for $\gamma \gg 1$. In this case, the value of a is time-independent up to $\beta_\perp^4$ and the time dependence of the invariant χ enters through the dependence $\mathscr{E} = \mathscr{E}(t)$. The right-hand side of formulas (2), as well as the subsequent consideration, applies to the case when the electron mean velocity is in opposition to the laser beam direction. The parameter χ has been well studied in terms of the theory of electromagnetic processes in constant external fields [5, 12, 16].

In the ultrarelativistic case $\gamma \gg 1$, the classical theory gives the following results for an electron in the field of a plane-polarized laser field [21]:

$$N_0^{-1}\frac{d^3N}{dudzd\varphi} = \frac{3}{\pi a}\sum_{k=1}^{\infty}\left[j_{xk}^2 + \frac{\eta_k^2}{2\nu_0^2}j_{zk}^2 - \sqrt{2}\frac{\eta_k}{\nu_0}j_{xk}j_{zk}\cos\varphi\right], \tag{3}$$

where $u = \hbar\omega/E$, $\eta_k^2 = ak/u - \nu_0^2 - 1$, and $z = ct$.

For a given harmonic number k, the frequency varies in the range $0 < u < ak/(1 + \nu_0^2)$. The quantities j_{zk} and j_{xk} are expressed through the Bessel functions $J_n(x)$, as was done in similar problems [3, §3.3; 5, Appendix A; 22]):

$$j_{zk} = B^{-1}\sum_{m=-\infty}^{\infty}(k+2m)J_m(A)J_{k+2m}(B), \qquad j_{xk} = \sum_{m=-\infty}^{\infty}J_m(A)J_{k+2m}(B), \tag{4}$$

where $A = (2a)^{-1}\nu_0^2 u$ and $B = 2\sqrt{2}\,a^{-1}\nu_0 u\eta_k\cos\varphi$.

The parameter N_0 in (3) is the number of photons emitted per unit length in a weak laser field ($\nu_0 \ll 1$):

$$N_0 = \frac{2}{3c}\alpha\nu_0^2\Omega_0, \tag{5}$$

where $\alpha = 1/137$.

Expression (3) was obtained without considering quantum recoil and spin effects upon emitting a hard photon (see below).

For the emission due to electron planar channeling in a parabolic potential $U = 4U_m x^2/d_p^2$, where U_m is the

potential well depth on a crystallographic plane and d_p is the interplanar spacing, the classical approach ($\gamma \gg 1$) gives similar results. In view of quantum recoil upon emission (which is reflected by the substitution $u \longrightarrow u' \equiv u(1-u)$ in classical expressions), the associated formula has the form

$$\frac{d^3N}{dudzd\varphi} = \frac{\alpha\nu^2}{2\pi\lambda_c\gamma}\sum_{k=1}^{\infty}\left[4j_{xk}^2 + \frac{\eta_k^2}{\nu^2}j_{zk}^2 - 4\frac{\eta_k}{\nu}j_{xk}j_{zk}\cos\varphi\right], \quad (6)$$

where $u = \hbar\omega/E$, $\eta_k^2 = 4\nu k/b - 2\nu^2 - 1$, $b = u'x_m/\lambda_c$, $\nu = \Theta_0\gamma$, $\Theta_0 = x_m\omega_0/2c$, the quantities ω_0 and x_m have the meaning of the frequency and amplitude of positron transverse oscillations in the channel $x_m = d_p(E_\perp/U_m)^{1/2}/2$, $E_\perp$ is the electron transverse energy, and $\lambda_c = \hbar/m_0c$ is the Compton wavelength.

For a given harmonic number k, the frequency lies in the range

$$0 < u < u_k/(1 + 2\nu^2 + u_k).$$

Here, $u_k = 4k\nu\lambda_c/x_m$ and j_{xk} and j_{yk} are given by formulas (4), where $A = b\nu/4$ and $B = b\eta_k\cos\varphi$.

The distributions calculated by (3)–(6) are very close to the spectral distributions of undulator radiation [22].

The Baĭer–Katkov quasi-classical theory [1, §5; 12] gives the following result for the spectrum of photons moving toward a linearly polarized plane wave:

$$N_0^{-1}\frac{d^3N}{dudzd\varphi} = \frac{3}{\pi a}\sum_{k=1}^{\infty}\left[g_{\text{class}}^{(k)} + \frac{uu'}{2}\left(g_{\text{class}}^{(k)} + \frac{j_{zk}^2}{2\nu_0^2}\right)\right], \quad (7)$$

$$g_{\text{class}}^{(k)} = j_{xk}^2 + \frac{\eta_k^2}{2\nu_0^2}j_{zk}^2 - \sqrt{2}\frac{\eta_k}{\nu_0}j_{xk}j_{zk}\cos\varphi, \quad (8)$$

where $u = \hbar\omega/E$, $u' = u/(1-u)$, $\eta_k^2 = ak/u' - \nu_0^2 - 1$, and $z = ct$.

For a given harmonic number k, the frequency lies in the range $0 < u' < ak/(1 + \nu_0^2)$ and j_{xk} and j_{yk} are given by formulas (4), where $A = (2a)^{-1}\nu_0^2u'$ and $B = 2\sqrt{2}\,a^{-1}\nu_0u'\eta_k\cos\varphi$.

Formulas (7) and (8) are the exact solution within the applicability domain of the Baĭer–Katkov quasi-classical approach, i.e., for $\gamma \gg 1$. According to (7), the form of the spectrum depends, in general, on the two invariants given by (2), which is consistent with Ritus's result [5, Chap. 2, formulas (38) and (39)]. The term $g_{\text{class}}^{(k)}$ in the brackets in (7) reflects the result of classical electrodynamics (see (3) and (4)) where the substitution $u \longrightarrow u/(1-u)$ has been made; the other term, including the product uu', is spin-related.

General formula (7) is greatly simplified in two limiting cases: $\nu_0 \ll 1$ (weak external field) and $\nu_0 > 1$ (strong external field).

The former case corresponds to the dipole approximation. Here, (7) passes to

$$N_0^{-1}\frac{d^2N}{dudz} = \frac{3}{2a}\left(1 - 2\frac{u'}{a} + 2\frac{u'^2}{a^2} + \frac{uu'}{2}\right), \quad (9)$$

where $0 < u' < a$ and the spin-related term is $uu'/2$.

Thus, in weak fields, the emission spectrum depends only on the invariant a. Quantum dipole formulas like (9) are well known in the theory of interaction between a plane wave and moving electrons [23, §101]. Unlike the electron motion in ESTFs, formula (9) is valid at electron energies as high as desired with the proviso that $\nu_0 \ll 1$. Specifically, upon positron planar channeling, the emission spectrum, as was shown above, is described by formulas similar to (7)–(9) (the classical component in emission (6)) [18]. However, the situation where the spin term would play a significant role in dipole formula (9) is unrealistic in ESTFs, because the condition for emission dipolity is inevitably violated at high energies, when the spin effect becomes essential (for positrons, this energy $E > 50$–100 GeV). In this sense, the spectrum of hard polarized photons with an energy $\hbar\omega \sim E$ generated by lasers that have $\nu_0 \sim 1$ is more advantageous than that obtained in OCs, since the former is more monochromatic and the multiphoton background in its soft range is absent.

The aforesaid is illustrated in Figs. 1–3, showing the spectra of emitted photons for electron energies of hundreds of GeV (in this case, $a = 3$ corresponds to the energy 150 GeV) in the field of a 1-μm laser. Figure 1 demonstrates the near-dipole situation. At the same parameters of an electron beam in OCs, the monochromatism of the emission is much lower [24–26] because of a high concentration of low-energy photons. At a fixed electron energy, the contribution of higher harmonics increases with laser beam intensity (cf. Figs. 2 and 3). Conversely, at a fixed laser beam intensity, the emission becomes more dipole with increasing electron energy. Such a situation is the reverse of that in ESTFs, where the contribution of higher harmonics grows with electron energy. Figures 4 and 5 show the spectra of positrons with energies of 10 and 100 GeV calculated with regard for all the above effects. In the spectrum in Fig. 5, the number of higher harmonics is so large that some of them are indistinguishable and the synchrotron approximation describes the spectrum quite adequately. The positron spectrum in Fig. 4 is close to undulator radiation spectra in shape [22].

Figures 1–3 demonstrate that lasers with $\nu_0 \leq 1$ are the most appropriate for generating individual hard

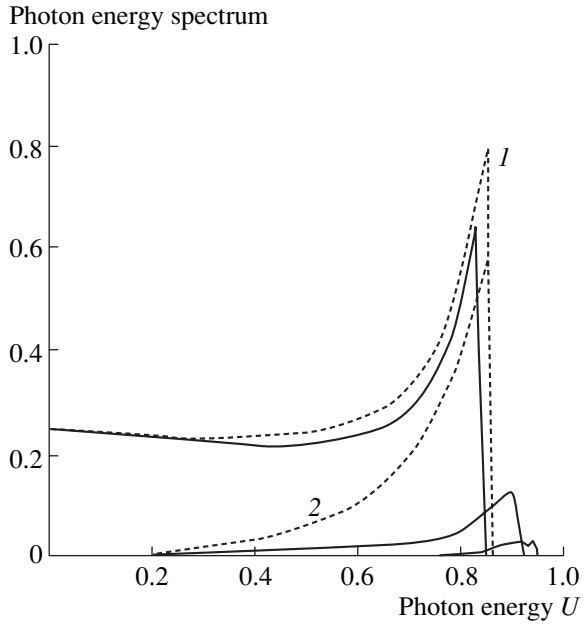


Fig. 1. Emission energy spectra of electrons moving toward a linearly polarized plane wave in a weak laser field for a = 6 and ν_0 = 0.5. Dashed lines *1* and *2*, spectrum in the dipole approximation and the spin contribution in the same approximation, respectively; continuous curves, exact calculation. On the ordinate axis functions (7) (continuous curves) and (9) (dashed lines) multiplied by u are plotted.

photons. In this case, the monochromatism is the highest and the number of photons emitted is fairly large. At lower or higher values of ν_0, both advantages of the emission in the IPW field disappear. For example, at $\nu_0 \ll 1$, the total probability of the emission defined by formula (5) is relatively small. At $\nu_0 > 1$, the role of nonlinear effects in Compton scattering increases and the spectrum becomes less monochromatic. Moreover, in superintense laser fields, the emission multiplicity [9] and the effect of electron expulsion from the laser beam because of the nonuniform spatial distribution of the intensity over its cross section become important [27].

The shape (7) of the emission spectrum is simplified when $\nu_0 > 1$, i.e., when harmonics with $k \gg 1$ play a decisive role. In this case, the spectrum also depends on one parameter (now on χ) and has the form

$$\frac{d^2N}{dudz} = \int_0^{2\pi} \frac{d^2N^{(\mathrm{syn})}(\alpha_0)}{dudz}\frac{d\alpha_0}{2\pi}, \tag{10}$$

where the integrand is the well-known quantum synchrotron formula

$$N_0^{-1}\frac{d^2N^{(\mathrm{syn})}}{dudz} = \frac{\sqrt{3}}{\pi a \nu_0^2}\left[(2+uu')K_{2/3}(\xi) - \int_\xi^\infty K_{1/3}(\eta)d\eta\right], \tag{11}$$

with $\xi = 2u'/(3\chi)$. Here, the parameter χ is expressed through the invariants a and ν_0: $\chi = a\nu_0 \sin\alpha_0/\sqrt{2}$.

Integration over α_0 in (10) means averaging (11) over the period of electron transverse motion. Formulas (10) and (11) are applicable if

$$\nu_0^2 u a^{-1}(1-u)^{-1} > 1. \tag{12}$$

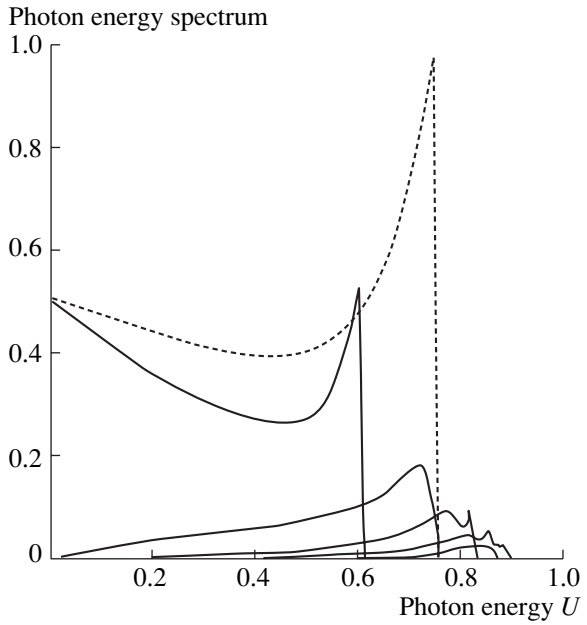


Fig. 2. The same as in Fig. 1 for a = 3 and ν_0 = 1.

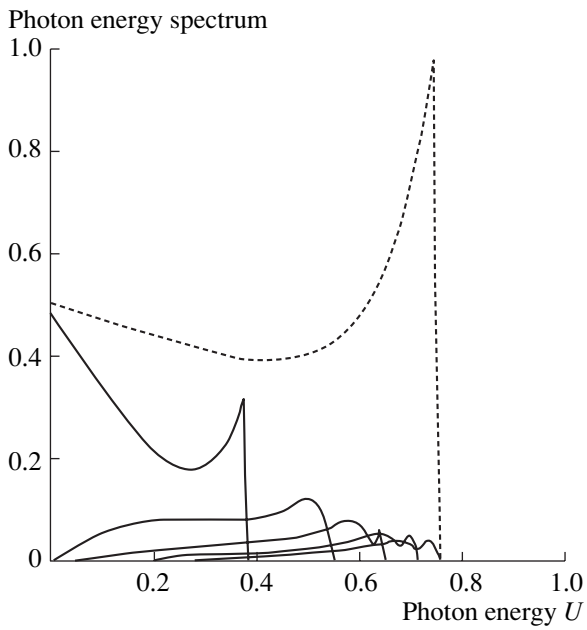


Fig. 3. The same as in Fig. 1 for a = 3 and ν_0 = 2.

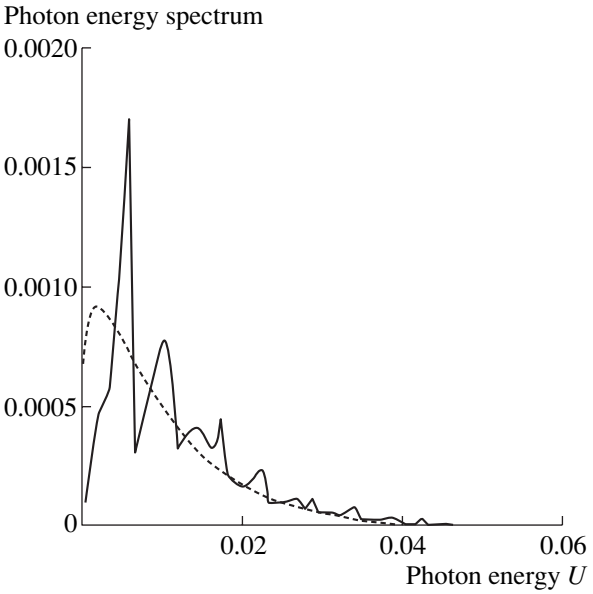


Fig. 4. Energy spectra of positron emission upon planar channeling in Si⟨110⟩. $\theta_0 = \theta_L\sqrt{\varepsilon}/2$, where $\varepsilon = 0.9U_m$ and θ_L is the Lindhard angle. The positron initial energy is E_0 = 10 GeV. Dashed line, spectrum in the synchrotron approximation; continuous curve, exact calculation.

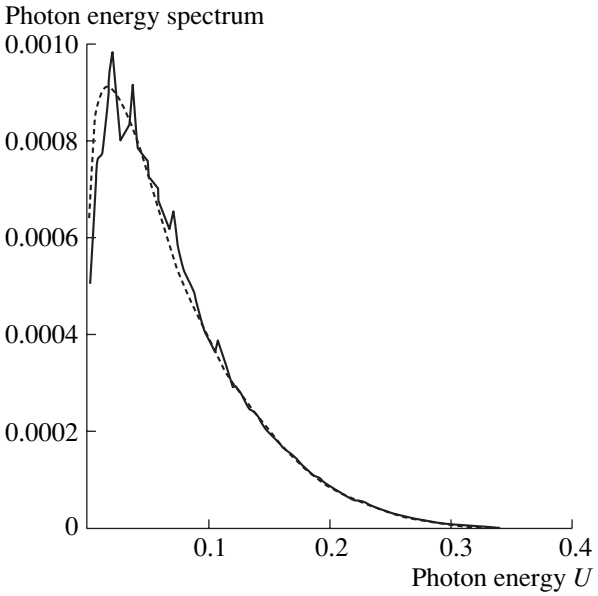


Fig. 5. The same as in Fig. 4 for E_0 = 100 GeV.

The synchrotron approximation, based on formula (11), is successfully used for calculating the emission spectra of electrons with energies above 100 GeV in OCs [19, 24, 25]; the higher the electron energy, the higher the accuracy of the synchrotron approximation (Fig. 5). For IPWs, conversely, at a fixed ν_0, the condition for synchrotron approximation applicability is violated as the electron energy grows, as follows from (12).

Figure 6 plots the total number of photons emitted per unit length (in units of N_0, see (5)) against the parameter a for various ν_0. The calculation by formulas (7)–(11) for a laser beam with a wavelength of 1 μm shows that the exact solution differs from the dipole approximation twofold even at $\nu_0 = 2$ (Figs. 3, 6).

In recent CERN experiments with 35–243 GeV electron beams [29], the effect of the spin on the emission in OCs has been shown for the first time. It has turned out that the spin contribution in the hard range of the spectrum is 20–25% at electron energies of several hundreds of GeV. For IPWs, the spin contribution will be shown to be much higher at the same energies. The emission spectra of electrons with various energies in the field of a laser wave with $\nu_0 = 2$ are shown in Fig. 7. From these spectra, it follows that an increase in the energy raises the dipolity of the emission. Simultaneously, the contribution of the spin term to the emission grows and becomes dominant at $\hbar\omega/E \sim 1$. Even at $\nu_0 = 2$ and $a = 6$, the spin term in the hard range of the spectrum prevails (Fig. 7, curve *2*). Note for comparison that in OCs, the effect of the spin term is much smaller at the same electron energies, as follows from theoretical calculations [4, 24, 25]. In addition, it is seen that the synchrotron approximation describes adequately the hard range of the spectrum if condition (12) is met. At a given ν_0, a rise in the electron energy deteriorates the accuracy of the synchrotron approximation.

When comparing the OC and IPW efficiencies, one should properly determine the effective length of emission. Our calculations for OCs rely on realistic models,

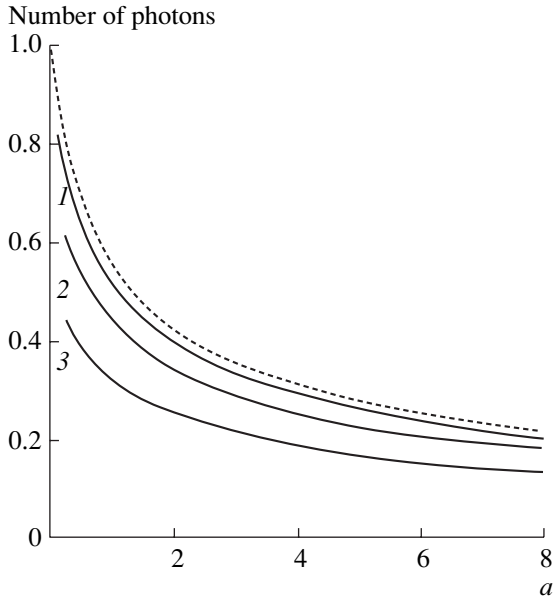


Fig. 6. Total number of emitted photons per unit length in units of N_0 (see (5)) vs. parameter a (see (2)). The dashed curve represents the dipole approximation: $\nu_0 \ll 1$. Continuous curves correspond to $\nu_0 = 0.5$ (*1*), 1 (*2*), and 2 (*3*). On the ordinate axis, the ratio N/N_0 is plotted.

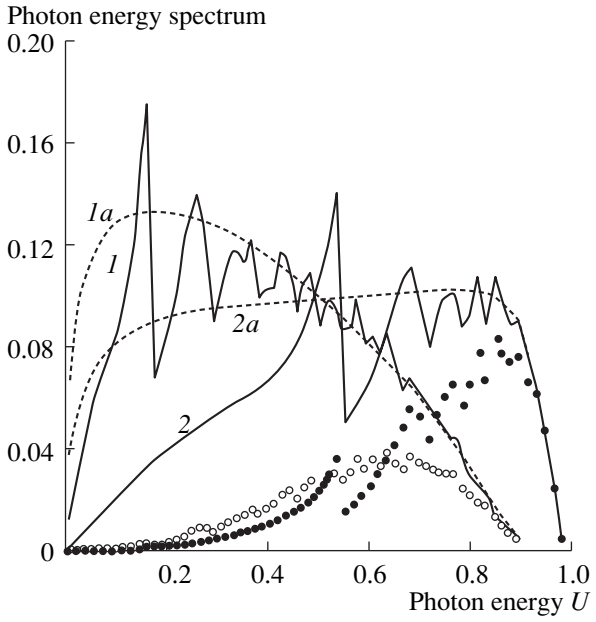


Fig. 7. The same as in Fig. 1. Continuous curves correspond to $a = 1$ (*1*) and 6 (*2*). The dashed curves *1a* and *2a* are the synchrotron approximations for the same values of a. (○) Spin contribution for $a = 1$ and (●) for $a = 6$.

which include secondary factors such as multiple scattering of electrons by individual atoms of a crystal, radiative damping of the transverse energy, etc. (for details, see [24, 25]). For terawatt lasers, the length of interaction with electrons of energies above several tens of GeV may reach several millimeters. In the experiment with a 46.6-GeV beam on the SLAC accelerator [9] (the laser wavelength ≈1 μm), the interaction length was found to be several hundreds of wavelengths (see also [18]). The associated emission yield was ≈1.8 photons per electron falling into the region of interaction with the laser. According to our analysis, only 10^{-3} of beam electrons took part in the interaction.

Of practical interest is the number of hard emitted photons per electron that have an energy in the range $0.8u_m < u < u_m$, where $u_m \sim a/(1 + a)$ is the maximal energy of a photon emitted (our estimates are restricted by the dipole approximation). According to (10), the number of such photons is $N_1 \sim 0.1N_0$ for $a > 3$, where N_0 is given by formula (5). For a 1-TW 1-μm laser interacting with 150-GeV electrons, we find that $N_1 \approx 0.6 \times 10^{-5}$ photons within a length $L = 0.5$ mm at a cross-sectional area of the laser beam $\Sigma = 1$ mm^2 (here, $u_m = 0.75$ E and $\nu_0^2 \approx 3.7 \times 10^{-5}$). This number increases significantly when the cross-sectional area of the beam shrinks. For example, when $\Sigma = 10^{-2}$–10^{-3} mm^2, the number of emitted photons per electron in the same spectral range becomes comparable to that obtainable at the same electron energy in OCs. This fact is of no surprise, since the field energy density in the focal spot of terawatt lasers may be comparable to the electrostatic energy density in the field of the crystal continuous potential [18]. It should be noted, however, that the photon yield in OCs at the energies considered far exceeds that obtainable with terawatt lasers. One reason is that a high degree of synchronization of laser field–electron interaction, which is necessary to increase the effective interaction time, is hardly feasible in the case of IPWs. In OCs, all electrons are involved in the interaction.

The number of individual hard photons with an energy $\hbar\omega$ lying in different intervals

	$0.5 < u < 0.8$	$0.8 < u < 0.9$	$u > 0.9$
C ⟨110⟩	0.085	2.0×10^{-3}	2.4×10^{-4}
150 GeV	(0.082)	(1.6×10^{-3})	(2.8×10^{-5})
Si ⟨110⟩	0.069	1.8×10^{-3}	5.2×10^{-4}
150 GeV	(0.062)	(7.0×10^{-4})	(2.0×10^{-5})
1000 GeV	0.22	2.2×10^{-2}	4.9×10^{-3}
	(0.215)	(2.1×10^{-2})	(4.4×10^{-3})
Amorphous target $z = L_{rad}/2$	0.10	0.021	0.014

Note: Here, $u = \hbar\omega/E$. The initial electron beam divergence angle is $0.2\Theta_L$, where Θ_L is the Lindhard critical angle. The thicknesses and orientation of the crystals are as follows: C ⟨110⟩, 2000 μm; Si ⟨110⟩, 1400 μm. The results for the amorphous target correspond to the thickness $z = L_{rad}/2$.

The table lists the average numbers of hard photons emitted by an electron in OCs in a given frequency range. Similar data are represented in [25], where the spectrum of individual photons most of which had low energies $u \equiv \hbar\omega/E < 0.5$ was looked for. Therefore, approximations made in [25] overestimate the noncoherent contribution to the hard range of the spectrum. The reason is that in [25] the coherent and incoherent contributions were simulated independently, which is actually a rather coarse approximation if one is interested in the hard range of the spectrum. The contribution from the coherent emission to the field of the atomic chain continuous potential is given in parentheses. For the soft range of the spectrum, $u < 0.5$, the number of photons emitted obtained in this work and in [25] exactly coincide; therefore, only data for hard photons are tabulated.

It follows from the table that, at an electron energy of 150 GeV, the coherent contribution plays a decisive part for photons with energies $u < 0.8$. As the electron energy grows to 1000 GeV, the coherent contribution prevails almost throughout the spectrum. Yet, even at energies of ~1 TeV, an amorphous thick target is more effective than OCs for generating photons with $u > 0.9$.

One more advantage of powerful lasers over OCs is that they allow for the generation of circularly polarized photons. In turn, OCs are preferred in the case when a large total number of photons are to be generated irre-

spective of the spectrum monochromatism (i.e., in order to produce intense positron beams [30]).

REFERENCES

1. V. N. Baĭer, V. M. Katkov, and V. M. Strakhovenko, *Electromagnetic Processes under High Energy in Oriented Single Crystals* (Nauka, Novosibirsk, 1989).
2. A. I. Akhiezer and N. F. Shul'ga, *Electrodynamics of High Energies in Matter* (Nauka, Moscow, 1993).
3. V. V. Beloshitsky and F. F. Komarov, Phys. Rep. **93** (3), 117 (1982).
4. A. H. Sorensen, Nucl. Instrum. Methods Phys. Res. B **119**, 1 (1996).
5. V. I. Ritus, Tr. Fiz. Inst. Akad. Nauk SSSR **111**, 5 (1979).
6. R. N. Milburn, Phys. Rev. Lett. **10**, 75 (1963).
7. I. I. Goldman, Phys. Lett. **128**, 664 (1962).
8. F. R. Arutyunian and V. A. Tumanian, Phys. Lett. **4**, 176 (1963).
9. C. Bula, K. T. McDonald, E. J. Prebys, *et al.*, Phys. Rev. Lett. **76**, 3116 (1996).
10. P. Eisenberger and S. Suckewer, Science **274**, 201 (1996).
11. R. W. Schoenlein, W. P. Leemans, A. H. Chin, *et al.*, Science **274**, 236 (1996).
12. V. N. Baĭer and V. M. Katkov, Zh. Éksp. Teor. Fiz. **53**, 1478 (1967) [Sov. Phys. JETP **26**, 854 (1968)]; Zh. Éksp. Teor. Fiz. **55**, 1542 (1968) [Sov. Phys. JETP **28**, 807 (1969)].
13. Y. I. Salamin and F. H. M. Faisal, Phys. Rev. A **54**, 4383 (1996).
14. A. Kh. Khokonov, M. Kh. Khokonov, and R. M. Keshev, Pis'ma Zh. Tekh. Fiz. **24** (20), 20 (1998) [Tech. Phys. Lett. **24**, 797 (1998)].
15. I. I. Goldman, Zh. Éksp. Teor. Fiz. **46**, 1412 (1964) [Sov. Phys. JETP **19**, 954 (1964)].
16. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **46**, 776 (1964) [Sov. Phys. JETP **19**, 529 (1964)]; Zh. Éksp. Teor. Fiz. **46**, 1768 (1964) [Sov. Phys. JETP **19**, 1191 (1964)]; Zh. Éksp. Teor. Fiz. **47**, 1130 (1964) [Sov. Phys. JETP **20**, 757 (1964)].
17. T. J. Englert and E. A. Rinehart, Phys. Rev. A **28**, 1539 (1983).
18. M. Kh. Khokonov and R. A. Carrigan, Nucl. Instrum. Methods Phys. Res. B **145**, 133 (1998).
19. J. C. Kimball and N. Cue, Phys. Rev. Lett. **52**, 1747 (1984).
20. J. Lindhard, Phys. Rev. A **43**, 6032 (1991).
21. A. Kh. Khokonov, M. Kh. Khokonov, and R. V. Keshev, Nucl. Instrum. Methods Phys. Res. B **145**, 54 (1998).
22. A. D. Alferov, Yu. A. Bashmakov, and E. G. Bessonov, Tr. Fiz. Inst. Akad. Nauk SSSR **80**, 100 (1975).
23. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Course of Theoretical Physics*, Vol. 4: *Quantum Electrodynamics* (Nauka, Moscow, 1989; Pergamon, New York, 1982).
24. M. Kh. Khokonov, Pis'ma Zh. Éksp. Teor. Fiz. **56**, 349 (1992) [JETP Lett. **56**, 333 (1992)]; Zh. Éksp. Teor. Fiz. **103**, 1723 (1993) [JETP **76**, 849 (1993)].
25. A. Kh. Khokonov and M. Kh. Khokonov, Zh. Tekh. Fiz. **68** (9), 37 (1998) [Tech. Phys. **43**, 1043 (1998)].
26. K. Kirsebom, R. Medenwaldt, U. Mikkelsen, *et al.*, Nucl. Instrum. Methods Phys. Res. B **119**, 79 (1996).
27. T. W. B. Kibble, Phys. Rev. Lett. **16**, 1054 (1966); Phys. Rev. **150**, 1060 (1966).
28. N. G. Klepikov, Zh. Éksp. Teor. Fiz. **26**, 19 (1954).
29. K. Kirsebom, U. Mikkelsen, E. Uggerhoj, *et al.*, Phys. Rev. Lett. **87**, 054801 (2001).
30. X. Artru, V. N. Baier, T. V. Baier, *et al.*, Nucl. Instrum. Methods Phys. Res. B **119**, 246 (1996).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1420–1426. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 76–82.

OPTICS, QUANTUM ELECTRONICS

Properties of Oxide Nanopowders Prepared by Target Evaporation with a Pulse–Periodic CO_2 Laser

Yu. A. Kotov, V. V. Osipov, M. G. Ivanov, O. M. Samatov, V. V. Platonov, E. I. Azarkevich, A. M. Murzakaev, and A. I. Medvedev

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

e-mail: max@iep.uran.ru

Received February 27, 2002; in final form, May 6, 2002

Abstract—The design and characteristics of a setup for producing metal oxide nanopowders with an output of up to 20 g/h are discussed. The grain mean size in the powders is 15 nm, and the radiation power consumption is 30–40 (W h)/g. Y_2O_3-stabilized ZrO_2 (YSZ) and Al_2O_3 + YSZ nanopowders are prepared by target evaporation with a pulse–periodic CO_2 laser, followed by vapor condensation in an air stream. The mean power, peak power, and efficiency of the pulse–periodic CO_2 laser, excited by a combined discharge, are, respectively, 1 kW, 10 kW, and ≈10%. Data for the powder specific surface, grain shape, and grain size distribution, as well as results of X-ray phase and structure analysis, are reported. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Recent technologies of synthesizing bulk nanostructured materials with improved mechanical and novel electromagnetic and optical properties [1, 2] have generated interest in producing powders with nanometer-sized grains (nanoparticles). Nanoparticles can be produced by a variety of methods (for their detailed analysis, see, e.g., [2, 3]). Laser radiation has not found wide recognition for this purpose because of the low output and high energy consumption of related techniques. At the same time, laser-assisted material evaporation and subsequent vapor condensation has been known for a long time [4]. This approach to material production provides high-purity and fine-grain deposits, is cheap, and can be applied to various targets. However, its competitiveness with other approaches has been proved only recently.

For example, the output Δm of a CO_2 laser in this process has been shown [3, 5, 6] to depend considerably on the radiation power P and intensity I, as well as on the target velocity V_t relative to the laser beam. For example, for $P = 3270$ W, $I = 4.2 \times 10^5$ W/cm^2, and $V_t = 28$ cm/s, Δm, the energy consumption ΔW, and the grain size d_{BET} were found to be 130 g/h, 25 (W h)/g, and 60 nm, respectively, for ZrO_2 powders.

To raise the intensity (to 10^7 W/cm^2), the laser was switched to the pulsed mode with a mechanical chopper or a Fabry–Perot interferometer [2, 3]. In this case, the grain mean size was reduced to $d_{BET} \approx 15$ nm; however, at the same radiation mean power $P = 600$ W, Δm and ΔW were 11 g/h and 54.5 (W h)/g, respectively, for the pulse duration $t_p = 25$ μs and $\Delta m = 3.8$ g/h and $\Delta W = 158$ (W h)/g for $t_p = 100$ μs. These values are much lower than those obtained in the cw mode with $P = 600$ W: $\Delta m = 23.8$ g/h and $\Delta W = 25$ (W h)/g.

Yet, we believe that the pulsed mode must provide not only finer grains (because of an increased intensity and more efficient vapor removal from the hot zone) but also at least no higher energy consumption than in the cw mode. Our opinion relies on the fact that the probability of an elementary event of evaporation grows exponentially with the melt temperature and, hence, with the radiation peak power density [4]. Such a situation takes place up to optical breakdown. In [3], the pulsed mode was obtained by switching the Q factor of the laser cavity. Q switching imparts the specific shape to a laser pulse, because of which the advantages of the pulse–periodic mode of preparing nanodispersed powders cannot be revealed in full measure. Moreover, Q switching decreases appreciably the laser efficiency.

In this work, nanograin powders are produced with a CO_2 laser excited by a pulse–periodic combined discharge [7].

EXPERIMENTAL

In our experiments, the lasing medium was excited by a pulse–periodic combined discharge. In this case, a plasma in the discharge gap is initiated by a short-term self-sustained discharge and a major fraction (97–98%) of the energy is delivered during a long-term non-self-sustained discharge under a reduced field strength that is optimal to excite the laser upper level [4]. The excitation scheme makes it possible to separate the high- and low-voltage feed circuits without using discharge-current-limiting elements, which therefore successively

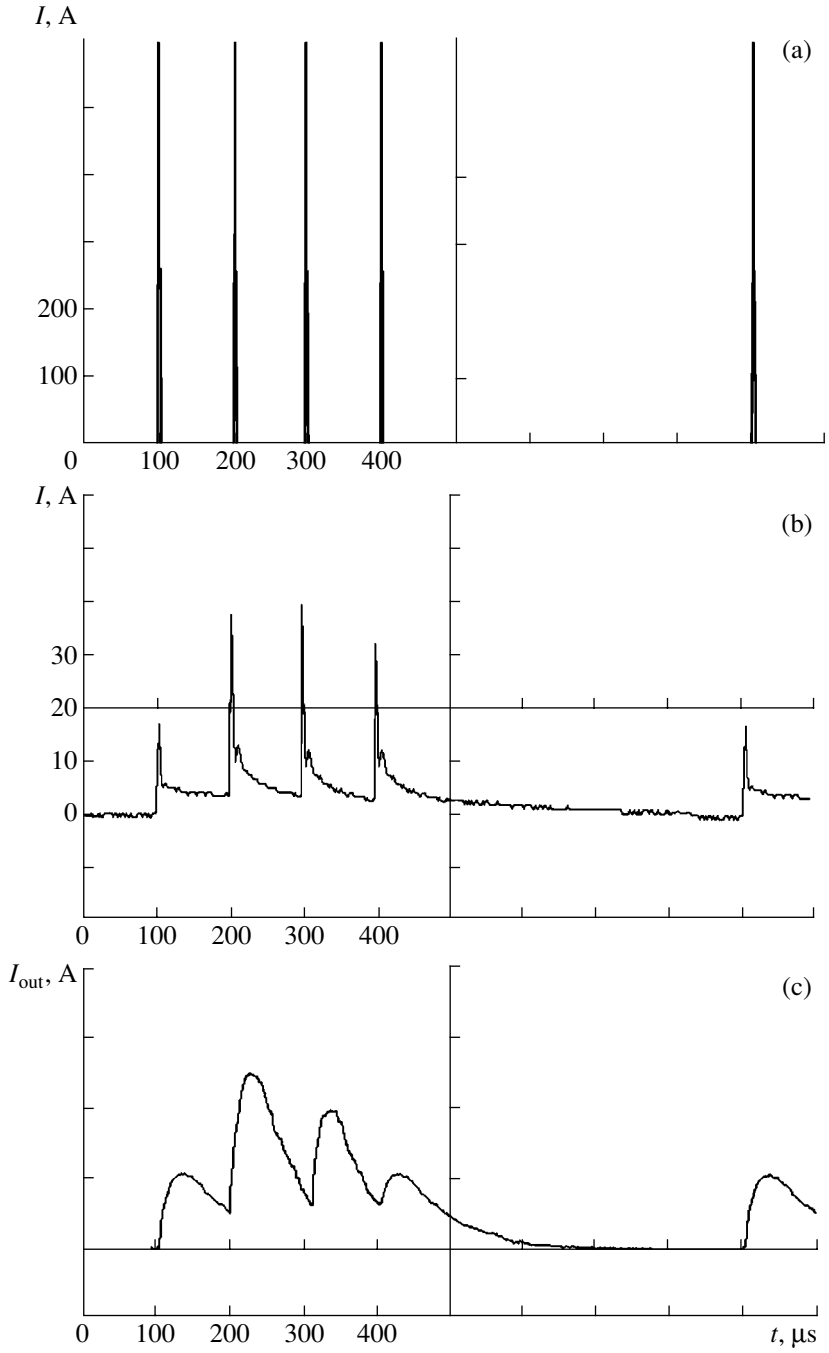


Fig. 1. Waveforms of current pulses for (a) self-sustained discharge, (b) non-self-sustained discharge, and (c) laser radiation.

excites the self-sustained and non-self-sustained discharges.

The waveforms of discharge and radiation current pulses illustrating the operation of the laser and the possibility of varying the lasing duration are shown in Fig. 1. A self-sustained discharge (duration ≈100 ns; Fig. 1a) in the two-electrode gap was initiated by a Blumlein generator. Once the self-sustained discharge had been switched off, the non-self-sustained discharge (duration 100 μs) was maintained in the gap (Fig. 1b). The voltage across the gap was optimal for the energy to be transferred to the laser upper level. After the current of the non-self-sustained discharge had declined to a certain level, a high-voltage pulse was applied to the gap again and the process was repeated.

The lasing medium could be pumped by both single shots and a pulse packet. The packet frequency could be varied from 50 to 1000 Hz; the number of pulses in a packet, from 1 to 4; and the pulse-to-pulse time in a packet, from 50 to 200 μs according to the number of

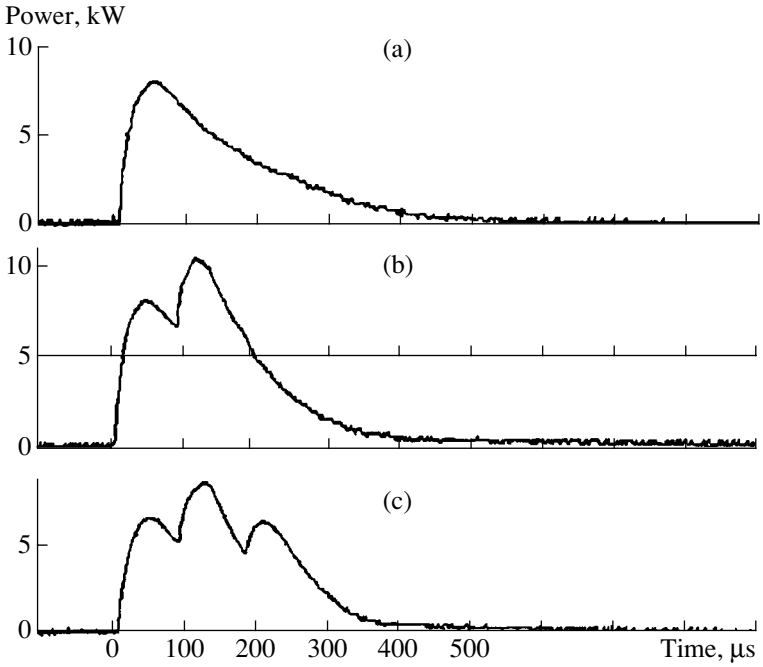


Fig. 2. Waveforms of radiation pulses for (a) one, (b) two, and (c) three self-sustained discharge pulses in a packet.

pumping pulses. Typical waveforms of pulses in a packet are shown in Figs. 1 and 2. If the number of pulses per packet is no more than three, the lasing duration remains virtually unchanged (only the shape of laser pulses was changed). The peak power P_{max} could reach 10 kW, even if the mean power did not exceed 1 kW (Fig. 2b).

The radiation power was maximum at a frequency of 750–900 Hz for H_2- and CO-free mixtures and 400–600 Hz for mixtures containing H_2 or CO. We found that a mixture of composition $CO_2 : N_2 : He = 1 : 4 : 8$ provides efficiency as high as 22% at a pressure of 80 torr. However, during the operation, the efficiency dropped noticeably because of CO_2 dissociation and electronegative species appearing in the mixture. Therefore, the search for long-term operating conditions was carried out with CO-containing mixtures, where the initial composition can be maintained by means of the reducing reaction $CO + O \longrightarrow CO_2$ to compensate for the dissociation. In a $CO_2 : CO : N_2 : He = 1.5 : 1 : 10 : 20$ mixture at a pressure of 60 torr, a radiation mean power of 600–800 W was obtained (at an efficiency of 10%), which did not decrease during the operation.

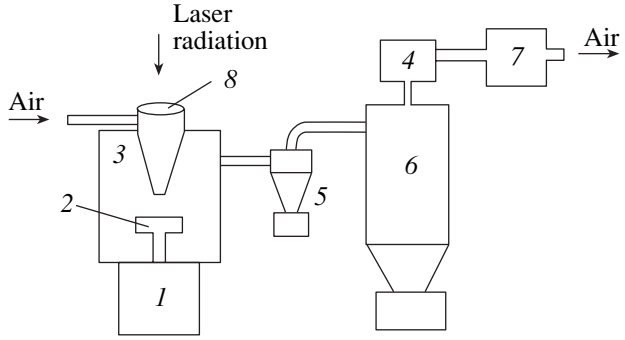


Fig. 3. Experimental setup for nanopowder production.

An experimental setup for producing nanopowders is depicted in Fig. 3. Laser radiation was focused on target *2* with lens *8*, which serves simultaneously as the entrance window of chamber *3*. Drive *1* rotates and moves target *2* in the horizontal plane so that the laser beam scans its surface with a constant rate, providing uniform sputtering of the target surface. During the experiment, the target was displaced in the axial direction so that its surface remained in the plane of the focal spot. The focal length of a KCl lens was 10 cm; the focal spot diameter, 0.45 mm.

The targets used were YSZ, $ZrO_2 + Y_2O_3$, and $YSZ + Al_2O_3$ powders with a grain size of several tens of micrometers and a desired oxide ratio that were pressed into pellets. The beam scan rate was 20 cm/s. Between pulses, the target was displaced by $D \approx 0.045$cm. Fan *4* pumped a working gas (air or an $N_2 : O_2 = 79 : 21$ mixture) into sealed evaporation chamber *3* and transported the powder to cyclone *5* and electrostatic precipitator *6*, where the powder was trapped. Before being discharged into the atmosphere, the air was purified by mechanical filter *7*. Near the target surface, the gas velocity was ≈15 m/s.

We prepared 2.8YSZ, 4.1YSZ, 8.6YSZ, and 9.85YSZ powders (figures mean the mole percentage of Y_2O_3), as well as powders of *A*41.1 + 1.45YSZ58.9 and *A*88.8 + 1.1YSZ11.2 mixtures ($A = Al_2O_3$, and fig-

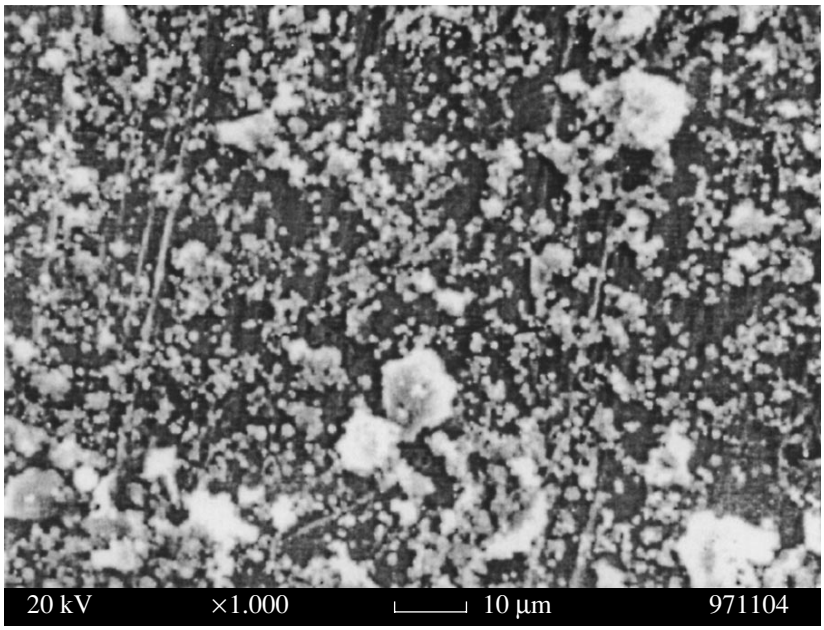


Fig. 4. Typical micrograph of the powder sediment.

ures at the end mean the weight percentage of the component). Both the raw materials and powders prepared were analyzed by X-ray techniques. The elemental composition was determined by spectral methods with a Jobin Yvon 48 instrument. Phase and structure analyses were performed with a DRON-4 X-ray diffractometer. The specific surface was found by the BET method using an argon–helium mixture with a GKh-1 chromatograph. The grain shape and grain size distribution were studied with JEM-200 transmission and JSM-T220A scanning electron microscopes. The fractional composition of the powders was investigated by sedimentation analysis. The content of moisture and volatiles was determined thermogravimetrically with a Q1500 instrument.

In preliminary experiments on choosing optimal conditions for preparing nanopowders, we varied the number of pulses per packet from 1 to 4 (Figs. 1, 2). The energy and the peak power of a pulse were varied by varying the pumping pulse energy. Under these conditions, the peak power density of the radiation varied from 1 to 5.2 MW/cm^2. The maximum output in nanopowder production, 20 g/h, was observed for 4-h continuous operation with a two-pulse packet (Fig. 2b) and a maximal power density of 5.2 MW/cm^2. This finding is to some extent unexpected, since, according to [8], optical breakdown is inevitable for such long pulses with power densities exceeding 2×10^6 W/cm^2.

RESULTS AND DISCUSSION

For all the compositions studied, the output was 15–20 g/h and the energy consumption, 40–30 (W h)/g. These values are very close to those obtained in [3]. Bearing in mind that the energy required for ZrO_2 evaporation is about 2.7 (W h)/g, we conclude that our results and those obtained in [3] indicate the low efficiency of using the radiation energy. Therefore, it is necessary to find reasons for such high losses and ways to reduce them.

The analyses of the powders showed that they are weakly agglomerated and have two fractions with greatly differing sizes. The first fraction (Fig. 4) consists of spherical grains with sizes from 0.2 to 2.0 μm and accounts for 3–7 wt %. Presumably, it forms by splashing the liquid phase, since the composition of these grains either is close to that of the raw material or is enriched by the component with a higher point of evaporation (Y_2O_3), while the other fraction, nanofraction, always has a deficiency of this component compared with the raw material (Table 1). The coarser fraction sometimes contains shapeless grains of size to 10 μm (possibly, target fragments).

From the grain size distribution (Fig. 5b) in the fraction remaining in the suspension (Fig. 5a), it follows that 98% of grains have sizes <40 nm and only 0.25% of grains have a size from 65 to 100 nm. The latter grains are nearly spherical and slightly faceted.

Upon sedimentation, the grain specific surface usually decreases by 10%, which indicates that the powders are slightly agglomerated. However, after the sedimentation, the powders have a bulk weight that is one order of magnitude higher, becoming more producible for subsequent applications.

The study of the phase composition and structure of the YSZ powders (Table 1) showed that they are single-phase and their averaged lattice parameter is well

Table 1

No.	Raw material for target preparation	Powder after sedimentation
1	Powder mixture: ZrO_2, S = 20 m^2/g; 3.1Y_2O_3, S = 4.5 m^2/g	2.8YSZ: S = 68 m^2/g; T: a = 5.106 and c = 5.1638 Å, grain size D = 19 nm, weight of volatiles $m(V)$ = 2.7 wt %. Sediment 3.4YSZ: T: a = 5.1084 and c = 5.1674; D = 26 nm. 6% of M phase. Melted surface layer of target 5.2YSZ
2	Powder mixture: ZrO_2, S = 20 m^2/g, T = 45 and M = 55 wt %; 4.5Y_2O_3, S = 4.5 m^2/g	4.15YSZ: S = 64.4 m^2/g, T: a = 5.115 and c = 5.161 Å, D = 17 nm, $m(V)$ = 2.6 wt %. Sediment 4.35YSZ; $C - a$ = 5.13 Å, D = 25 nm. 7% of M phase.
3	Powder mixture: ZrO_2, S = 51 m^2/g, T = 60 and M = 40 wt %; 9.1Y_2O_3, S = 4.5 m^2/g	8.6YSZ: S = 86 m^2/g; $C - a$ = 5.1405 Å D = 17 nm, $m(V)$ = 2.4 wt %. Sediment 8.9YSZ: $C - a$ = 5.144 Å, D = 25 nm. 7% of M phase
4	Powder 10.15YSZ, S = 6.1 m^2/g, $C - a$ = 5.1448E	9.85YSZ: S = 79 m^2/g, $C - a$ = 5.1459 Å, D = 18 nm, $m(V)$ = 2.8 wt %. Sediment 10.4YSZ: $C - a$ = 5.15 Å, D = 41 nm

Note: In raw material mixtures, the mole percentage of Y_2O_3 powder is indicated. M, T, and C stand for monoclinic, tetragonal, and cubic lattice, respectively.

Table 2

Target composition	Target material	Powder after sedimentation
Powder mixture A40 + 1.65YSZ60	Al_2O_3: S = 74 m^2/g; γ ≈ 20 wt %; δ ≈ 80 wt % 1.65YSZ: S = 7.76 m^2/g. M = 58 and T = 42 wt %. Grain size D = 70 nm.	A41.1 + 1.45YSZ58.9, S = 80.6 m^2/g, $m(V)$ = 4.9 wt %. Composition and structure: 1.45YSZ: T—31 wt %: a = 5.095 and c = 5.156E, D = 11 nm. K—28 wt %: a = 4.924 Å, D = 6 nm. γ = Al_2O_3—20 wt %, D = 10 nm. Amorphous Al_2O_3—21 wt %.
Powder mixture A85 + 1.65YSZ15	"	A88.8 + 1.15YSZ11.2, S = 86 m^2/g, $m(V)$ = 4.8 wt %. Preliminary composition and structure: cubic and tetragonal 1.15YSZ and γ-Al_2O_3, amorphous Al_2O_3; diffraction pattern interpretation is being continued

described by a straight line (Fig. 6) whose initial portion coincides with data for pure ZrO_2 [9] obtained by the plasma–chemical method and containing 90% of the tetragonal phase. Our data for ZrO_2 nanopowders obtained by electrical explosion and containing 55 to 65 wt % of the tetragonal phase [10] fall in the range shown in Fig. 6.

The study of the elemental composition of the powder mixtures (Table 2) suggests that a decrease in the high-temperature component content in these samples, prepared by coevaporation, is much more significant than in the YSZ powders. This decrease becomes more pronounced as the content of the low-temperature component (Al_2O_3, point of evaporation T_e = 3800 K, melting point T_m = 2320 K) grows. For example, in mixture 1, the YSZ content decreased by ≈2% and the Y_2O_3 content in YSZ, by 12%. In mixture 2, the corresponding values are 25 and 30%, respectively. In other words, the initial equality of the component volumes (mixture 1) and weights (mixture 2) in the targets is violated in the powders. The phase composition of the powders also turned out to be unusual. For example, in mixture 1, about 40 wt % of 1.45YSZ is in the cubic phase. Moreover, the lattice parameter of both the cubic and the tetragonal phases is much smaller than could be expected from the concentration dependence (Fig. 6). The monoclinic phase is completely absent.

It can be proposed that in the liquid state some amount of Al_2O_3 dissolves in YSZ, forming the solid solution. Assuming that the YSZ lattice parameter varies linearly from the value 5.123 Å for pure 1.45YSZ according to the zirconium-to-aluminum atomic radius ratio (≈1.2), one can estimate the amount of Al_2O_3 dissolved in the YSZ lattice. These estimates show that the tetragonal phase occupies ≈0.7 wt % and the cubic phase, ≈19.4 wt % Al_2O_3. From the diffraction patterns, it also follows that Al_2O_3 is partially in the amorphous state. If it is assumed that 1.45YSZ is entirely crystalline, the amount of the amorphous Al_2O_3 in the power considered is ≈21 wt %.

The diffraction pattern for the A88.8 + 1.15YSZ11.1 powder was still harder to interpret, because the line asymmetry and the background level were very high. It

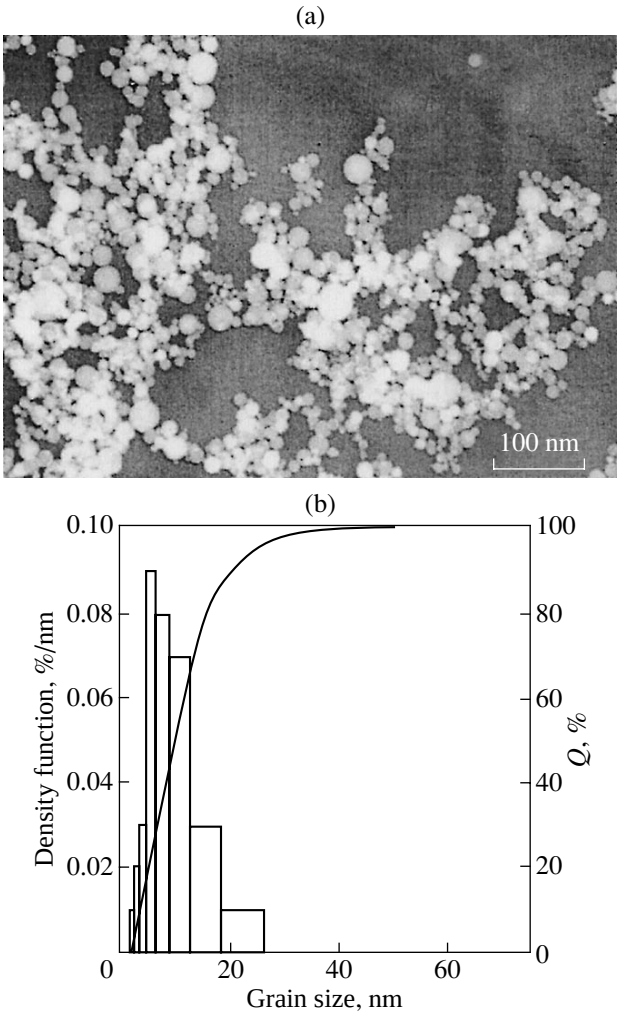


Fig. 5. (a) Typical micrograph of the powder fraction from the suspension and (b) grain size distribution.

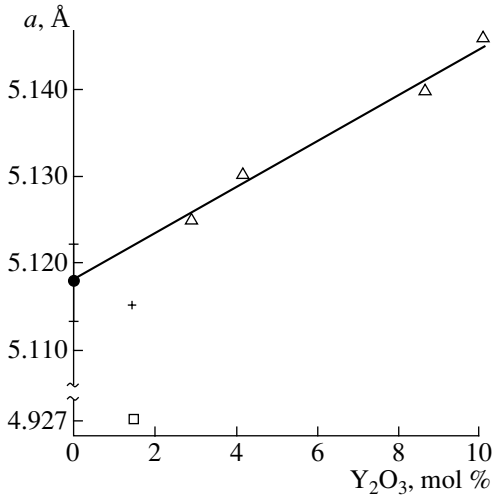


Fig. 6. YSZ lattice parameter vs. Y_2O_3 content: (△) our data for the YSZ powder, (●) published data for ZrO_2 with 90 wt % of the tetragonal phase, (+) tetragonal phase of the *A*40 + 1.45YSZ mixture, and (□) cubic phase of the *A*40 + 1.45YSZ mixture.

is only clear that an increase in the Al_2O_3 content does not increase the fraction of the cubic modification relative to the tetragonal one in the powder after laser synthesis. However, upon sedimentation, a significant fraction of the powder in the tetragonal phase precipitates. The high asymmetry of the diffraction lines, especially for the tetragonal phase, leads to an error in determining the weight composition of the phases and grain size. This, and also the fact that the YSZ tetragonal phase, which has a grain size of 6 nm, almost completely precipitated, lead us to suggest that YSZ tetragonal grains are covered by aluminum oxide and seem to have a greater size. That is why they precipitate upon sedimentation.

Thus, the above data lead us to conclude that the powders prepared in this work are of interest for further investigating their structure, which depends on the Al_2O_3/YSZ ratio in the starting mixture.

CONCLUSIONS

(1) A pulsed CO_2 laser makes it possible to decrease fourfold the grain size in air under normal conditions. The power consumption is no higher than the lowest power consumption observed for cw lasers [1].

(2) The efficiency of using the radiation for preparing nanopowders is less than 10%. This parameter can be improved. The output can also be increased even at the same mean power by shrinking the radiation pulses and increasing the pulse repetition rate.

(3) A method and a setup for preparing pure nanopowders of complex composition with a narrow grain size distribution were developed. Targets can be made of coarser grain powders of both available chemical compounds and mechanical mixtures of the components. To obtain a desired composition, it is necessary to take into account that in the powder, the content of a component with a higher point of evaporation is lower than in the raw material.

(4) A powder prepared by evaporating targets made of YSZ + Al_2O_3 mixtures contains YSZ crystallites with a lattice parameter that is much smaller (by ≈4%) than in pure YSZ of the same composition. This decrease in the lattice parameter is likely to be due to the formation of the Al_2O_3–YSZ solid solution.

ACKNOWLEDGMENTS

The authors are indebted to T.M. Demina for specific surface measurements, V.M. Tel'nova for carrying out sedimentation experiments, and N.I. Moskalenko for performing elemental analysis.

This work was supported by the INCO COPERNICUS Foundation (contract no. IC15-CT97-0713) and the Ministry of Science of the Russian Federation (project no. 30).

REFERENCES

1. Nanostruct. Mater. **12** (1999).
2. A. I. Gusev and A. A. Rempel', *Nanocrystalline Materials* (Nauka, Yekaterinburg, 2000).
3. E. Muller, Ch. Oestreich, U. Popp, *et al.*, Kona (Hirakata, Jpn.), No. 13, 79 (1995).
4. S. I. Anisimov, Ya. A. Imis, G. S. Romanov, and Yu. V. Khodyko, *High Power Radiation Effect on Metals* (Nauka, Moscow, 1970).
5. E. Muller, Ch. Oestreich, and U. Popp, in *Proceedings of the 4th EuroCeram. Conference* (Galassic, Itali, 1995), Vol. 1, p. 291.
6. G. Michel, G. Staupendahl, G. Eberhardt, *et al.*, in *Proceedings of the 5th EuroCeram. Conference on Key Engineering Materials, France, 1995*, Vol. 1, p. 161.
7. V. V. Osipov, M. G. Ivanov, and V. V. Lisenkov, Opt. Atmos. Okeana **10**, 1266 (1997).
8. N. G. Basov, V. V. Bashenko, and E. V. Glotov, Izv. Akad. Nauk SSSR, Ser. Fiz. **48**, 2310 (1984).
9. V. N. Troitskiĭ, E. N. Kurkin, V. I. Torbov, *et al.*, Neorg. Mater. **30**, 1436 (1994).
10. Y. A. Kotov, I. V. Beketov, A. M. Murzakaev, *et al.*, Mater. Sci. Forum, 913 (1996).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1427–1433. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 83–89.

ACOUSTICS, ACOUSTOELECTRONICS

Theory and Analysis of a Hybrid SAW-Resonator Filter with Enhanced Out-of-Band Suppression

V. F. Dmitriev

Avangard–Élionika Closed Joint-Stock Company, St. Petersburg, 195271 Russia

e-mail: elionica@rol.ru

Received March 27, 2002; in final form, April 23, 2002

Abstract—Modified equations for coupled modes are used to develop a theory of longitudinal resonant mode filters. Theoretical and experimental results are compared. The design of a hybrid SAW-resonator filter where both longitudinal resonant modes and resonant modes of a ladder structure are used is proposed. This design provides enhanced out-of-band suppression. *© 2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Due to the rapid advance in mobile communications in recent years, great interest has arisen in filters on SAW resonators. The main advantage of these filters is small insertion losses (1–4 dB) at a reasonable out-of-band suppression (25–60 dB). These properties make these filters promising for communications systems, in particular, for cellular phones. Most cellular phones, irrespective of the standard they use (including GSM, PCS, PCN, etc.), incorporate frequency-selective elements implemented as SAW resonant filters. The basic types of SAW resonant filters are SAW-resonator ladder filters [1], NCRF filters (based on transverse resonant modes) [2], and DMSF filters (based on longitudinal resonant modes) [3]. Each of them has advantages and disadvantages of its own. The main disadvantage of a ladder filter is its relatively low out-of-band suppression when the frequency is offset from the center frequency by several bandwidths. Filters on transverse resonant modes can only be made on materials with a low electromechanical coupling coefficient and therefore have a narrow relative bandwidth (no more than 0.15%). Also, NCRF filters require matching elements. The transfer function of filters on longitudinal resonant modes exhibits a shoulder (at a level of –10…–15 dB in filters with a two-pole input admittance) on its high-frequency slope. Therefore, these filters must include two or three sections connected in series, and the number of poles in their input admittance increases to four or six, respectively.

A technique for decreasing the shoulder level in the DMSF filter transfer function by connecting a capacitive element between the input and output of the filter has been proposed and experimentally verified in [4]. An approximate theoretical analysis of such a filter and a comparison between the theoretical and experimental results were reported in [5]. However, this technique not only decreases the shoulder level but also degrades the overall out-of-band suppression.

In this paper, the modified equations for coupled modes [1] are used to construct a theory of filters operating with longitudinal resonant modes. Theoretical results are compared with experimental data. A topological method is proposed to remove the shoulder in the DMSF filter transfer function by using an SAW-resonator ladder filter on the same substrate instead of the second section of the filter.

MODIFIED COM METHOD

The conventional COM theory (see, e.g., [6]), which derives and solves inhomogeneous differential equations, involves too complex a procedure for SAW device analysis. In the framework of this theory, it is difficult to include such factors as the variable period of the structure, apodization, and nonuniform charge distribution over the electrodes. All these factors are easy to take into account with the modified COM method, which deals with an elementary unit of the structure (an electrode of the interdigital transducer (IDT) or of the reflecting structure). This method also allows for the use of more complex models of the structure. The parameters of the SAW structure component (the IDT, reflecting structure, or their arbitrary combination) are calculated by multiplying the P matrices of individual electrodes (as is usually done in the well-known theory of two-port networks with scattering matrices).

Consider an SAW structure in the form of arbitrarily spaced alternating-polarity electrodes overlapping each other. Assume that the source of a signal with an amplitude U_0 is on the left. Consider the Kth electrode of the IDT (Fig. 1a). Let $R(Z, \omega)$ and $S(Z, \omega)$ be two coupled plane waves with a wave number κ propagating through the IDT in the positive and negative Z directions, respectively. We write these homogeneous plane waves as

$$R(Z, \omega) = R(\omega)\exp(-j\kappa Z), \quad (1)$$

1063-7842/02/4711-1427$22.00

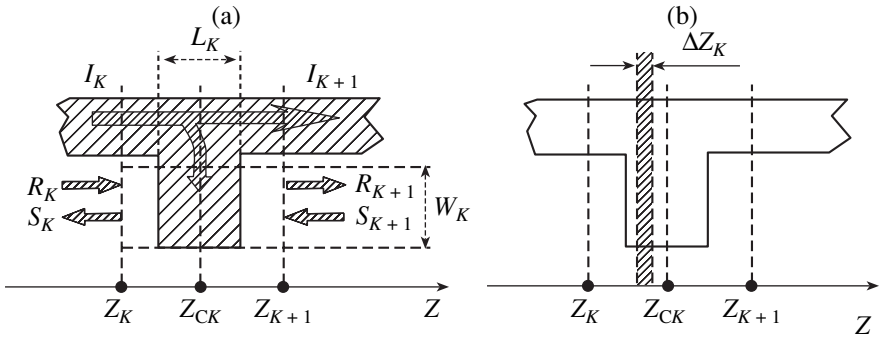


Fig. 1. *K*th electrode of an IDT.

$$R(Z, \omega) = S(\omega)\exp(+j\kappa Z), \tag{2}$$

where $R(\omega)$ and $S(\omega)$ are the complex amplitudes of the waves.

Let a wave incident on the *K*th electrode from the left be $R_K(R, \omega)$ and that from the right, $S_{K+1}(Z, \omega)$. Then, with allowance for reflection, transmission, and transformation with a coefficient ξ_K, the complex amplitudes of the transmitted waves are given by

$$\begin{aligned} S_K(\omega) &= r_K\eta_{1K}\exp[-j(\kappa_E-\kappa_0)p_K]R_K(\omega) \\ &+ \eta_{1K}(1-|r_K|^2)^{1/2}\exp(-j(\kappa_E-\kappa_0)p_K]S_{K+1}(\omega) \\ &+ \xi_K(\omega)\eta_{2K}\exp[-j(\kappa_E-\kappa_0)p_K/2]U_0, \end{aligned} \tag{3}$$

$$\begin{aligned} R_{K+1}(\omega) &= \eta_{1K}(1-|r_K|^2)^{1/2}\exp[-j(\kappa_E-\kappa_0)p_K]R_K(\omega) \\ &+ r_K\eta_{1K}\exp[-j(\kappa_E-\kappa_0)p_K]S_{K+1}(\omega) \\ &+ \xi_K(\omega)\eta_{2K}\exp[-j(\kappa_E-\kappa_0)p_K/2]U_0, \end{aligned} \tag{4}$$

where r_K is the complex coefficient of reflection from the *K*th electrode, κ_E is the SAW effective wave number, $\kappa_0 = 2\pi/p_K$, $p_K = Z_{K+1} - Z_K$, ξ_K is the SAW transformation coefficient at the *K*th electrode, $\eta_{1K} = W_{1K}/W_0$, $\eta_{2K} = W_{2K}/W_0$, W_0 is the maximal aperture, W_{1K} is the overlap of adjacent electrodes, and $W_{2K} = W_0$ if dead electrodes are used and $W_{2K} = W_{1K}$ otherwise.

The phase factors by the terms responsible for wave reflection (transformation) define the phase advance within the interval between the reflection (transformation) center and the respective boundary: Z_K for $S_K(\omega)$ and Z_{K+1} for $R_K(\omega)$. The SAW reflection (transformation) center is assumed to be at the electrode center. We calculate the effective wave number as $\kappa_E = 2\pi/\lambda_E = \omega/[V_0 + L_K(V_M - V_0)/p_K] - j\alpha$, where V_0 is the SAW velocity on the free surface, V_M is the SAW velocity under the metallized surface, and α is the SAW total loss in the electrode structure per unit length.

The variation of the current in the IDT conductor is due to the transformation of the forward and backward waves and the voltage drop across the electrode capacitance:

$$\begin{aligned} \Delta I_K(\omega) &= I_K(\omega) - I_{K+1}(\omega) \\ &= 2\xi_K(\omega)\exp[-j(\kappa_E-\kappa_0)p_K/2]R_K(\omega) \\ &+ 2\xi_K(\omega)\exp[-j(\kappa_E-\kappa_0)p_K/2]S_K(\omega) \\ &+ j\omega(C_2/2)U_0. \end{aligned} \tag{5}$$

Consider the terms associated with SAW transformation on the IDT electrode (Fig. 1b) with regard for the fact that the excitation is distributed. We assume that the efficiencies of the direct and inverse transformations are equal, i.e., that the transformation is reciprocal. Let the current distribution $J(Z)$ over the IDT electrode surface be given. Assume that the mechanisms of SAW transformation by a small element ΔZ_K of the electrode and by the electrode as a whole are the same. Then, adding all the contributions to SAW transformation over the electrode width relative to its center Z_C and passing to the limit $\Delta Z_K \longrightarrow 0$, we obtain

$$\xi_K = G_a \int_{-L_K/2}^{L_K/2} J(Z)\exp[-j(\omega/V_M - \kappa_0)Z]dZ, \tag{6}$$

where G_a is the radiation acoustic admittance at the synchronism frequency (see, e.g., [6]).

The self-consistent analysis of the current distribution $J(Z)$ over the electrode surface, i.e., the analysis with allowance for edge effects, finite IDT length, and back response of the piezoelectric material, is given in our earlier paper [7]. Relationships (3)–(5) can be written in matrix form:

$$\begin{vmatrix} S_K(\omega) \\ R_{K+1}(\omega) \\ \Delta I_K(\omega) \end{vmatrix} = \begin{vmatrix} P(1,1) & P(1,2) & P(1,3) \\ P(2,2) & P(2,2) & P(2,3) \\ P(3,1) & P(3,2) & P(3,3) \end{vmatrix} \begin{vmatrix} R_K(\omega) \\ S_{K+1}(\omega) \\ U_0 \end{vmatrix}. \tag{7}$$

Now, the *P* matrix of the IDT as a whole can be found as the product of *P* matrices for each of the electrodes.

Applying the system of equations (3)–(5) with arbitrary coefficients to two series-connected SAW structures, we can easily obtain the components of the resultant P matrix:

$$P^{(r)}(1,1) = P^{(1)}(1,1) + P^{(1)}(1,2)P^{(2)}(1,1)P^{(1)}(2,1)/P_0; \quad (8)$$

$$P^{(r)}(1,2) = P^{(1)}(1,2)P^{(2)}(1,2)/P_0; \quad (9)$$

$$P^{(r)}(1,3) = P^{(1)}(1,3) + P^{(1)}(1,2)[P^{(2)}(1,3) + P^{(2)}(1,1)P^{(1)}(2,3)]/P_0; \quad (10)$$

$$P^{(r)}(2,1) = P^{(1)}(2,1)P^{(2)}(2,1)/P_0; \quad (11)$$

$$P^{(r)}(2,2) = P^{(2)}(2,2) + P^{(2)}(2,1)P^{(1)}(2,2)P^{(2)}(1,2)/P_0; \quad (12)$$

$$P^{(r)}(2,3) = P^{(2)}(2,3) + P^{(2)}(2,1)/P^{(1)}(2,3) + P^{(2)}(1,3)P^{(1)}(2,2)]/P_0; \quad (13)$$

$$P^{(r)}(3,1) = P^{(1)}(3,1) + P^{(1)}(2,1)[P^{(1)}(3,1) + P^{(2)}(1,1)P^{(1)}(3,2)]/P_0; \quad (14)$$

$$P^{(r)}(3,2) = P^{(1)}(3,2) + P^{(2)}(1,2)[P^{(1)}(3,2) + P^{(1)}(2,2)P^{(2)}(3,1)]/P_0; \quad (15)$$

$$P^{(r)}(3,3) = P^{(1)}(3,3) + P^{(2)}(3,3) + \{P^{(1)}(3,2)[P^{(2)}(1,3) + P^{(2)}(1,1)P^{(1)}(2,3)] + P^{(2)}(3,1)[P^{(1)}(2,3) + P^{(1)}(2,2)P^2(1,3)]\}/P_0, \quad (16)$$

where $P_0 = 1 - P^{(2)}(1,1)P^{(1)}(2,2)$ and the superscripts r, 1, and 2 refer to the resultant P matrix, the P matrix of the left-hand SAW structure, and the P matrix of the right-hand SAW structure, respectively. Here, a single electrode or group of electrodes with a known P matrix can be viewed as an SAW structure.

The procedure for calculating the components of the SAW structure's P matrix as the product of the corresponding components in accordance with relationships (8)–(16) can symbolically be written as

$$P(l,m) + \prod_{n=1}^{N} \mathbf{F}\{P_n(l',m')\}, \quad (17)$$

where N is the number of the last electrode in the SAW structure and the function $\mathbf{F}\{P_n(l', m')\}$ implies the consecutive calculation of the products of the matrix components according to (8)–(16).

The above relationships allow one to calculate the input admittance of the IDT (as a part of a filter or resonator) with arbitrarily varying electrode spacings and widths with regard for the actual current (charge) distribution over the electrode surface. Note that the IDT input admittance is defined by the element $P(3, 3)$ of the resultant P matrix.

ANALYSIS OF FILTERS BASED ON LONGITUDINAL RESONANT MODES BY THE MODIFIED COM METHOD

The configuration of a DMSF filter [3] is illustrated in Fig. 2. The filter has one input transducer ITD-A and two output transducers IDT-B and IDT-C connected in parallel. Note that the filter can also use one output transducer instead of two. To provide the optimal excitation of the resonance modes in the resonator, the IDT-B and IDT-C transducers are completed with reflecting arrays RA-B and RA-C at their edges.

We will analyze the DMSF filter by the modified COM method. The acoustoelectric equivalent circuit of the DMSF filter is constructed by representing each of the IDTs as a device with two electrical and four acoustic inputs (outputs), as shown in Fig. 3. The potentials U^A, U^B, and U^C are applied to the transducers IDT-A, IDT-B, and IDT-C, through which the currents I^A, I^B, and I^C, respectively, flow. Acoustic waves with complex amplitudes R^{A1}, R^{B1}, and R^{C1} are incident on the transducers IDT-A, IDT-B, and IDT-C from the left. The complex amplitudes of the respective reflected acoustic waves are S^{A1}, S^{B1}, and S^{C1}. The complex amplitudes of waves incident from the right are S^{B2}, S^{A2}, and S^{C2}; the complex amplitudes of the respective reflected waves are R^{B2}, R^{A2}, and R^{C2}.

Writing equations like Eqs. (3)–(5) for the Kth electrode of the IDT-A and calculating the products of the matrix components according to (8)–(16), we find the components of the matrix P^A for the IDT-A transducer:

$$\begin{vmatrix} S^{A1}(\omega) \\ R^{A2}(\omega) \\ I^A(\omega) \end{vmatrix} = \begin{vmatrix} P^A(1,1) & P^A(1,2) & P^A(1,3) \\ P^A(2,2) & P^A(2,2) & P^A(2,3) \\ P^A(3,1) & P^A(3,2) & P^A(3,3) \end{vmatrix} \begin{vmatrix} R^{A1}(\omega) \\ S^{A2}(\omega) \\ U^A \end{vmatrix}. \quad (18)$$

The elements of the matrices P^B and P^C for the transducers IDT-B and IDT-C, respectively, will be calculated with regard for the reflecting arrays RA-B and RA-C located on the left and right. Let us denote the

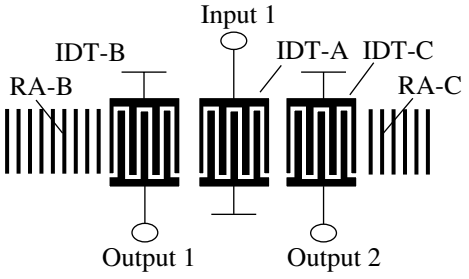


Fig. 2. Layout of the filter on longitudinal resonant modes.

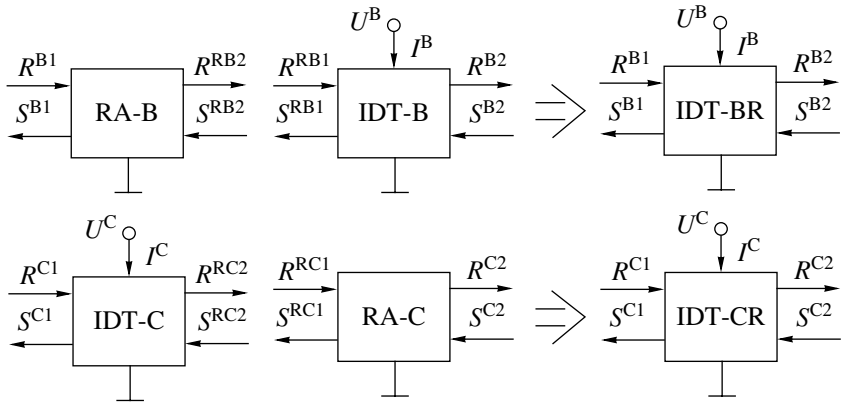


Fig. 3. Equivalent acoustoelectric circuit of the DMSF filter transducers.

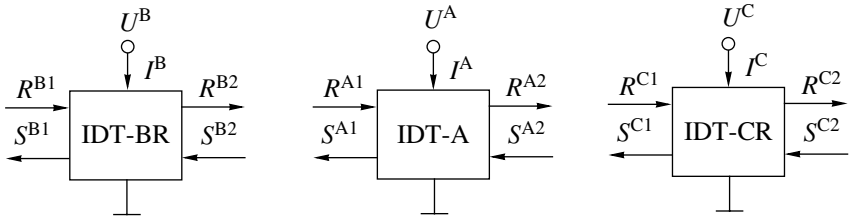


Fig. 4. Equivalent acoustoelectric circuit of a DMSF filter.

IDT-B transducer with the reflecting array RA-B as IDT-RB and the IDT-C transducer with the reflecting array RA-C as IDT-RC. The principle of finding the resultant P matrices for the IDT-RB and IDT-RC transducers is illustrated in Fig. 3. When calculating the products of the matrix components according to formulas (8)–(16), we take into account that, for the reflecting arrays, $\xi_K = 0$, $U = 0$, and $C_2 = 0$. Hence, the matrix elements $P(1, 3)$, $P(2, 3)$, $P(3, 1)$, $P(3, 2)$, and $P(3, 3)$ equal zero.

It should also be noted that the P matrix of the gap between the IDT and the reflecting array or between two IDTs can be found from formulas (3)–(5) with $\xi_K = 0$, $U_0 = 0$, $C_2 = 0$, $r_K = 0$, and $W_K = 0$. In this case, only the elements $P(1, 2)$ and $P(2, 1)$ of the P matrix will be nonzero. From Fig. 3, it is clear that $R^{RB1} = R^{RB2}\exp(-j\kappa d)$, $S^{RB2} = S^{RB1}\exp(j\kappa d)$, $R^{RC1} = R^{RC2}\exp(-j\kappa d)$, and $S^{RC2} = S^{RC1}\exp(j\kappa d)$, where d is the spacing between the reflecting array and the IDT. Then, the components of the P^{BR} matrix, which describes the IDT-B transducer with the reflecting array RA-B, can be written as

$$\begin{vmatrix} S^{B1}(\omega) \\ R^{B2}(\omega) \\ I^B(\omega) \end{vmatrix} = \begin{vmatrix} P^{BR}(1,1) & P^{BR}(1,2) & P^{BR}(1,3) \\ P^{BR}(2,2) & P^{BR}(2,2) & P^{BR}(2,3) \\ P^{BR}(3,1) & P^{BR}(3,2) & P^{BR}(3,3) \end{vmatrix} \begin{vmatrix} R^{B1}(\omega) \\ S^{B2}(\omega) \\ U^B \end{vmatrix}. \quad (19)$$

The components of the P^{CR} matrix, which describes the IDT-C transducer with the reflecting array RA-C, can be calculated in a similar manner:

$$\begin{vmatrix} S^{C2}(\omega) \\ R^{C1}(\omega) \\ I^C(\omega) \end{vmatrix} = \begin{vmatrix} P^{CR}(1,1) & P^{CR}(1,2) & P^{CR}(1,3) \\ P^{CR}(2,2) & P^{CR}(2,2) & P^{CR}(2,3) \\ P^{CR}(3,1) & P^{CR}(3,2) & P^{CR}(3,3) \end{vmatrix} \begin{vmatrix} R^{C2}(\omega) \\ S^{C1}(\omega) \\ U^C \end{vmatrix}. \quad (20)$$

With the above transformations taken into account, the equivalent circuit of an DMSF filter can finally be represented as shown in Fig. 4.

Equations (18)–(20), which correspond to the equivalent acoustoelectrical circuit of Fig. 4, describe the complex amplitudes of the incident and reflected waves, as well as the complex amplitudes of the currents that flow through the transducers. Note that $R_{B1} = 0$, $S_{C2} = 0$, $R_{A1} = R_{B2}\exp(-j\kappa d_1)$, $S_{B2} = S_{A1}\exp(j\kappa d_1)$, $R_{C1} = R_{A2}\exp(-j\kappa d_1)$, and $S_{A2} = S_{C1}\exp(j\kappa d_1)$, where d_1 is the spacing between the IDT-A and IDT-BR and d_2 is that between the IDT-A and IDT-CR.

Solving Eqs. (18)–(20) for unknown currents, one obtains

$$\left\{\begin{aligned} I^A &= U^A Y^A + U^B Y^{AB} + U^C Y^{AC} \\ I^B &= U^A Y^{BA} + U^B Y^B + U^C Y^{BC} \\ I^C &= U^A Y^{CA} + U^B Y^{CB} + U^C Y^C \end{aligned}\right\}, \quad (21)$$

where Y^{A}, Y^{AB}, Y^{AC}, Y^{BA}, Y^{B}, Y^{BC}, Y^{CA}, Y^{CB}, and Y^{C} have the meaning of admittances and can be found from Eqs. (18)–(20).

The equivalent electric circuit of the filter that corresponds to the system of equations (21) is shown in Fig. 5. Using the equivalent circuit and the admittances that enter into system (21), one can calculate the filter's *S* parameters.

The filter of configuration shown in Fig. 2 and the equivalent circuit in Fig. 5 provide a low out-of-band suppression (about –30 dB at a shoulder level of –10…–15 dB); therefore, in applications, one should use no less than two such sections connected in series.

IMPLEMENTATION OF THE DMSF FILTER

In order to experimentally verify the theory proposed above, we designed and fabricated a filter with a center frequency of 95.5 MHz and a bandwidth of 3.5 MHz. The equivalent electrical circuit of the filter with two sections connected in series is shown in Fig. 6. Note that the filter's input admittance has four poles.

In accordance with the bandwidth specified, 36° $LiTaO_3$ was chosen as the substrate material [1]. The basic design parameters of the filter were as follows: the IDT aperture was $45\lambda_0$; the half-periods of the transducers IDT-A, IDT-B, and IDT-C were 10.3 μm; and the transducers IDT-A, IDT-B, and IDT-C and the reflecting arrays RA-B and RA-C consisted of 39, 27, 27, 130, and 90 electrodes, respectively. The electrodes of the transducers and the reflecting arrays represented four-layer structures. First, a 0.01-μm-thick vanadium sublayer was applied directly on the piezoelectric substrate by vacuum magnetron deposition. Then, a 1.15-μm-thick aluminum layer, 0.01-μm-thick vanadium sublayer, and a 0.5-μm-thick copper layer were consecutively deposited. When calculating the filter frequency responses, we used the function $r_K(h)$ and the electromechanical coupling coefficient squared $k^2(h)$ (it is needed when calculating G_a), which were taken from [8]. The loss α in the structure was assumed to be equal to 0.005 dB/λ.

The transfer function (*S*21) of the filter built into a 50-Ω transmission line that was calculated by the above method is plotted in Fig. 7a; that measured by a K4-37 instrument, in Fig. 7b. The theoretical minimal losses within the passband are 0.8 dB, and the measured value is 2.5 dB. The difference in the theoretical and measured minimal losses is apparently because our calculations neglected the ohmic loss in the electrodes. The irregular run of the experimental function near the shoulder is due to the inaccurate (within 1.5–2.0 μm) positioning of the photomask features.

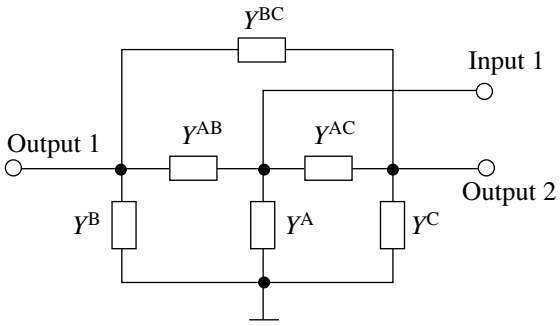


Fig. 5. Equivalent electric circuit of a DMSF filter.

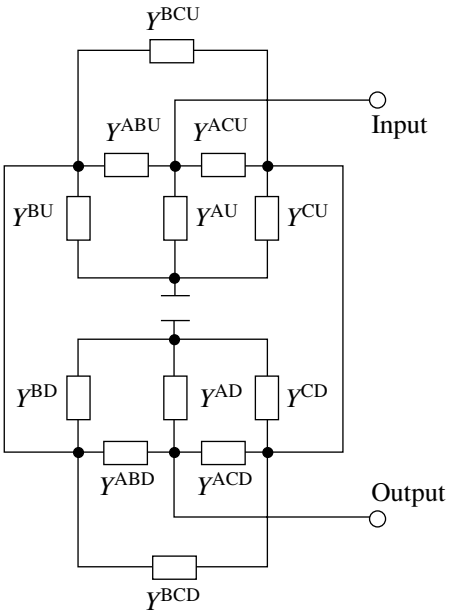


Fig. 6. Equivalent electric circuit of a two-section DMSF filter used in GSM devices.

A HYBRID DMSF–LADDER FILTER

A disadvantage of the DMSF filter is the shoulder at a level of about –30 dB on the high-frequency slope of its transfer function. A method has been proposed [4, 5] to mitigate this disadvantage by connecting a capacitive element between the input and output. The introduction of capacitive feedback, however, not only enhances the out-of-band suppression in the shoulder region but also degrades the overall out-of-band suppression outside the shoulder. Below, we propose a topological method for removing the shoulder by replacing the second section of the DMSF filter by an SAW-resonator-based ladder filter. The total number of poles in the input admittance remains equal to four. Figure 8a illustrates the layout of a hybrid DMSF–ladder filter, and Fig. 8b shows its equivalent electric circuit. As it has the same number of poles in its input admittance as the filters proposed earlier [3–5], the hybrid filter exhibits a higher out-of-band suppression, in particular, in the

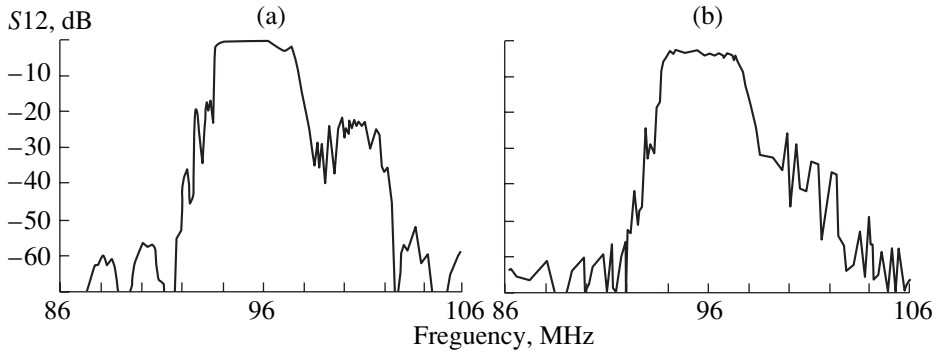


Fig. 7. (a) Calculated and (b) measured transfer functions of the DMSF filter.

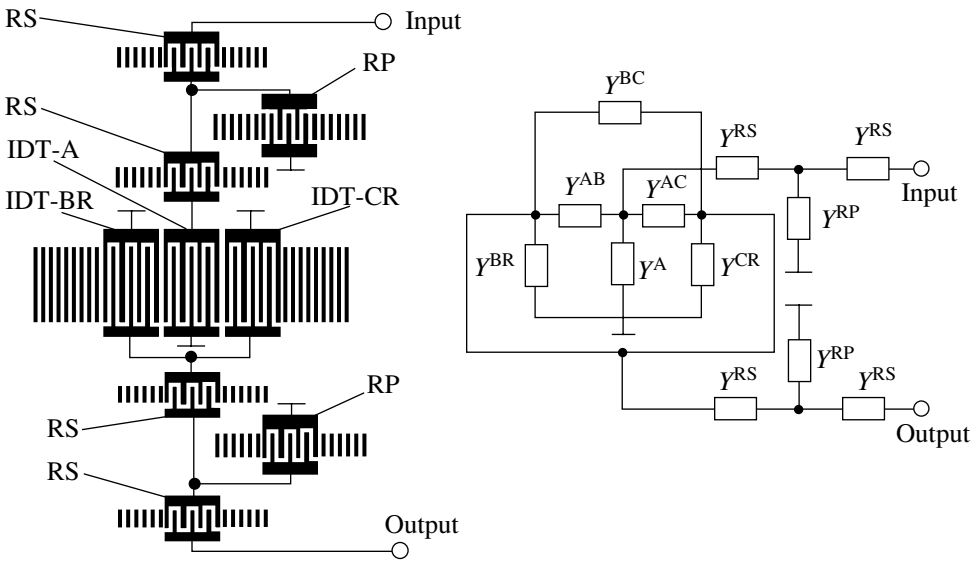


Fig. 8. Layout and equivalent electric circuit of the hybrid DMSF–ladder filter with enhanced out-of-band suppression.

shoulder region. The geometry of the resonators of the ladder filter was chosen by the ladder filter synthesis method proposed in [1]. To obtain the optimal frequency response of the hybrid filter, the passband of the ladder filter must coincide with that of one DMSF section and the input (output) impedance of the ladder filter must be matched to the input (output) impedance of the DMSF filter.

As a piezoelectric material, 36° $LiTaO_3$ was chosen. The basic geometrical parameters of the DMSF stage were as follows: the aperture was $65\lambda_0$; the half-periods of the IDT-A, IDT-B, and IDT-C transducers were 1.0 μm; and the transducers IDT-A, IDT-B, and IDT-C and reflecting arrays RA-B and RA-C consisted of 39, 27, 27, 125, and 85 electrodes, respectively. The electrodes were made of a 0.34-μm-thick aluminum film. The geometrical parameters of resonators of the ladder section of the hybrid filter were as follows: the transducer of the resonator from the parallel arm (RP) had a period 0.975 times that of the IDT-A, consisted of 151 electrodes, and had an aperture of $41\lambda_0$. The transducer of the resonator from the series arm (RS) had a period 1.018 times that of the IDT-A, consisted of 157 electrodes, and had an aperture of $21\lambda_0$. The reflecting arrays of the RP and RS resonators had 120 electrodes. The ratio between the reflecting array period and the transducer half-period was 1.015 for the RP resonator and 0.985 for the RS resonator. The losses α in the structure were assumed to be 0.02 dB/λ. Note that the whole structure was integrated on a single substrate.

Figure 9 plots the transfer functions of a 947-MHz filter (curve *1*) commonly used in GSM devices, which consists of two sections like those in Fig. 2 connected in series. Curve *2* shows a similar filter in which one of the sections is removed and the input and output of the remaining section are completed with T-section SAW-resonator-based ladder filters. Both curves were calculated by the above method. As we see from Fig. 9, the shoulder suppression is improved by 20 dB. The out-of-band suppression in the neighborhood of the passband on its low-frequency side is also improved.

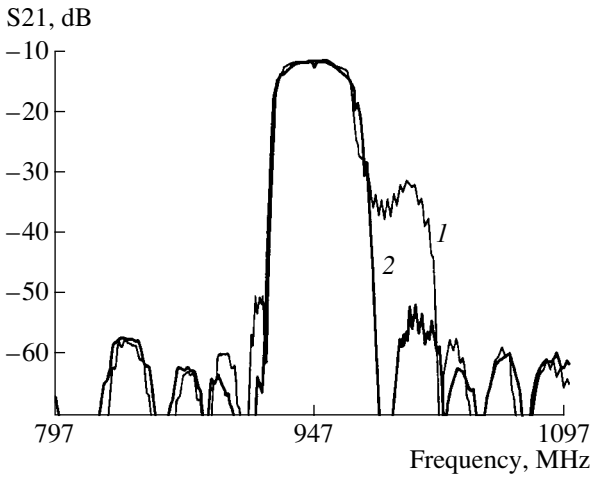


Fig. 9. Transfer function of (*1*) a usual DMSF filter employed in GSM devices and (*2*) the hybrid DMSF--ladder filter illustrated in Fig. 8

Similar results can be obtained with Π-section ladder filters instead of T-section ones or with a combination of the two.

CONCLUSION

Owing to the combination of the properties of ladder filters and filters on longitudinal resonant modes, a hybrid SAW-resonator-based filter proposed in this paper features good out-of-band suppression (higher than 50 dB) throughout the frequency range: both near the passband and at a significant offset from the center frequency.

The filter on longitudinal resonant modes incorporated into the hybrid filter may consist of both two and three IDTs.

The ladder part of the hybrid filter can employ sections of various types. For optimal design parameters, the difference in the passband insertion loss ripple and in the voltage standing-wave ratio is insignificant. However, according to our analysis, the optimum ladder section at the input and output of the DMSF filter is a T section.

REFERENCES

1. V. F. Dmitriev, Zh. Tekh. Fiz. **72** (8), 95 (2002) [Tech. Phys. **47**, 1024 (2002)].
2. V. F. Dmitriev and N. P. Osipova, in *Proceedings of the International Forum on Wave Electronics and Its Application, St. Petersburg, 2000*, p. 360.
3. T. Morita, Y. Watanabe, N. Tanaka, *et al.*, Proc.-IEEE Ultrason. Symp. 95 (1992).
4. S. Beaudin, S. Damphousse, and T. Cameron, Proc.-IEEE Ultrason. Symp., 389.
5. C. K. Campbell and P. J. Edmondson, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **48**, 1298 (2001).
6. S. V. Birykov, G. Martin, V. G. Polevoi, *et al.*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 612 (1995).
7. V. F. Dmitriev, Zh. Tekh. Fiz. **72** (9), 93 (2002) [Tech. Phys. **47**, 1158 (2002)].
8. U. Rosler, D. Cohrs, and A. Dietz, Proc.-IEEE Ultrason. Symp. 247 (1995).

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1434–1439. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 90–95.

RADIOPHYSICS

Multiplication of the Drive Signal Frequency in a Relativistic Microwave Amplifier with a Rod Slow-Wave Structure

A. S. Shlapakovskiĭ*, E. Schamiloglu, and I. I. Grushin***

** Nuclear Physics Institute, Tomsk Polytechnical University, pr. Lenina 2a, Tomsk, 634050 Russia*

*** University of New Mexico, Albuquerque, NM 87131, USA*

e-mail: shl@npi.tpu.ru

Received March 25, 2002

Abstract—A theory is developed for the one-dimensional nonlinear multimode simulation of a Cerenkov maser of special configuration (antenna–amplifier [1, 2]) where a tubular relativistic electron beam propagates along a circular waveguide with a dielectric rod. The operating mode is the fundamental azimuthally asymmetric HE_{11} mode. Harmonics at the input signal frequency can be amplified because they fall into the amplification bands for higher modes. At certain parameters, the output power at the second or third harmonic may considerably exceed that of the amplified signal at the fundamental frequency. The powers of output harmonics can be efficiently controlled by varying the point of electron beam extraction from the interaction region, as well as by varying the input frequency or by switching the polarization of the input signal from linear to circular. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Consider a relativistic electron beam propagating along a surface-wave antenna. Then, an antenna feed signal will be amplified because of Cerenkov interaction. Such a hybrid of a surface-wave antenna and a traveling-wave tube (TWT) is referred to as an antenna–amplifier [1, 2]. In the case of a cylindrical antenna–amplifier (Fig. 1), a tubular beam "flows around" a dielectric rod antenna. The beam propagation is provided by an external coaxial conductor, onto which the beam is dumped with a guiding magnetic field. The rf drive signal enters into the interaction region through a feed waveguide, which, at the same time, serves as the hollow cathode holder of the linear induction accelerator (LIA) module. The outside of the LIA cathode holder is at the ground potential; therefore, it can be directly connected to the waveguide output of an rf drive source. The operating mode in this system is the fundamental azimuthally asymmetric HE_{11} mode, which is unusual for TWTs. In [1, 2], the basic idea and attractive properties of the antenna–amplifier design have been discussed and a linear theory for the general case (without azimuthal symmetry) has been developed. In this paper, an additional important feature, namely, the multiplication of the input signal frequency, resulting from nonlinear analysis in the multimode regime is reported.

Even from the earlier numerical solution of the dispersion relation for various eigenmodes of the system [1, 2], one could see that frequencies that are multiples of frequencies from the amplification band of the fundamental HE_{11} mode may fall in the amplification bands of higher modes. This implies that the following mechanism of beam–electromagnetic field interaction is possible. The input signal at the fundamental mode induces a high-frequency (HF) current in the beam at a frequency ω. The nonlinear interaction generates the nth harmonic of the HF current, which excites a higher mode of frequency $n\omega$. In consequence, the amplified signal involves both the fundamental frequency and its harmonics.

The effect of harmonic generation in wide-band TWTs, where the amplification band exceeds an octave, is well known [3, 4]. Recently, this effect has been discussed as applied to plasma-filled TWTs and relativistic BWOs [5]. In these cases, harmonic radiation is considered as stray and its power is several orders of magnitude lower than that at the fundamental frequency. This, however, may not be the case for an

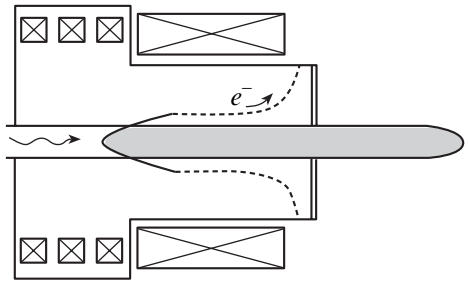


Fig. 1. Antenna–amplifier in the linear induction accelerator module.

1063-7842/02/4711-1434$22.00 © 2002 MAIK "Nauka/Interperiodica"

antenna–amplifier because of its specific operating mode, namely, the asymmetric HE_{11} mode.

Indeed, the rate of growth of the field amplitude along any slow-wave system excited by a monochromatic current of density $\mathbf{j}$ is proportional to the integral $\int_S \mathbf{j}\mathbf{E}_\alpha^{0*}\,dS$ over the system cross section [3], where $\mathbf{E}_\alpha^0$ is the field eigenfunction for an α mode. Hence, if the HF current is induced by the field of a symmetric TM mode, as occurs in conventional TWTs, none of azimuthally asymmetric modes is excited, because the field eigenfunctions and the symmetric current are orthogonal to each other. However, if the input signal is represented by an asymmetric wave (the HE_{11} mode for an antenna–amplifier), any azimuthal harmonic of the HF current may appear in the beam in the large-signal regime as a result of the nonlinearity. Therefore, in the case of an antenna–amplifier, the spectrum of modes that may interact with the beam is not restricted to a family of modes with the same azimuthal index. In this case, the beam coupling with higher modes at harmonic frequencies may be essential because of the specific difference in the radial distributions of the transverse components of the HF field for the fundamental and higher modes. The power transferred by the HE_{11} mode is distributed over the whole cross section of the dielectric rod, whereas in the case of higher modes, the power is concentrated largely near the rod surface. Thus, the longitudinal electric field near the beam for a higher mode may be higher than that for a lowest one of the same power. Calculations of the spatial instability increment [1, 2] show that at certain parameters the HE_{11} mode does not dominate over higher modes, although the frequencies corresponding to instability for other modes are substantially higher. In the large-signal regime, this may result in the fact that the harmonic power turns out to be of the same order of magnitude or even higher than that at the fundamental frequency.

Thus, in contrast to conventional wide-band TWTs, the effect of multiplication of the input signal frequency in an antenna–amplifier may be used to generate an output signal with a spectrum of tunable-power harmonics. In this paper, some aspects of such a control are studied. Below, we present an approach allowing for the simulation of the aforementioned processes and report the results of calculation for the harmonic growth rates at various system parameters.

BASIC EQUATIONS

We will use a set of equations that is similar to that used in the nonlinear theory of TWTs [3] but involves an arbitrary number of modes each of which is synchronous to electrons at a certain frequency that is a multiple of the fundamental frequency. This implies that to each particle there must correspond an equation describing the variation of its phase with respect to each mode α involved in the interaction: $\Theta_\alpha = n_\alpha\omega t - k_{0\alpha}(n_\alpha\omega)z$, where n_α is the harmonic number and $k_{0\alpha}$ is the longitudinal wavenumber. If the beam is infinitely thin, this equation can be written as

$$\frac{d\Theta_\alpha}{dz} = \frac{n_\alpha\omega}{u(z, \varphi, t_0)} - k_{0\alpha}. \tag{1}$$

Here, u is the particle velocity; t_0, the time instant the particle enters into the interaction region; and φ, its azimuthal position. The φ dependence of the particle motion and HF fields is another important feature of our simulation. The excitation equation for a mode with an azimuthal number l_α can be represented in the form

$$\frac{dE_\alpha}{dz} = -\frac{\Phi_\alpha^2(r_b)}{4P_\alpha}\frac{I}{2\pi^2 n_\alpha} \times \int_0^{2\pi} d\varphi \int_0^{2\pi n_\alpha} e^{i[\Theta_\alpha(z, \varphi, \Theta_{0\alpha}) - l_\alpha\varphi]}\,d\Theta_{0\alpha}, \tag{2}$$

where E_α is the complex amplitude of the longitudinal electric field of the mode for a beam radius r_b, the ratio between the squared z component of the field Φ_α and the power P_α characterizes mode–beam coupling, I is the beam current, and $\Theta_{0\alpha} = n_\alpha\omega t_0$ is the particle initial phase with respect to the mode. Eventually, a one-dimensional motion equation disregarding the asynchronous HF field of the space charge is written as

$$\frac{du}{dz} = \frac{e}{mu}\left(1 - \frac{u^2}{c^2}\right)^{3/2} \times \sum_\alpha \mathrm{Re}\{E_\alpha(z)e^{-i[\Theta_\alpha(z, \varphi, \Theta_{0\alpha}) - l_\alpha\varphi]}\}, \tag{3}$$

where e, m, and c are the charge and rest mass of an electron and the velocity of light, respectively.

It is convenient to introduce dimensionless variables

$$\zeta = C_1\frac{\omega}{u_0}z, \quad \nu = \frac{\nu_0}{C_1}, \quad \nu_0 = 1 - \frac{u}{u_0}, \quad \tilde{E}_\alpha = -\frac{1}{C_1^2}\frac{eE_\alpha}{\gamma^3 m\omega u_0}, \tag{4}$$

as well as mismatch parameters and analogues to the Pierce parameter for each mode:

$$\delta_\alpha = \frac{1}{C_1}\left(\frac{k_{0\alpha}u_0}{n_\alpha\omega} - 1\right), \quad C_\alpha^3 = \frac{eI}{\gamma^3 m\omega^2}\frac{\Phi_\alpha^2(r_b)}{4P_\alpha}. \tag{5}$$

Here, u_0 is the electron initial velocity, γ is the corresponding relativistic factor, and C_1^3 is the Pierce parameter for the fundamental mode HE_{11}.

Assuming that $\nu_0 \ll 1$, from (1)–(5) we finally find the set of equations

$$\frac{d\tilde{E}_\alpha}{d\zeta} = \frac{1}{2\pi^2}\frac{C_\alpha^3}{n_\alpha C_1^3}\int_0^{2\pi} d\varphi \int_0^{2\pi n_\alpha} e^{i(\Theta_\alpha - l_\alpha\varphi)} d\Theta_{0\alpha},$$

$$\frac{d\Theta_\alpha}{d\zeta} = n_\alpha(\nu - \delta_\alpha), \qquad (6)$$

$$\frac{d\nu}{d\zeta} = [1 + 2(\gamma^2 - 1)C_1\nu]^{3/2}\sum_\alpha \mathrm{Re}\{\tilde{E}_\alpha e^{-i(\Theta_\alpha - l_\alpha\varphi)}\}.$$

The degeneracy of azimuthally asymmetric modes is correctly taken into account by including modes with both positive and negative azimuthal numbers in set (6). In this case, various polarizations (linear, circular, or elliptic) of the input signal field can be simulated by imposing appropriate initial conditions.

The actual values of the mismatch and Pierce parameters (5) for each mode involved in (6) are determined by numerically solving the dispersion relations for the waveguide with a dielectric rod and by integrating the distribution of the Poynting vector over the system cross section (see, e.g., [6]). In all the calculations considered in the following section, an ensemble of particles (48×48) is uniformly distributed over the intervals $(0, 2\pi)$ of azimuthal angles and initial phases with respect to the HE_{11} mode (for the modes excited at the harmonic frequencies, the initial phases are distributed between 0 and $2\pi n_\alpha$. The initial velocities of all the particles are set to be the same ($\nu = 0$). The power P_α transferred by an α mode through a given cross section is

$$\frac{P_\alpha}{P_b} = \frac{\gamma(\gamma+1)}{4}\frac{C_1^4}{C_\alpha^3}|\tilde{E}_\alpha|^2, \qquad (7)$$

where P_b is the initial kinetic power of the beam.

In the case of asymmetric modes, the power is given by the sum of the contributions from modes with positive and negative azimuthal numbers.

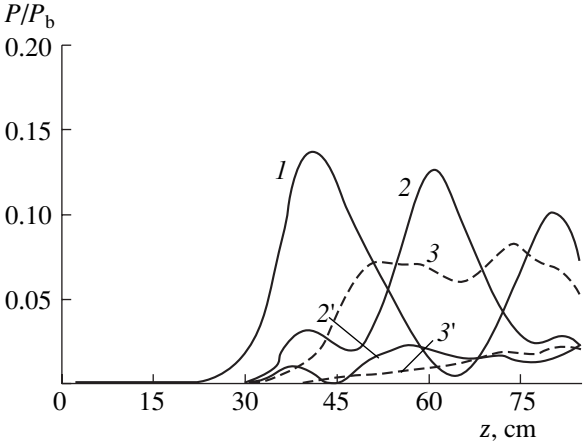


Fig. 2. Microwave power signal at the fundamental and multiple frequencies: (*1*) fundamental frequency, HE_{11} mode; (*2*) second harmonic, TM_{01} mode; (*2*') second harmonic, HE_{21} mode; (*3*) third harmonic, HE_{12} mode; and (*3*') third harmonic, HE_{31} mode. The parameters are as follows: waveguide radius 2 cm, rod radius 1 cm, $\varepsilon = 2.25$, beam–rod gap 2 mm, beam current 1.7 kA, and $\gamma = 1.8$. The input power is ≈13 kW, and the frequency (≈9.4 GHz) corresponds to the zero mismatch of the HE_{11} mode (linear polarization).

RESULTS AND DISCUSSION

As follows from the calculation of the dispersion for various modes of a waveguide with a dielectric rod and from the numerical solution of the dispersion relation for a system with a beam [1, 2], the inclusion of four modes at multiple frequencies in addition to the HE_{11} mode at the fundamental frequency is of significance for multimode simulation. The second harmonic of the HF current may excite the TM_{01} and HE_{21} modes; the third one, the HE_{12} and HE_{31} modes. Therefore, the simulation was performed with these five modes involved in the interaction.

The effect of harmonic generation is obvious at such system parameters when the HE_{11} mode does not dominate over higher modes in terms of the instability increment calculated in the framework of the linear theory [1, 2]. Figure 2 shows the dependences of the powers of the five modes included on the longitudinal coordinate. The geometric parameters, beam energy, and input frequency correspond to the case when the beam is in exact synchronism with the fundamental mode. The structure of the input signal field is time-invariable (linear polarization). Figure 2 also shows that the TM_{01} mode at the second harmonic and the HE_{12} mode at the third harmonic compete with the fundamental HE_{11} mode.

Reasons for such a competition can be illustrated by considering the particle ensemble in the phase space. Figure 3 depicts the phase portrait of the ensemble with the three high-power modes taken into account. The longitudinal coordinate is chosen so as to provide maximum bunching. It is of importance that the harmonic fields here are rather low; therefore, the beam dynamics up to this coordinate is virtually defined only by the signal at the fundamental frequency. As is clearly seen, the beam is bunched basically near two phases of the wave with the input frequency (Fig. 3a), unlike conventional TWTs. Indeed, the azimuthal structure of the input signal field was fixed in this experiment. This implies that the whole ensemble can be divided into two equivalent parts; i.e., each particle has its "twin" at the opposite azimuth, which moves in the same HF field but is of opposite sign. Therefore, bunches at the opposite azimuths form around the phases rotated by π with respect to each other. At the same time, both bunches are in phase with respect to the synchronous wave at the second harmonic. Hence, Fig. 3b shows bunching largely

around one phase. In the case of the symmetric TM_{01} mode, it turns out that this phase slows-down all the particles, whereas in the case of the asymmetric HE_{21} mode, which has almost the same phase velocity and an even higher coupling with the beam, this phase slows-down only that half of the ensemble covering the corresponding azimuths. That is why the TM_{01} mode in this case builds up faster than the HE_{21} one at the second harmonic. For the third harmonic, the situation is more complicated; however, similar arguments can be put forward to explain the faster growth of the HE_{12} mode compared with that of the HE_{31} mode. In addition, the synchronism detuning for the HE_{31} mode is rather large at such parameters.

Return now to Fig. 2. This figure can be viewed as showing the dependences of the powers of the interacting modes on the length of the interaction region. For a beam propagating outside the rod, the point of beam extraction from the interaction space can easily be controlled by varying the length of the guiding magnetic field. As follows from Fig. 2, the shift of the extraction point within 20 cm considerably changes the output spectrum. Indeed, at one length of the interaction region, the fundamental frequency dominates. At another length, the powers of the fundamental, second, and third harmonics are approximately equal. At the third frequency, the second harmonic dominates, whereas the power at the fundamental frequency is the lowest. Thus, not only frequency multiplication but also the control of the powers of various harmonics at the output of an antenna–amplifier are provided.

It is worth noting that the output spectrum can be changed by varying the input frequency without changing the beam transport length. In the linear theory, the amplification band for the HE_{11} mode is wider than those for higher modes. Therefore, multiple frequencies may not fall in the amplification bands of higher modes if the mismatch from the fundamental frequency is large. Figure 4 shows the results of calculation in this case for a positive detuning. The maximum efficiency and the associated longitudinal coordinate increase compared with those in Fig. 2 (which is the well-known behavior), whereas the harmonic powers decrease considerably. The frequency dependences of the powers for all the modes at a fixed interaction length are demonstrated in Fig. 5. Here, the possibility of controlling the harmonic contributions to the output signal by varying the input frequency is clearly seen. The fundamental frequency dominates over the harmonics both at positive and negative (however, with a substantially smaller efficiency) detunings. Note that the input signal frequency is easy to vary.

There is one more way to reduce the power levels at multiple frequencies without varying either the interaction length or the input frequency. It is possible because the growth of the field at harmonics depends on whether the input field structure is fixed or rotates azimuthally. In the case of a rotating structure (the circular

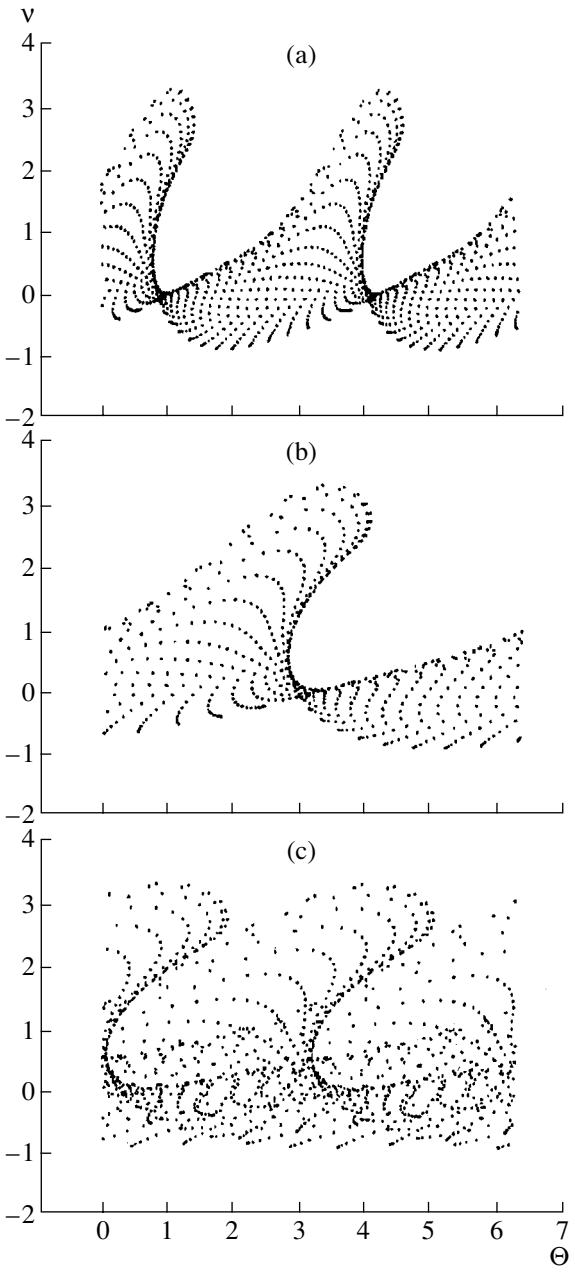


Fig. 3. Configuration of the particle ensemble in the phase space at $z \approx 36$ cm. The parameters are the same as in Fig. 2. The phases are shown with respect to the (a) HE_{11} mode (fundamental frequency), (b) TM_{01} mode (second harmonic), and (c) HE_{12} mode (third harmonic).

polarization of the input wave), the modes with azimuthal numbers that do not coincide with the harmonic number do not grow at all. The result of simulation for associated initial conditions is shown in Fig. 6. Indeed, if the input signal is circularly polarized, all particles, irrespective of their azimuthal position, are "equivalent" on average. Therefore, they are bunched over the same length but at different phases for different azimuths. In other words, a beam transforms into a helix

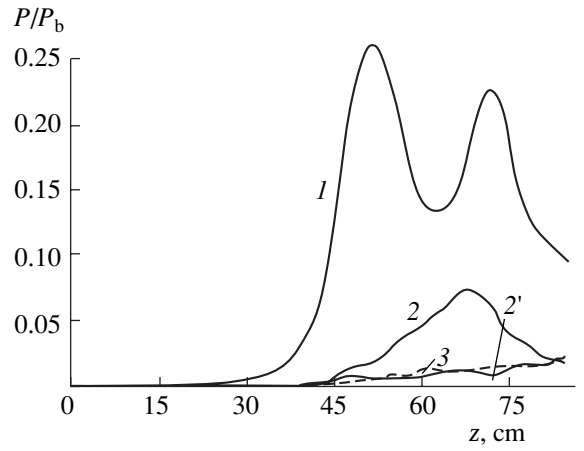


Fig. 4. The same as in Fig. 2 for the input signal frequency 11.94 GHz.

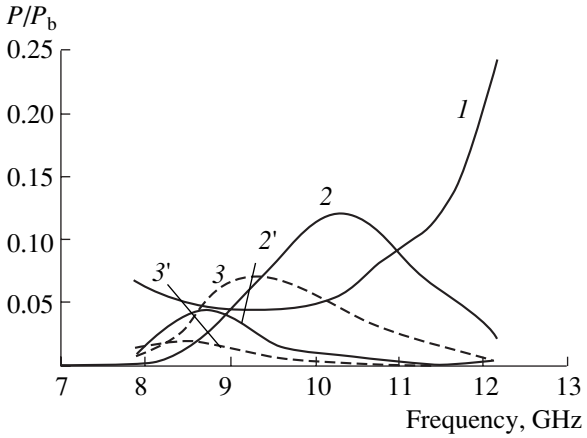


Fig. 5. Microwave power in the cross section at $z = 55$ cm as a function of the input frequency. The parameters and notation are the same as in Fig. 2.

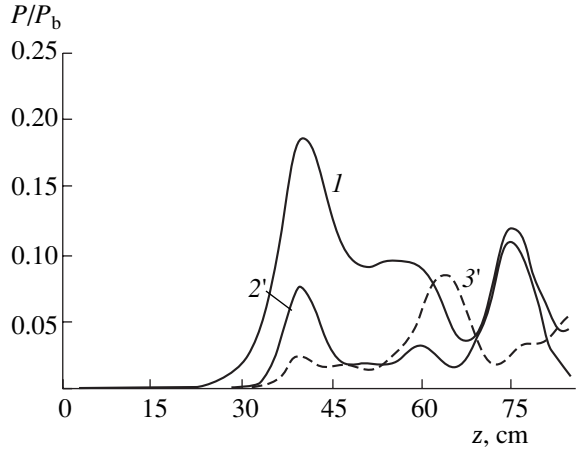


Fig. 6. The same as in Fig. 2 for the circularly polarized input signal.

in the field of the input wave as a result of bunching. Therefore, a wave that has the same phase angular velocity as the fundamental frequency can grow at the harmonic frequency. Such is the HE_{21} mode at the second harmonic and the HE_{31} mode at the third one. Comparing Figs. 6 and 2, one sees that the input spectrum can change if the initial polarization changes from linear to circular.

Note also that, along with the possibility of controlling the harmonic amplitudes in the output spectrum of an antenna–amplifier, there is the possibilily of controlling the phases of harmonics. Calculations show that, if the phase of an input wave changes by $\Delta\Theta$, the phase of any mode involved in the interaction in a given cross section of the interaction region changes by $n_\alpha\Delta\Theta$ depending on the harmonic number.

To conclude, we note that intense harmonic generation is not an inherent property of an antenna–amplifier. If the parameters of the beam and waveguide with a rod are such that the HE_{11} mode dominates in the instability spectrum of the system, the powers at multiple frequencies are very low. This can be exemplified by the set of parameters used in [1, 2], where the dominance of the fundamental mode over higher ones in terms of the linear increment was demonstrated (waveguide radius 2 cm, rod radius 6 mm, $\varepsilon = 5$, beam–rod gap 4 mm, beam current 1 kA, and $\gamma = 1.4$). At these parameters in the amplification frequency range 9.5–9.9 GHz, the power of the second harmonic is lower than that of the fundamental one by at least 20 dB. In this case, the efficiency of an antenna–amplifier is 30%.

CONCLUSION

One-dimensional nonlinear simulation of an antenna–amplifier (a Cerenkov maser with a rod slow-wave structure and the azimuthally asymmetric operating mode), which was performed in view of the possibility of exciting higher modes, showed that the effect of input signal frequency multiplication due to the generation of the second and third harmonics is feasible. The possibilities of controlling the spectral distribution of the output signal were studied. They open up new ways for electronically tuning microwave devices. Such devices, which allow for the variation of the harmonic amplitudes in the radiation spectrum, could have many applications. It would be of importance to confirm this effect by three-dimensional simulation.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (grant no. 00-02-17720) and the NATO Science Commitee (grant no. PST.CLG 975604).

The authors are grateful to G.P. Fomenko for helpful discussions.

REFERENCES

1. A. S. Shlapakovskiĭ, Pis'ma Zh. Tekh. Fiz. **25** (7), 43 (1999) [Tech. Phys. Lett. **25**, 267 (1999)].
2. A. S. Shlapakovski, in *Intense Microwave Pulses VI*, Ed. by Howard E. Brandt; Proc. SPIE **3702**, 108 (1999).
3. L. A. Vaĭnshteĭn and V. A. Solntsev, *Lectures on Microwave Electronics* (Sov. Radio, Moscow, 1973).
4. N. J. Dionne, IEEE Trans. Electron Devices **17**, 365 (1970).
5. G. S. Nusinovich, Y. Carmel, T. M. Antonsen, Jr., *et al.*, IEEE Trans. Plasma Sci. **26**, 628 (1998).
6. Yu. A. Ilarionov, S. B. Raevskiĭ, and V. Ya. Smorgonskiĭ, *Analysis of Corrugated and Partially Filled Waveguides* (Sov. Radio, Moscow, 1980).

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 11, 2002, pp. 1440–1443. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 96–99.

ELECTRON AND ION BEAMS, ACCELERATORS

Mass-Spectroscopic Study of the Diffusion and Solubility of Helium in Submicrocrystalline Palladium

A. N. Zhiganov*, A. Ya. Kupryazhkin*, R. R. Mulyukov, and I. Kh. Bitkulov****

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

*** Institute of Problems of Metal Superplasticity, Russian Academy of Science, ul. Khalturina 39, Ufa, 450001 Bashkortostan, Russia*

e-mail: kupr@dpt.ustu.ru

Received April 9, 2002

Abstract—By the method of helium thermal desorption from submicrocrystalline palladium presaturated in the gaseous phase, the diffusion coefficient D_{eff} and solubility coefficient C_{eff} of helium are measured in the range P = 0–3 MPa and T = 293–508 K. The pressure dependence of C_{eff} flattens at high pressures. At low saturation pressures, the temperature dependences of the diffusion and solubility coefficients may be divided into (1) high-temperature (400–508 K) and (2) low-temperature (293–400 K) ranges described by the exponentials $D_{1,2} = D_0\exp(-E^D_{1,2}/kT)$ and $C_{1,2} = C_0\exp(-E^S_{1,2}/kT)$. The energies of diffusion activation are $E^D_2 = 0.0036 \pm 0.0015$ eV and $E^D_1 = 0.33 \pm 0.03$ eV, and the solution energies are $E^S_2 = -0.025 \pm 0.008$ eV and $E^S_1 = 0.086 \pm 0.008$ eV in the low- and high-temperature ranges, respectively. Mechanisms behind the diffusion and solution of helium are discussed. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The interaction of helium with metals usually attracts interest in the context of reactor material embrittlement (see, e.g., [1]). Most of the works are devoted to the interaction of helium with native and radiation-induced defects in metals, and they study helium release from linearly heated samples presaturated by helium via neutron or α particle irradiation or tritium saturation. As a rule, associated techniques imply the saturation of the samples by helium under nonequilibrium conditions (the concentration of helium is higher than the equilibrium value) and the performance of measurements also under nonequilibrium conditions (defect annealing). In the majority of cases, this prevents the determination of the transport coefficients of helium in a metal for a particular diffusion or solution mechanism.

Of interest are methods for measuring the diffusion and solubility coefficients of helium in polycrystalline metals. In this work, such measurements are carried out at low temperatures on samples of submicrocrystalline palladium, which is a basic component of membranes used for the separation of hydrogen isotopes from each other and from helium. Moreover, studying the properties of metals with a submicrocrystalline structure is of independent interest.

EXPERIMENTAL SETUP AND TECHNIQUE

Samples used were made of 99.99%-pure submicrocrystalline palladium produced under high plastic deformations up to the true logarithmic degree e = 7 by torsion under a quasi-hydrostatic pressure with a Bridgeman anvil [2].

The microstructure was examined with a JEM 2000EX transmission electron microscope (TEM). According to the TEM data, the intense plastic deformation of the samples resulted in a highly dispersed imperfect structure with an average grain size of 150 nm. The density of lattice dislocations was $\approx 3 \times 10^{10}$ cm^{-2}. The samples were plates with a thickness $h = (6.1 \pm 0.3) \times 10^{-3}$ cm, total geometrical surface area $2S = 2.5 \pm 0.1$ cm^2, and weight $m = 90.9 \pm 0.1$ mg.

The solubility and diffusion of helium in palladium were studied in the experiment on helium thermal desorption [3] from samples presaturated in the helium atmosphere at given saturation temperature T and pressure P (Fig. 1). The experimental procedure was as follows. The sample was placed into saturation chamber *9*, which was evacuated to a high vacuum (<10^{-6} Pa) by diffusion pump *13* with nitrogen trap *14*. Then, the chamber was filled with helium to a given pressure P (which is indicated by manometer *8* to an accuracy of 0.05 MPa) and heated by oven *5* to a given temperature T (which is kept constant with an accuracy of 0.1 K). Our setup allowed for the saturation of the samples at high pressures. By cooling volume *10* by liquid nitrogen and subsequently heating it to room temperature or above, the pressure in the saturation chamber can be raised to 40 MPa at a pressure in helium vessel *11* of ~10 MPa.

The helium-saturated samples were quenched by rapid cooling to room temperature and placed into measuring chamber *4*. The saturation time was determined experimentally. With the sample inside, the measuring chamber was evacuated to a high vacuum ($<10^{-6}$ Pa) and heated to the same temperature T; then, the desorption of helium from the sample was measured. According to our calculations, the loss of helium from the sample during quenching, chamber-to-chamber transportation, and heating to the working temperature was negligibly small.

The kinetics of helium evolution from the sample was studied with an MI-1201B mass spectrometer operating in the quasi-static (with respect to helium) regime of evacuation. Ions were detected by a secondary-emission multiplier [4] at a pressure in the measuring chamber that was no higher than 10^{-6} Pa. When studying the desorption of helium, the gas from the measuring chamber was delivered into mass spectrometer *6* portion by portion at regular intervals, with each portion pumped off after the measurements. We checked the residual helium pressure in the degassing cell both before and after the desorption experiment.

The measuring system was calibrated by doubly expanding a certain amount of the gas contained in gauged volume V_1 (*3*). The pressure in this volume was determined by optical manometer *1* scaled from 0 to 1 torr [4].

The desorption curves were processed by solving the second Fick equation for an infinite plate of thickness h in the case of degassing into a vacuum (zero boundary conditions) with the effective diffusion coefficient D_{eff} and the effective solubility C_{eff}. With these assumptions, the total desorption flux of helium versus the degassing time t should fit the dependence

$$J(t) = \frac{8SC_{\text{eff}}D_{\text{eff}}}{h}\sum_{k=0}^{\infty}\exp\left(-\frac{\pi^2(2k+1)^2D_{\text{eff}}}{h^2}t\right). \quad (1)$$

To eliminate the effect of annealing under high-temperature experimental conditions, we repeated the measurements at low temperatures. The diffusion and solubility coefficients were determined during heating and cooling the samples.

Typical data obtained for the highest and lowest saturation pressures and temperatures are plotted in Fig. 2. Throughout the range of pressures and temperatures, the time dependences of the flux logarithm are straight lines within the accuracy of the analysis. Such a behavior corresponds to solution (1) at large times of desorption $t > h^2/(\pi^2 D_{\text{eff}})$ and suggests that we are dealing with a single helium transport mechanism with the effective parameters C_{eff} and D_{eff}. Thus, the diffusion coefficient can be determined from the slope of the dependences $\ln(J) = f(t)$. The accuracy of D_{eff} determination is 7%.

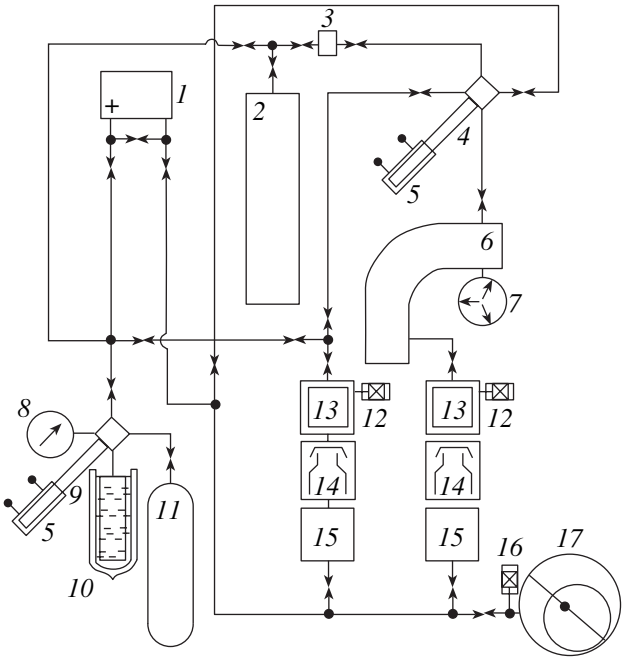


Fig. 1. Experimental setup: (*1*) optical manometer, (*2*) gauged volume V_2, (*3*) gauged volume V_1, (*4*) measuring chamber, (*5*) ovens for heating the chambers, (*6*) MI-1201B mass spectrometer, (*7*) ion-getter pump, (*8*) spring gauge, (*9*) saturation chamber, (*10*) high-pressure chamber with Dewar vessel, (*11*) gas cylinder, (*12*) magnetoionization pressure transducer, (*13*) nitrogen trap, (*14*) diffusion pump, (*15*) forevacuum container, (*16*) PMT-4M thermocouple converter, and (*17*) fore pump.

The solubility of helium in the palladium was found for the completely degassed samples. The result is accurate to within 10%.

RESULTS AND DISCUSSION

The solubility of helium in palladium was measured as a function of the saturation pressure at seven different temperatures. Three of these isotherms are shown in Fig. 3. At all the temperatures, the dependence $C_{\text{eff}} =$

Parameters of helium transport (diffusion and solution) along the grain boundaries of submicrocrystalline Pd

No.	ΔT, K	$D_{\text{eff}0}$, cm^2 s^{-1}	E^D, eV	$C_{\text{eff}0}$, cm^{-3}/(10^5 Pa)	E^S, eV
1	293–400	$(0.98^{+1.1}_{-0.09}) \times 10^{-9}$	0.0036 ± 0.0015	$(0.7^{+0.31}_{-0.2}) \times 10^{15}$	-0.025 ± 0.008
2	400–508	$(1.1^{+0.9}_{-0.5}) \times 10^{-5}$	0.33 ± 0.03	$(1.5^{+0.41}_{-0.3}) \times 10^{16}$	0.086 ± 0.008

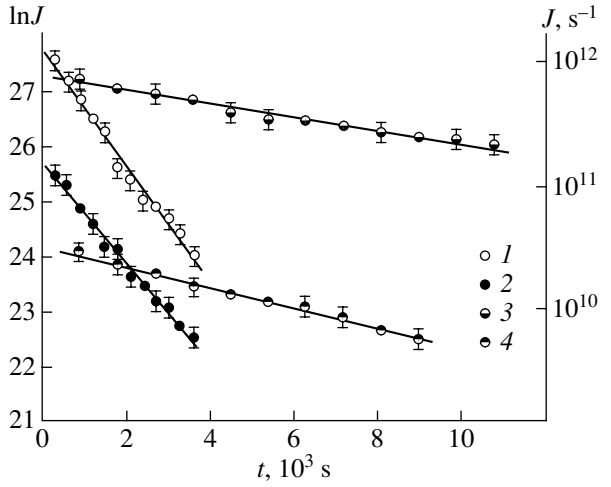


Fig. 2. Typical time dependences of the helium desorption flux at (*1*) T = 508 K, P = 25 MPa; (*2*) T = 508 K, P = 2 MPa; (*3*) T = 293 K, P = 25 MPa; and (*4*) T = 293 K, P = 2 MPa. The solid line is the approximation.

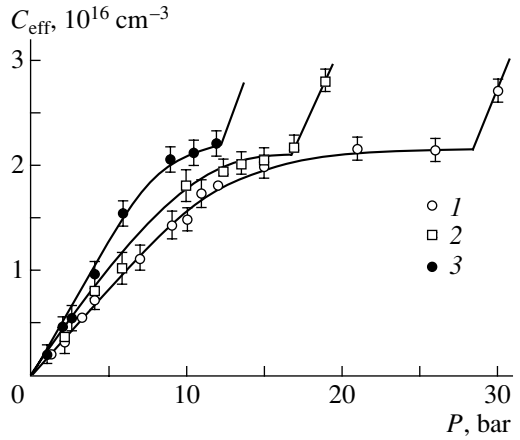


Fig. 3. Solubility of helium vs. saturation pressure for the saturation temperature T = (*1*) 387, (*2*) 433, and (*3*) 508 K. The solid line is the approximation.

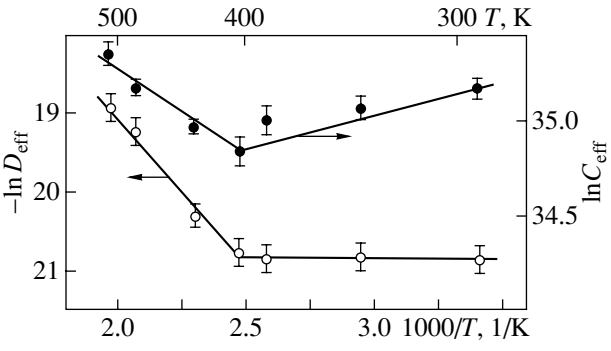


Fig. 4. Temperature dependences of the diffusion coefficient D_{eff} (P = 0.1–0.25 MPa) and solubility coefficient C_{eff} (P = 0.1 MPa) at a low site occupation. The solid line is the approximation.

$f(P)$ has a plateau in accordance with the expression [4]

$$C_{eff}(P) = \frac{C^*\Gamma P}{1 + \Gamma P}, \quad (2)$$

where C^* is the concentration of sites being saturated and Γ is a quantity that is only slightly dependent on pressure.

Throughout the temperature range studied, $C^* = (2.0 \pm 0.2) \times 10^{16}$ cm^{-3}. The higher the temperature, the shorter the plateau and the lower the pressure at which the abrupt rise in the solubility starts.

At low saturation pressures, $P = 10^3$–10^5 Pa (a low occupation of the sites), the temperature dependences of the diffusion and solubility coefficients (Fig. 4) have a clear-cut kink between low-temperature (293–400 K) and high-temperature (400–508 K) sections, which are characterized by the exponential dependences

$$\begin{aligned} D_{eff1,2} &= D_{eff0}\exp(-E^D_{1,2}/kT), \\ C_{eff1,2} &= C_{eff0}\exp(-E^S_{1,2}/kT). \end{aligned} \quad (3)$$

The values of the preexponentials D_{eff0} and C_{eff0}, as well as the activation energies of diffusion and solution $E^S_{1,2}$ and $E^D_{1,2}$, are listed in the table.

Two typical curves $D_{eff} = f(P)$ shown in Fig. 5 for the temperatures T = 403 K and T = 433 K can be divided into three characteristic parts: the range of small pressures, where the diffusion coefficient is independent of the saturation pressure (i.e., of the occupation of the sites) within the experimental accuracy; the transition region of medium pressures, where D_{eff} grows with pressure; and the plateau, where D_{eff} is again constant and pressure independent.

According to our results, which are in line with data provided by the electron–positron annihilation [5] and magnetic susceptibility [6] methods, the palladium samples studied feature a considerably higher concentration of vacancies and vacancy clusters (6–12 vacancies [5]) as compared to that in a usual Pd polycrystal. A nearly twofold increase in the grain size was observed after the annealing of the samples at 473 K [6]. At this annealing temperature, the concentration of the clusters becomes comparable to the detectability of the technique [5], while the concentration of vacancies remains constant up to ≈823 K [6].

The values of D_{eff} and C_{eff} were found to be independent of the annealing temperature up to 508 K and, consequently, of the grain size. The position of the plateau on the dependence $C_{eff} = f(P)$ (i.e., the concentration of the sites being saturated) is also the same for any temperature from the range considered. Therefore, the helium transport detected in the experiment takes place mainly along the grain boundaries. As is evident from curve *3* in Fig. 3, the concentration of sites that are saturated by helium C^* is temperature independent. We assume that, in submicrocrystalline palladium, these

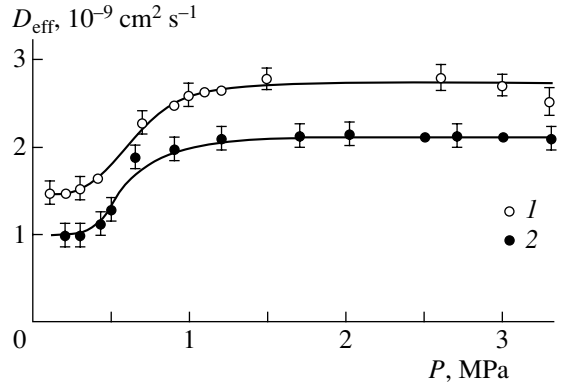


Fig. 5. Diffusion coefficient vs. saturation pressure at T = (*1*) 403 and (*2*) 433 K. Solid line is approximation.

are deformation-induced grain-boundary vacancies and vacancy clusters that are stable against low-temperature annealing.

Defects that fill up at low temperatures have a low solution energy. For the samples under study, such defects may be unannealed vacancy clusters at grain boundaries. The energy of helium solution in the defect is $E_1^S = -0.025 \pm 0.008$ eV, which is comparable to the energy of helium adsorption on a solid surface [7]. In this temperature range, the activation energy of helium grain-boundary diffusion is $E_1^D = 0.0036 \pm 0.0015$ eV. As the temperature rises, the mechanism of helium solution with the energy $E_2^S = 0.086 \pm 0.008$ eV in less complex defects, including vacancies, becomes prevalent. These defects have a higher solution energy and, at the same time, a higher concentration. Diffusion also proceeds over these defects with the activation energy $E_2^D = 0.33 \pm 0.03$ eV.

The growth of D_{eff} with saturation pressure may be explained by the increase in the number of higher mobility He–divacancy and He–He complexes, as well as by the smoothing of the potential relief for diffusing helium as it fills up the sites.

The refinement of mechanisms behind helium diffusion and solution in polycrystals and the effect of vacancies and vacancy clusters on helium transport in polycrystals call for further investigation.

ACKNOWLEDGMENTS

R.R. Mulyukov and I.Kh. Bitkulov are grateful to the Russian Foundation for Basic Research (grant no. 02-01-97924-r2002agidel a) and the Integrated program “Nanocrystals and Supramolecular Systems” of the Russian Academy of Sciences for partial financial support.

REFERENCES

1. A. G. Zaluzhnyĭ and A. L. Suvorov, Zh. Tekh. Fiz. **71** (2), 55 (2001) [Tech. Phys. **46**, 186 (2001)].
2. R. R. Mulyukov and M. D. Starostenkov, Acta Metall. Sin. **13** (1), 301 (2000).
3. A. G. Dudorov and A. Ya. Kupryazhkin, Zh. Tekh. Fiz. **68** (12), 85 (1998) [Tech. Phys. **43**, 1479 (1998)].
4. A. Ya. Kupryazhkin and A. Yu. Kurkin, Fiz. Tverd. Tela (St. Petersburg) **35**, 3003 (1993) [Phys. Solid State **35**, 1475 (1993)].
5. R. Würschum, A. Kübler, *et al.*, Ann. Chim. Sci. Mater. **21**, 471 (1996).
6. A. A. Rempel’, A. I. Gusev, S. Z. Nazarova, *et al.*, Dokl. Akad. Nauk **345**, 330 (1995) [Phys. Dokl. **40**, 570 (1995)].
7. *The Solid–Gas Interface*, Ed. by E. A. Flood (Dekker, New York, 1967; Mir, Moscow, 1970).

Translated by A. Sidorova

Technical Physics, Vol. 47, No. 11, 2002, pp. 1444–1452. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 100–108.

ELECTRON AND ION BEAMS, ACCELERATORS

Reflection of keV-Energy Electrons from Multilayer Surfaces

V. P. Afanas'ev*, A. V. Lubenchenko*, S. D. Fedorovich*, and A. B. Pavolotskiĭ**

** Moscow Power Institute (Technical University), ul. Krasnokazarmennaya 17, Moscow, 111250 Russia*

*** Research Institute of Nuclear Physics, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*

e-mail: lub@phns.mpei.ac.ru

Received February 22, 2002; in final form, May 14, 2002

Abstract—The energy spectra of electrons reflected from multilayer targets are studied theoretically and experimentally. A self-consistent theory of electron reflection from multilayer surfaces is constructed. Simple analytical models of electron reflection that illustrate the feasibility of the depth profiling of multilayer targets are presented. The energy spectra of electrons normally incident on Nb/Si and Nb/Al/Nb/Si targets and reflected from them at an angle of 45° to the normal are taken. A method for the depth profiling of such structures is elaborated. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Electron spectroscopy is one of the most efficient and commonly used method for solid surface analysis [1]. In this work, we use a method that analyzes double-differential spectra of electrons reflected into a given solid angle. Instruments for studying such spectra usually contain an electron probe and an energy analyzer. However, unlike AES and characteristic loss spectroscopy, where similar devices are used, our method does not require a high accuracy. Information on the target is derived from depths on the order of the electron transport length l_{tr}, which is roughly proportional to the beam energy squared. Varying the probing beam energy from 4 to 32 keV, one can penetrate, for example, a copper target to depths from 25 to 800 nm.

Currently available experimental data for energy spectra of electrons with an energy E_0 in the range $1 < E_0 < 100$ keV reflected from homogeneous targets are very scarce and differ quantitatively and even qualitatively. The discrepancy is explained by differences in target purity and energy analyzer resolutions, as well as by an uncertainty in the point-spread function of recording devices. Energy spectra recorded in the 1950s–1990s were smooth curves with a single peak [2–7]. In further studies [8], electron energy spectra had a complex shape of a dome with a peak. Figure 1 compares the spectra obtained by different authors.

The absence of an elastic peak in the energy spectra recorded in [2–4] is due to the fact that those spectra were obtained in a poor vacuum (hence, from the contaminated surface) and the recording devices had a poor resolution. The elastic peak was also not observed in spectra taken in [5–7], because the insufficient surface purity greatly affected the high-energy part of the spectrum. Note the absence of published data for the spectra of electrons reflected from multilayer targets.

The aim of this work is to demonstrate the spectra of electrons reflected from multilayer targets and to construct a computing procedure for restoring the depth distribution of the target composition.

EXPERIMENTAL ENERGY SPECTRA OF ELECTRONS REFLECTED FROM HOMOGENEOUS AND MULTILAYER TARGETS

The stainless-steel vacuum chamber of the experimental setup withstood heating up to 300°C. The preevacuation of the chamber was carried out with a turbomolecular pump. Oil-free evacuation during measurements was accomplished by means of two magnetic-discharge pumps (with the turbomolecular pump cut off), and the residual pressure during the experiment was no higher than 10^{-6} Pa. The residual gas composition in the vacuum chamber was monitored with a mass spectrometer. Carboniferous contaminants were not detected in the course of the experiment. The target surface was precleaned by rinsing in a solution of formic acid and hydrogen peroxide. At the final stage of surface preparation, the surface was decontaminated by ion milling with the use of a defocused argon ion beam of energy 5 keV and current density 1.5 mA/m^2. The target was heated by a defocused electron beam of current 300 μA and accelerating voltage 20 kV. Under these conditions, the target temperature reached 500°C. The design of the setup allowed for the simultaneous action of the probing electron and decontaminating ion

1063-7842/02/4711-1444$22.00

beams. The purity of the target surface was monitored by observing the peak of elastically reflected electrons. The electron-assisted growth of carboniferous films on the target surface after the ion milling and a change in the near-surface composition of the target due to argon ion penetration was not revealed.

The probing electron beam of energy E_0 between 3 and 30 keV was generated by an electron gun. The working electron beam current was kept at a level of 125 μA (the target temperature did not exceed 500°C for these energies of the beam). The beam was directed to the target at a right angle. The target was grounded through a calibrated resistor. The experiment geometry was such that the energy analyzer received electrons reflected from the target surface at an angle $\vartheta = 135°$ (that is, at 45° to the normal to the target).

The energies of reflected electrons were analyzed with a 180° spherical electrostatic energy analyzer with the energy resolution $\Delta E/E_0 = 0.8\%$. The voltage applied to the analyzer plates was programmably varied from 0 to 4 kV. The electrons passing through the analyzer fell into a Faraday cup. Their current (from 10^{-12} to 10^{-10} A) was measured by an electrometric amplifier with an absolute resolution of 10^{-13} A.

Nb/Si and Nb/Al/Nb/Si targets were made on a KÉF-4.5 (phosphorus-doped, resistivity 4.5 Ω cm) Si(100) substrate using an L560 installation (Leybold). The substrate was rinsed first in bichromate at 90°C for 1 h and then in running deionized water. In the installation, the substrates were decontaminated in an rf argon glow at a pressure of 5×10^2 Pa for 150 s. Then, the chamber was evacuated to 2.4×10^{-2} Pa with the subsequent evaporation of Nb and Al.

The thicknesses of the Nb and Al layers were measured with a Taylor–Hobsen profilometer using the two-beam interferometry method. Multiple measurements showed an error within 10%. The rates of Nb and Al evaporation were found to be 2.34 and 0.16 nm/s, respectively.

Figures 2–4 show the experimental results. The spectra were obtained after subtracting the background signal and with regard for the point-spread function of the analyzer. The spectra of electrons reflected from homogeneous targets are depicted in Fig. 2. Figure 3 shows the energy spectra of electrons reflected from the Nb/Si target; Fig. 4, reflected electron spectra from two niobium films with an aluminum layer in between (Nb/Al/Nb/Si target) that is located at different depths. The depth of the aluminum layer was specified by varying the thickness of the upper film. The effect of the silicon substrate on the spectra for the given electron probe energies is virtually absent.

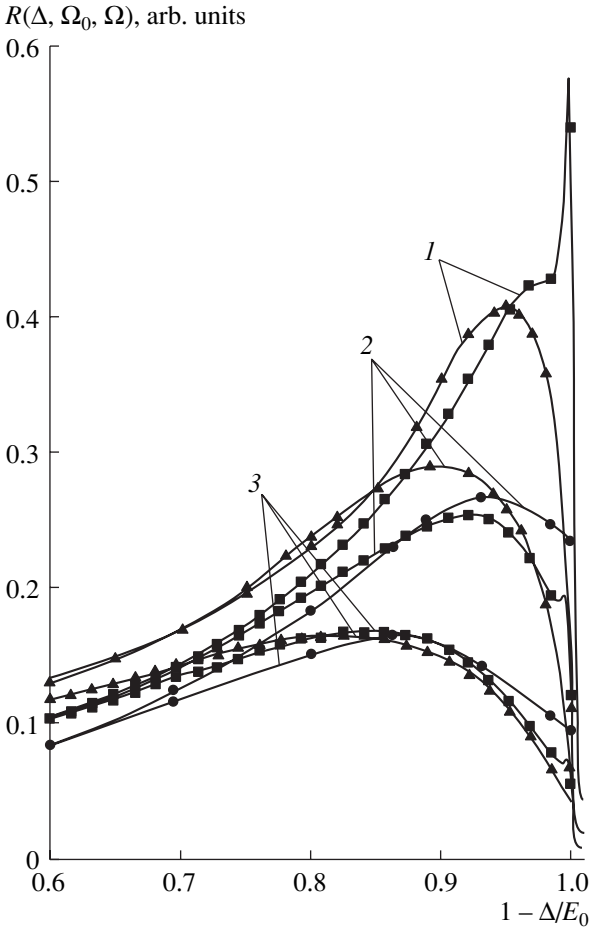


Fig. 1. Experimental data of different authors for the spectra of electrons reflected from a semi-infinite target: initial energy 30 keV, right angle of incidence, angle of observation 45° to the normal (the same angles are in Figs. 2–4). Squares, data from [8]; circles, [4]; triangles, [2]. The target material is (*1*) Pt, (*2*) Ag, and (*3*) Cu.

Below, we elaborate a model to treat the spectra and to restore the elemental profile across the target depth.

THEORETICAL ANALYSIS OF ENERGY SPECTRA OF ELECTRONS REFLECTED FROM HOMOGENEOUS AND MULTILAYER TARGETS

Consider a homogeneous solid layer of thickness x. Figure 5 represents a scheme of electron scattering. Let us increase the thickness of the layer by adding a thin film of thickness dx to its lower boundary. The reflection function changes in this case by the value

$$dR(x, \Delta, \Omega_0, \Omega) = R(x + dx, \Delta, \Omega_0, \Omega) - R(x, \Delta, \Omega_0, \Omega)$$

according to the process depicted by thick lines. Math-

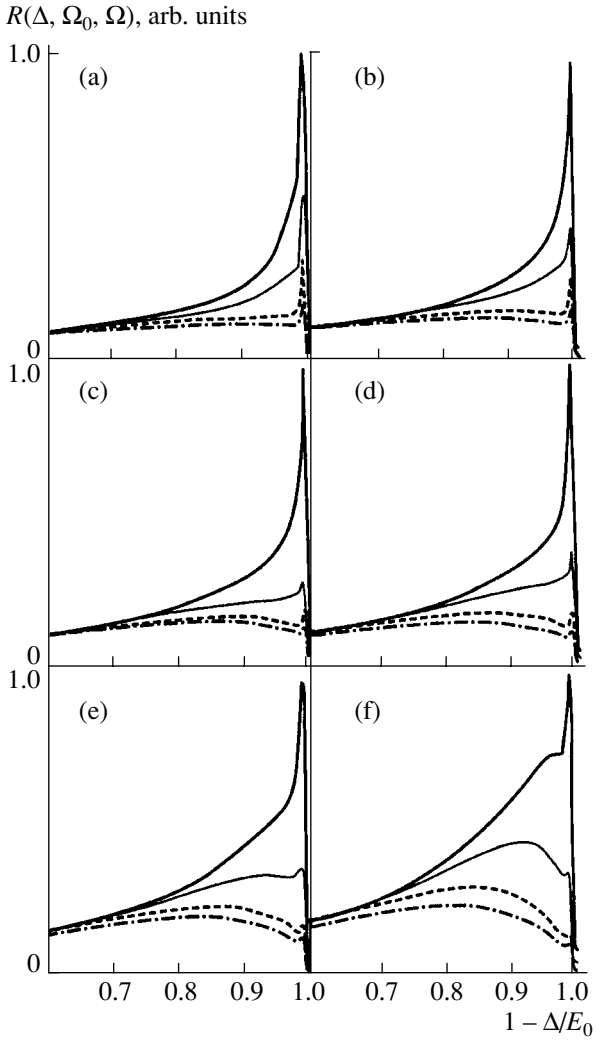


Fig. 2. Spectra of electrons reflected from a homogeneous semi-infinite target. The energy of normal incidence is (a) 4, (b) 6, (c) 8, (d) 10, (e) 16, and (f) 32 keV. Shown are spectra of electrons reflected from the Au target (continuous thick line), Ag target (continuous thin line), Cu target (dashed line), and V target (dash-and-dot line).

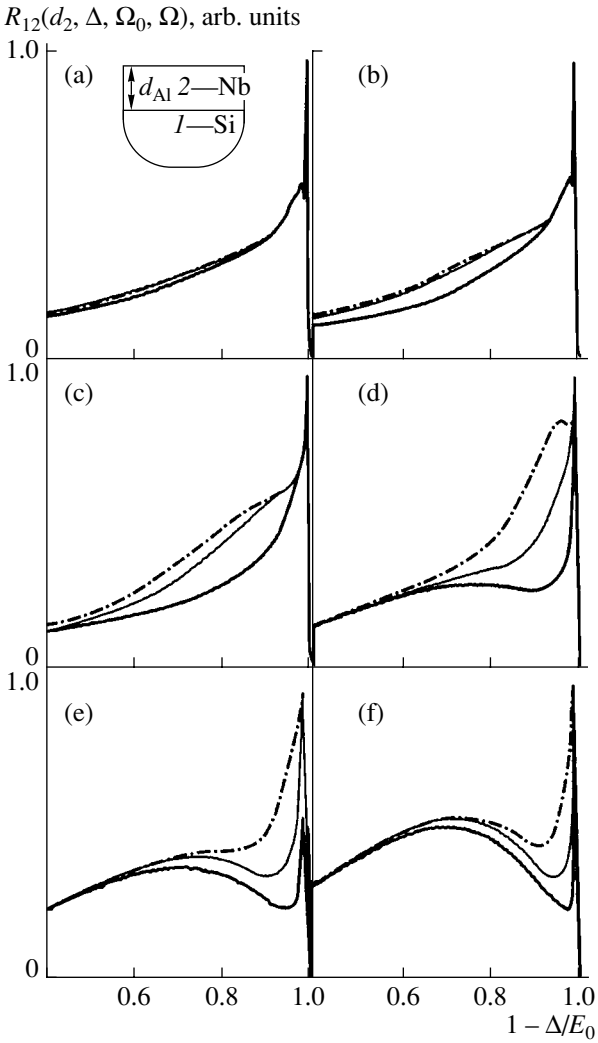


Fig. 3. Spectra of electrons reflected from the Nb/Si two-layer target. The initial energy is (a) 4, (b) 6, (c) 8, (d) 10, (e) 24, and (f) 32 keV. Shown are the spectra of electrons reflected from the targets with the Nb layers evaporated for different time intervals: 20 s (continuous thick line), 40 s (continuous thin line), and 60 s (dash-and-dot line).

ematically, the process is described as

$$n dx \iiint T(x, \Delta', \Omega_0, \Omega')\omega_{el}(\Omega', \Omega'') \times T(x, \Delta - \Delta', \Omega'', \Omega) d\Delta' d\Omega' d\Omega''. \quad (1)$$

Here, $T(x, \Delta, \Omega_0, \Omega)$ is the transmission function; $\Delta = E_0 - E$ is the energy loss; E_0, E, $\Omega_0 = \{\vartheta_0, \varphi_0\}$, and $\Omega = \{\vartheta, \varphi\}$ are the electron energies and motion directions at the entrance to and the exit from the layer, respectively; $d\Omega = \sin\vartheta d\vartheta d\varphi$; $0 < \vartheta < \pi/2$; $0 < \varphi < 2\pi$; $0 < \Delta' < E_0$; n is the atomic concentration in the target; and $\omega_{el}(\Omega', \Omega'')$ is the differential elastic scattering cross section. Equating $dR(x, \Delta, \Omega_0, \Omega)$ to expression (1), we arrive at the differential equation for $R(x, \Delta, \Omega_0, \Omega)$

$$\frac{dR(x, \Delta, \Omega_0, \Omega)}{dx} = n \iiint T(x, \Delta', \Omega_0, \Omega')\omega_{el}(\Omega', \Omega'') \times T(x, \Delta - \Delta', \Omega'', \Omega) d\Delta' d\Omega' d\Omega'' \quad (2)$$

with the boundary condition $R(x, \Delta, \Omega_0, \Omega) = 0$ at $x = 0$. Equation (2) in a set with a similar equation for $T(x, \Delta, \Omega_0, \Omega)$, which in its turn depends on $R(x, \Delta, \Omega_0, \Omega)$, is exact.

In general form, this set is impossible to solve; therefore, our goal is to select an approximation that will analytically describe electron backscattering most adequately. The additivity of the elastic and inelastic collision integrals in the transport equation allows us to

represent the transmission function in multiplicative form:

$$T(x, \Delta, \Omega_0, \Omega) = T_{in}(x, \Delta)T_{el}(x, \Omega_0, \Omega), \quad (3)$$

where $T_{in}(x, \Delta)$ is the inelastic transmission function and $T_{el}(x, \Omega_0, \Omega)$ is the elastic transmission function.

Let us restrict ourselves to the one-velocity approximation. Solving linear differential equation (2) and substituting the transmission function in the form of (3) into the solution, we come to an expression that can be written as

$$R(d, \Delta, \Omega_0, \Omega) = \int_0^d A(x(1/\mu_0 + 1/\mu), \Omega_0, \Omega) \times T_{in}(x(1/\mu_0 + 1/\mu), \Delta)dx. \quad (4)$$

Here, d is the thickness of the layer, $\mu_0 = \cos\vartheta_0$, $\mu = \cos\vartheta$,

$$A(u, \Omega_0, \Omega) = n\iint T_{el}(x/\mu_0, \Omega_0, \Omega')\omega_{el}(\Omega', \Omega'') \times T_{el}(x/\mu, \Omega'', \Omega)d\Omega' d\Omega'' \quad (5)$$

is the free path distribution of electrons reflected, and $u = x(1/\mu_0 + 1/\mu)$ is the electron free path in the low-angle approximation. The function $A(u, \Omega_0, \Omega)$ is the solution to the elastic part of the problem. It is responsible for reflected electrons that traveled the path u in the target. The representation of the solutions in the form of (4) and (5) became possible because of the low-angle approximation. This approximation deals with the projected range of a particle, rather than with its actual travel, and hence fails to adequately describe particles whose free path exceeds the transport length.

Next, consider the approximation of single deflection in the elastic channel and the approximation of continuous slowing-down in the inelastic channel. Such an approach is only applicable to the description of the energy spectra when the electron energy exceeds several megaelectronvolts. However, it illustrates most vividly the essence of our electron spectroscopy method for depth profiling and so merits consideration.

In the single-deflection approximation, a change in the particle direction is ignored up to an event of "strong" collision, when the descending motion of a particle in the target becomes ascending:

$$T_{el}(x, \Omega', \Omega'') = \delta(\Omega' - \Omega''), \quad (6)$$

where $\delta(x)$ is the Dirac function.

Under these conditions, the range distribution takes the form

$$A(u, \Omega_0, \Omega) = m\omega_{el}(\Omega_0, \Omega). \quad (7)$$

In the single-deflection approximation, the reflection function, according to (4)–(7), can be written as

$$R(d, \Delta, \Omega_0, \Omega) = n\omega_{el}(\Omega_0, \Omega)\int_0^{d}\int_0^{\Delta} T_{in}\left(\frac{x}{\mu_0}, \Delta'\right) \times T_{in}\left(\frac{x}{\mu}, \Delta - \Delta'\right)d\Delta' dx. \quad (8)$$

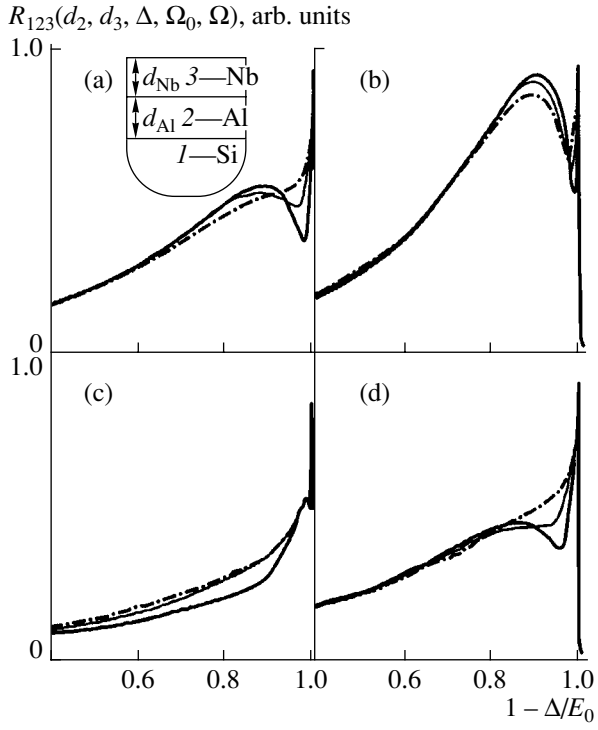


Fig. 4. Spectra of electrons reflected from the Nb/Al/Nb/Si multilayer target. The initial energy is (a) 4, (b) 8, (c) 10, and (d) 16 keV. The upper Nb layer was evaporated for 10 s (continuous thick line), 20 s (continuous thin line), and 30 s (dash-and-dot line). The intermediate Al layer was evaporated for 210 s.

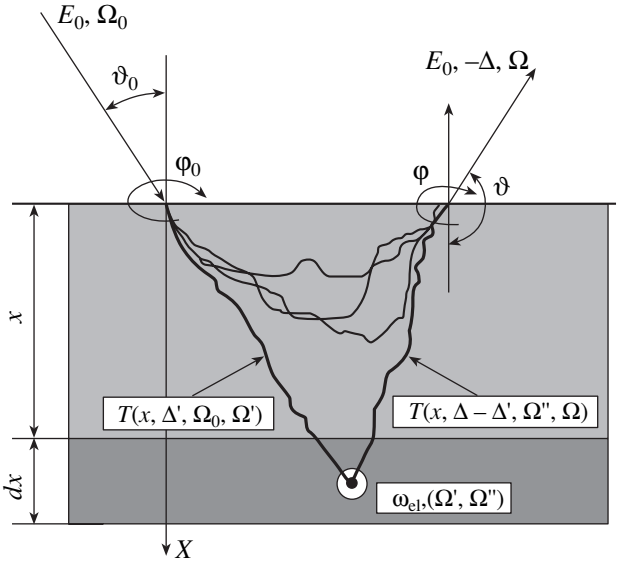


Fig. 5. Model illustrating reflection from the layer.

Integrating (8) with respect to Δ' and then with respect to x and representing the function $T_{in}(x, \Delta)$ within the continuous slowing-down approximation as $T_{in}(x, \Delta) = \delta(\Delta - \bar{\varepsilon} x)$, we arrive at

$$R(d, \Delta, \Omega_0, \Omega) = n\omega_{el}(\Omega_0, \Omega) \times \int_0^d dx \delta\left(\Delta - \bar{\varepsilon} x\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]\right) = \frac{n\omega_{el}(\Omega_0, \Omega)}{\bar{\varepsilon}\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]} \times \left\{\Theta\left(\Delta - \bar{\varepsilon} d\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]\right) - \Theta(\Delta)\right\}, \quad (9)$$

where $\Theta(x)$ is the Heaviside function and $\bar{\varepsilon}$ is the mean energy loss per unit length.

The self-consistent calculation of the spectrum of electrons reflected from a thin solid layer ($d \ll l_{tr}$) will be based on the Gaudsmith–Saundersen equation, which describes the electron motion in the elastic channel. In a form convenient for numerical calculations, it is given by

$$T_{el}^{G-S}(x, \Omega_0, \Omega) = \delta(\gamma)\exp(-n\sigma_{el}x) + \sum_{l=1}^{\infty}\frac{2l+1}{2}P_l(\gamma) \times (\exp(n(\omega_l - \sigma_{el})x) - \exp(-n\sigma_{el}x)), \quad (10)$$

where ω_l is the coefficient of the expansion of $\omega_{el}(\Omega_0, \Omega)$ in Legendre polynomials $P_l(\gamma)$, σ_{el} is the total elastic cross section, and γ is the angle between Ω_0 and Ω.

The angle γ is related to the angles $\Omega_0 = \{\vartheta_0, \varphi_0\}$ and $\Omega = \{\vartheta, \varphi\}$ by the well-known expression

$$\cos\gamma = \cos\vartheta_0\cos\vartheta + \sin\vartheta_0\sin\vartheta\cos(\varphi - \varphi_0).$$

Below, along with (6), we will also use the formula for T_{el} obtained in the diffusion approximation,

$$T_{el}^{D}(x, \Omega_0, \Omega) = \sum_{l=1}^{\infty}\frac{2l+1}{2}P_l(\gamma)\exp\left(-n\sigma_{tr}\frac{l(l+1)}{2}x\right) \quad (11)$$

(σ_{tr} is the transport cross section). With this formula, we will derive expressions that have a clear physical meaning and simplify greatly the analysis of the reflected electron spectra without loss in the calculated data accuracy (as follows from the close coincidence of the first ten moments of distributions for (10) and (11)).

Let us represent the elastic cross section that transforms the descending electron flow into the ascending one in the form

$$\omega'_{el}(\Omega_0, \Omega) = \omega_{el}(\Omega_0, \Omega) - \sigma_{el}\delta(\Omega_0 - \Omega).$$

According to (10), the function $A(u, \Omega_0, \Omega)$ is given by

$$A^{G-S}(u, \Omega_0, \Omega) = n\exp(-n\sigma_{el}u)\omega_{el}(\Omega_0, \Omega) + n\sum_{l=1}^{\infty}\frac{2l+1}{2}(\omega_l - \sigma_{el})P_l(\gamma) \times (\exp(-n(\omega_l - \sigma_{el})u) - \exp(-n\sigma_{el}u)). \quad (12)$$

In the diffusion approximation, the range distribution in the target $A(u, \Omega_0, \Omega)$ is given by

$$A^{D}(u, \Omega_0, \Omega) = n\sum_{l=1}^{\infty}\frac{2l+1}{2}(\omega_l - \sigma_{el})P_l(\gamma) \times \exp\left(-\frac{u}{l_{tr}}\frac{l(l+1)}{2}\right), \quad (13)$$

where the transport length l_{tr}, the basic parameter in the problem of electron elastic scattering, appears in explicit form.

The detailed description of techniques for calculating the energy spectra $T_{in}(x, \Delta)$ of electrons passing through a solid layer of thickness x is given in [9]. As for energy loss fluctuations, they should be considered in the context of two problems of calculating the energy spectra of reflected electrons.

(i) Finding the electron spectrum in a narrow energy domain $E_0 - \Delta_1 \le E \le E_0$ ($\Delta_1 \approx 100$ eV) based on detailed knowledge of the single inelastic scattering cross section $\omega_{in}(\Delta)$ [10]. A corresponding experimental situation takes place when the characteristic electron energy loss (CEEL) spectra are measured.

(ii) Finding the reflected electron spectrum in a wide energy domain $0 < E < E_0 - \Delta_1$, where a detailed description of the single inelastic scattering cross section is inappropriate.

With exhaustive accuracy, the spectrum is described using the single inelastic scattering cross section $\omega_{in}(\Delta)$ determined in the two-threshold approximation [9, 10]:

$$\omega_{in}(\Delta) = \sum_k \lambda_k \omega_k^i(\Delta), \quad (14)$$

where

$$\omega_k^i(\Delta) = \begin{cases} 0, & \Delta < J_p \\ \sigma_{in}\dfrac{(3+\alpha_p)J_p^{3+\alpha_p}}{\Delta^{4+\alpha_p}}, & J_p \le \Delta \\ \sigma_{in}\dfrac{(1+\alpha_i)J_i^{1+\alpha_i}}{\Delta^{2+\alpha_i}}, & J_i \le \Delta, \end{cases} \quad (15)$$

σ_{in} is the total inelastic cross section, λ_k is the probability of inelastic scattering, J_p is an adjustable parameter

($J_p \approx \varepsilon_{pl}$, where ε_{pl} is the energy of bulk plasmon excitation), J_i are the ionization thresholds, and α_p and α_i are adjustable factors.

Then, the function $T_{in}(x, \Delta)$ is given by [9]

$$T_{in}(\tau, \Delta) = \frac{\exp(-\tau)}{\pi}\int_0^\infty \exp(\tau[\lambda_p cf(pJ_p, -3-\alpha_p) + \lambda_i cf(pJ_i, -1-\alpha_i)])\cos(p\Delta - \tau[\lambda_p sf(pJ_p, -3-\alpha_p) + \lambda_i sf(pJ_i, -1-\alpha_i)])dp. \quad (16)$$

Here, $\tau = xn\sigma_{in}$, $\lambda_i = \sigma_{ion}/\sigma_{in}$, σ_{ion} is the ionization cross section, $\lambda_p = 1 - \lambda_i$,

$$cf(s, a) = -as^{-a}C(s, a),$$
$$sf(s, a) = -as^{-a}S(s, a). \quad (17)$$

In turn, the functions $C(s, a)$ and $S(s, a)$ are related to the incomplete gamma-function as $\Gamma(a, is) = \exp(ia\pi/2)[C(s, a) - iS(s, a)]$.

Calculations based on formulas (16) and (17) are hard to perform. However, in our problem, we are interested in the domain $\tau > 1$; therefore, it is possible to consider the first three terms in the expansion of (17) in powers of s:

$$cf(s, a) = 1 - as^{-a}\Gamma(a)\cos(a\pi/2) - as^2/(4+2a),$$
$$sf(s, a) = -as^{-a}\Gamma(a)\sin(a\pi/2) + as/(1+a) - as^3/(18+3a). \quad (18)$$

In view of (4) and (13), we come to an expression for the function of electron reflection from a solid layer:

$$R(d, \Delta, \Omega_0, \Omega) = n\frac{\mu\mu_0}{\mu+\mu_0}\int_0^{d\left(\frac{1}{\mu}+\frac{1}{\mu_0}\right)}\sum_{l=1}^{\infty}\frac{2l+1}{2}(\omega_l - \sigma_{el}) \times P_l(\gamma)\exp\left(-\frac{u}{l_{tr}}\frac{l(l+1)}{2}\right)T_{in}(u, \Delta)du. \quad (19)$$

Inelastic losses will be described in terms of formulas (16) and (18).

Figure 6 shows the spectra of electrons reflected from the layer. Theoretical calculations performed with formula (19) are seen to be in good agreement with Kulenkampff and Ruttiger's experimental data [11]. The total inelastic cross section was determined from data in [12]. The elastic differential cross section $\omega_{el}(\gamma)$ was found from tables [13]. By way of example, the dashed line in Fig. 6 demonstrates the result of calculation by formula (9) for the thinnest layer (marked by a circle in Fig. 5). Without pretending to a high accuracy, formula (9), derived by simplest calculations, still gives an estimate of the layer thickness.

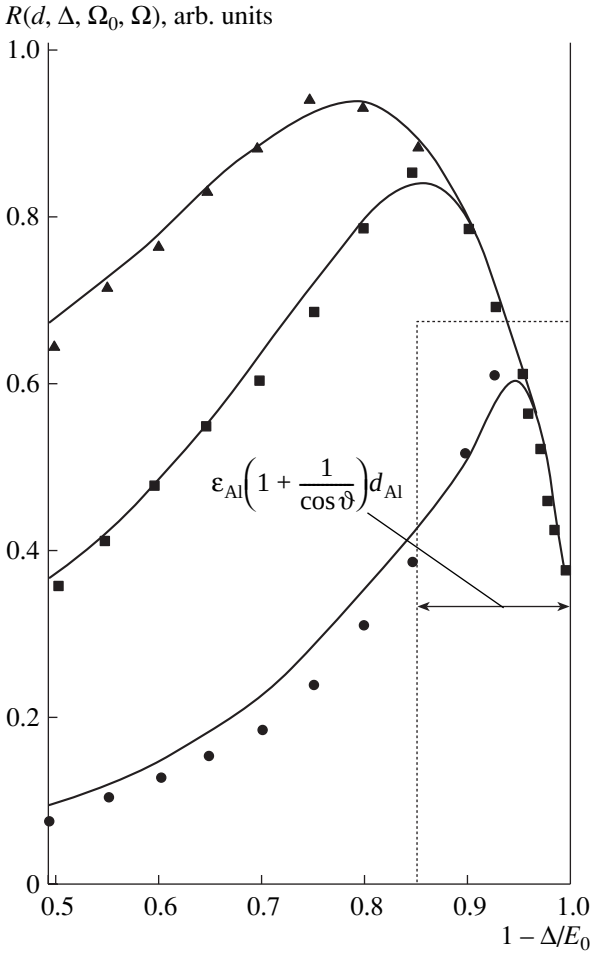


Fig. 6. Spectra of electrons reflected from the Al layer: right angle of incidence, angle of observation 63° to the normal, initial energy 30 keV. Data points are taken from [11]. The target thickness is 0.44 mg cm^{-2} (triangles), 0.31 mg cm^{-2} (squares), and 0.16 mg cm^{-2} (circles). Continuous lines depict the related spectra calculated by (19).

Now, let us calculate the spectra of electrons reflected from multilayer targets. Let a wide beam of particles be incident on a layered plane-parallel target at an angle Ω_0 to the normal. The target consists of a layer of material *2* (thickness d_2) applied on a layer of material *1* (thickness d_1). Let $d_1 \longrightarrow \infty$ (semi-infinite target). Also, let us know the following parameters: R_1, double-differential function of reflection from semi-infinite target *1* [$R_1(d_1 \longrightarrow \infty) \equiv R_1$]; and $R_2(d_2)$ and T_2, the double-differential functions of reflection from, and transmission through, the free layer of material *2*. Then, the reflection function $R_{12}(d_2)$ for the layered target is given by

$$R_{12}(d_2) = R_2(d_2) + T_2 \otimes R_1 \otimes T_2 + T_2 \otimes R_1 \otimes R_2(d_2) \otimes R_1 \otimes T_2 + \ldots. \quad (20)$$

Here, we use an abbreviated form of writing where the symbol $\otimes$ means integration over angular variables and convolution over energy losses.

Imagine that the homogeneous target made of material *2* is two-layer. Applying the same approach as for the inhomogeneous target, we arrive at

$$\begin{aligned} R_{22}(d_2) &= R_2 = R_2(d_2) + T_2 \otimes R_2 \otimes T_2 \\ &+ T_2 \otimes R_2 \otimes R_2(d_2) \otimes R_2 \otimes T_2 + \ldots. \end{aligned} \tag{21}$$

Subtracting Eq. (21) from relationship (20) yields

$$\begin{aligned} R_{12}(d_2) - R_2 &= T_2 \otimes (R_1 - R_2) \otimes T_2 \\ &+ T_2 \otimes (R_1 \otimes R_2(d_2) \otimes R_1 \\ &- R_2 \otimes R_2(d_2) \otimes R_2) \otimes T_2 + \ldots. \end{aligned} \tag{22}$$

In the first approximation, we obtain the reflection function for the inhomogeneous (layered) target:

$$R_{12}(d_2) = R_2 + T_2 \otimes (R_1 - R_2) \otimes T_2. \tag{23}$$

Expression (23) meets the invariance conditions throughout the range of the thickness d of the inhomogeneous coating. We have

$$\begin{aligned} T_2 &\longrightarrow 1 \quad \text{and} \quad R_{12}(d_2) = R_1, \quad \text{if } d_2 \longrightarrow 0, \\ T_2 &\longrightarrow 0 \quad \text{and} \quad R_{12}(d_2) = R_2, \quad \text{if } d_2 \longrightarrow \infty. \end{aligned} \tag{24}$$

Let us write formula (23) in detail:

$$\begin{aligned} R_{12}(d_2, \Delta, \Omega_0, \Omega) &= R_2(\Delta, \Omega_0, \Omega) \\ + \int\int\int\int T_2(d_2, \Delta - \Delta', \Omega_0, \Omega')&[R_1(\Delta' - \Delta'', \Omega', \Omega'') \\ - R_2(\Delta' - \Delta'', \Omega', \Omega'')]&T_2(d_2, \Delta'', \Omega'', \Omega) \\ \times d\Delta' d\Delta'' d\Omega' d\Omega''&. \end{aligned} \tag{25}$$

Here, $T_2(d_2, \Delta, \Omega_0, \Omega)$ is the function of particle transmission that is responsible for the energy and angular distributions of the particles moving in layer *2* of thickness d_2, and $R_{1(2)}(\Delta, \Omega_0, \Omega)$ is the function of particle reflection from the semi-infinite target made of material *1* or *2*.

The model elaborated allows for simple generalization for the multilayer case. For example, for a three-layer target, we have

$$R_{123}(d_2, d_3) = R_3 + T_3 \otimes (R_{12}(d_2) - R_3) \otimes T_3, \tag{26}$$

where $R_{12}(d_2)$ is given by (23) and the layers are counted from the bottom.

Let us find the reflection function for a three-layer target under conditions of continuous slowing-down and single deflection. To do this, we substitute (6), (7), and the function $T_{in}(x, \Delta)$ derived in the continuous slowing-down approximation into (26) and take the integral to get

$$\begin{aligned} R_{123}(d_1, d_2, d_3, \Delta, \Omega_0, \Omega) &= \frac{n_1 \omega_{el3}(\Omega_0, \Omega)}{\bar{\varepsilon}_3\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]} \\ &\times \left[\Theta\left(\Delta - \bar{\varepsilon}_3 d_3\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]\right) - \Theta(\Delta)\right] \\ + \frac{n_2 \omega_{el2}(\Omega_0, \Omega)}{\bar{\varepsilon}_2\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]}&\left[\Theta\left(\Delta - (\bar{\varepsilon}_2 d_2 + \bar{\varepsilon}_3 d_3)\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]\right)\right. \\ \left. - \Theta\left(\Delta - \bar{\varepsilon}_3 d_3\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]\right)\right] &+ \frac{n_1 \omega_{el1}(\Omega_0, \Omega)}{\bar{\varepsilon}_1\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]} \\ \times \left[\Theta\left(\Delta - (\bar{\varepsilon}_1 d_1 + \bar{\varepsilon}_2 d_2 + \bar{\varepsilon}_3 d_3)\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]\right)\right.& \\ \left. - \Theta\left(\Delta - (\bar{\varepsilon}_2 d_2 + \bar{\varepsilon}_3 d_3)\left[\frac{1}{\mu_0} + \frac{1}{\mu}\right]\right)\right]&. \end{aligned} \tag{27}$$

In the first approximation, formula (27) describes the specific features of the spectra of electrons reflected from a three-layer target. A spectrum defined by (27) is represented by the continuous curve (Nb/Al/Nb/Si target) in Fig. 7. Simple mathematics demonstrates how the spectrum changes when a layer of light material appears at a certain depth. In this case, similarly to the above case of electron reflection from the aluminum free surface, estimates made under conditions of continuous slowing-down and single deflection give an idea of the target geometry.

In this work, we are also interested in the accuracy of locating a light material (Al in our case) marker buried in the bulk of a heavy material (Nb). Figure 8 compares experimental (data points) and calculated (formula (26)) spectra of electrons reflected from a homogeneous semi-infinite niobium target. The function $A_k(u, \Omega_0, \Omega)$ for either of the materials was calculated by (12); the function $T_{ink}(x, \Delta)$, from (16) and (18). The discrepancy between the experimental and analytical spectra near the peak of the elastically reflected electrons is associated with the simplified form of the function $T_{ink}(x, \Delta)$; namely, formulas (16) and (18) do not involve the singular term responsible for the electrons passing through the layer without scattering. The range of the peak of the elastically reflected electrons, which is carefully studied in CEEL spectroscopy, is of no interest in this specific case. The thickness of the aluminum marker and its depth (i.e., the thickness of the niobium layer) in different targets that were restored by the trial-and-error (fitting) procedure were found to be as follows: $d_{Al} = 38$ nm, $d_{Nb1} = 12$ nm, $d_{Nb2} = 26$ nm, and $d_{Nb3} = 45$ nm. The thicknesses of the layers determined with the Taylor–Hobsen profilometer were $d_{Al} = 34$ nm,

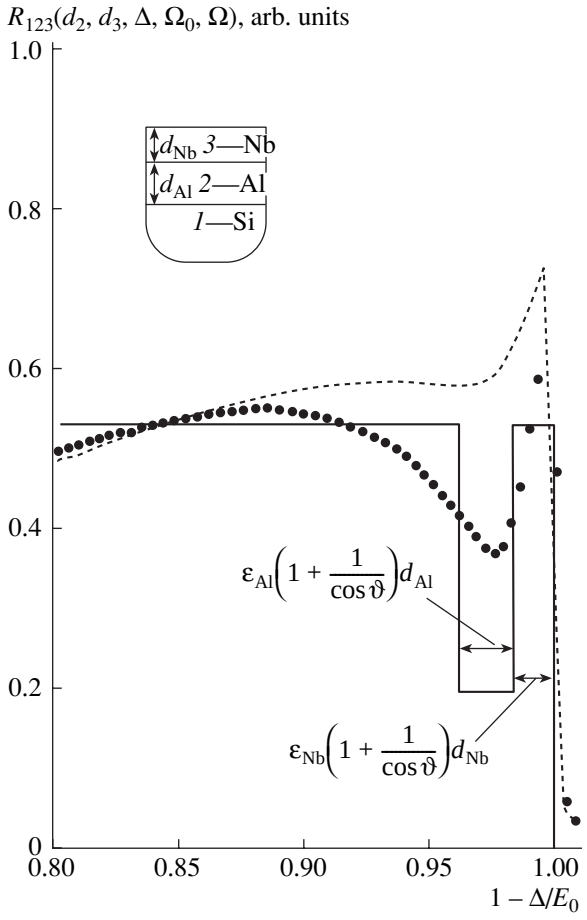


Fig. 7. Experimental spectrum of electrons reflected from the Nb/Al/Nb/Si multilayer target (circles) vs. spectrum calculated by (27) within the model of single deflection and continuous slowing-down (continuous line). The dashed line shows the Nb spectrum.

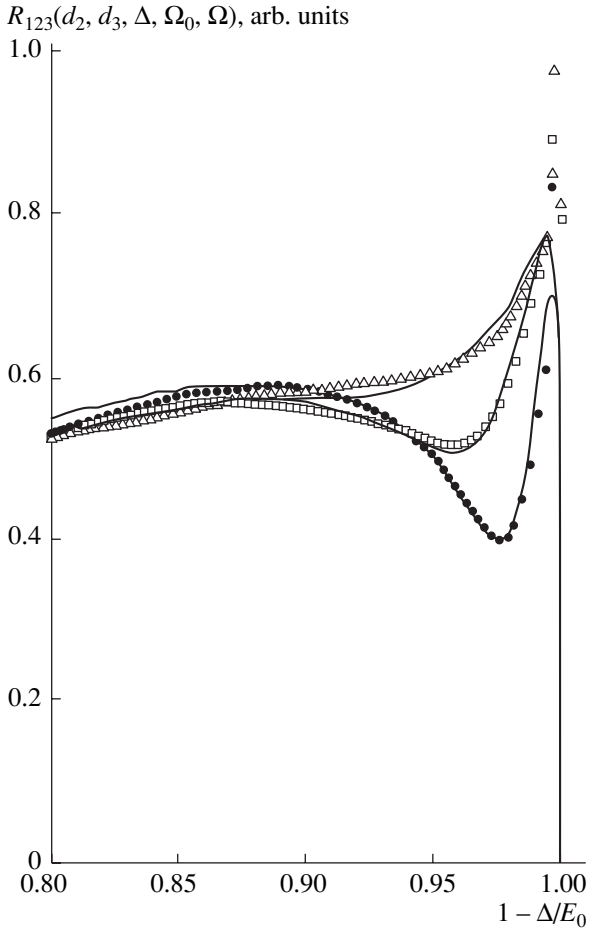


Fig. 8. Experimental spectra of electrons reflected from the Nb/Al/Nb/Si multilayer target (circles, d_{Nb1}; squares, d_{Nb2}; triangles, d_{Nb3}) vs. spectra calculated by (26) in terms of the self-consistent theory (continuous lines).

$d_{Nb1} = 23$ nm, $d_{Nb2} = 47$ nm, and $d_{Nb3} = 70$ nm. The most considerable discrepancy is observed for the Nb thickness. This fact, in our opinion, reflects today's lack of understanding of electron slowing-down parameters in solids.

The extraction of exact quantitative data for a sample under study (specifically, for the element distribution across its depth) with reflected electron spectroscopy is unlikely without knowing reliable information about the single inelastic electron energy loss cross section $\omega_{in}(\Delta)$ in a solid. At present, published data even for the first moments of $\omega_{in}(\Delta)$, total cross section σ_{in} and mean energy loss per unit length $\bar{\varepsilon}$, may differ severalfold. The most exhaustive information on the inelastic cross section $\omega_{in}(\Delta)$ can be found by analyzing the reflected electron spectra in the low energy loss range with a resolution of better than 1 eV. Such an approach is applied in CEEL spectroscopy. Related spectra have clear-cut peaks and are formed by electrons that have traveled a distance comparable to the mean free path for inelastic scattering l_{in}. Since $l_{in} \ll l_{tr}$, the electrons responsible for the CEEL spectrum also experience low-multiplicity scattering in the elastic channel. Formulas (12) and (13) describe the electron range distribution in the target at $u \ll l_{tr}$ with an extremely high accuracy that is limited only by the accuracy with which the elastic scattering cross section is determined. Because of this, CEEL spectra seem to be a reliable tool for studying an elementary event of inelastic scattering. A method for restoring the cross section $\omega_{in}(\Delta)$ from characteristic electron energy losses is described in [10].

CONCLUSION

Our experiments on the electron probing of targets show that the reflected electron spectra are highly informative. The spectrum analysis technique elaborated in this work forms a basis for a method of layer-by-layer analysis that can be used in commercial instruments for surface analysis, for example, in Auger spectrometers.

The self-consistent consideration of energy loss fluctuations makes it possible to properly treat experimental data both in the range of low losses and in that range of losses responsible for the spectrum as a whole (in other words, for the dome-shaped part of the spectrum).

In this work, we demonstrated the possibility of finding the depth and thickness of the aluminum marker in the niobium target. The adequate study of such a target in the conventional statement of the RBS method (most widely used for layer-by-layer analysis) is impossible (only the niobium spectrum will be observed). Electron spectroscopy, unlike the RBS method, can study light inclusions in a heavy matrix and heavy layers in a light material with equal efficiency [14].

Errors in determining the layer thickness arise from a number of factors.

(i) The basic (yet avoidable) error is associated with the large spread in data for inelastic scattering cross sections $\omega_{in}(\Delta)$.

(ii) The method presented compares spectra taken from a pure target and from this target with inclusions. The difference between these spectra is the most significant if the atomic numbers of a target material, Z_1, and of an impurity, Z_2, greatly diverge. As $\Delta Z = |Z_1 - Z_2|$ decreases, the error in thickness measurement grows. For the targets used in this work, the error was no more than 1 nm.

The relationship between the accuracy of depth profiling and the atomic number Z depends on the layer depth, initial energy, and energy resolution of the energy analyzer. This subject is beyond the scope of this work.

An important advantage of electron spectroscopy is the ease of varying the electron probe energy. The target thickness that can be analyzed is comparable to the electron transport length l_{tr}. Since $l_{tr} \sim E_0^2$, the change in the beam energy from $E_0 = 10$ keV to $E_0 = 32$ keV increases the target thickness from which information is derived tenfold.

If the size of a multilayer structure does not exceed l_{in}, one can study it with a submonolayer accuracy using CEEL spectra. The associated technique will be described in subsequent articles.

REFERENCES

1. M. V. Gomoyunova, Usp. Fiz. Nauk **136** (1), 105 (1982) [Sov. Phys. Usp. **25**, 58 (1982)].
2. H. Kulenkampff and K. Ruttiger, Z. Phys. B **137**, 426 (1954).
3. H. Kanter, Ann. Phys. (Leipzig) **20**, 144 (1957).
4. H. E. Bishop, in *Proceedings of the 4th International Conference on X-ray Optics and X-ray Microanalysis, Paris, 1965*, p. 153.
5. T. Matsukawa, R. Shimizu, and H. Hashimoto, J. Phys. D **7**, 695 (1974).
6. E. H. Darlington and V. E. Cosslet, J. Phys. D **5**, 1969 (1972).
7. P. Gerard, J. L. Balladore, H. Pinna, *et al.*, J. Phys. III **2**, 1015 (1992).
8. V. P. Afanas'ev, S. D. Fedorovich, M. S. Esimov, *et al.*, Zh. Tekh. Fiz. **64** (8), 180 (1994) [Tech. Phys. **39**, 845 (1994)].
9. V. P. Afanas'ev, A. V. Lubenchenko, and A. A. Ryzhov, Poverkhnost, No. 1, 6 (1996).
10. V. P. Afanas'ev, A. V. Lubenchenko, and A. V. Strizhov, Poverkhnost, No. 8, 16 (1999).
11. H. Kulenkampff and K. Ruttiger, Z. Phys. **152**, 249 (1958).
12. A. F. Akkerman, *Simulation of Charged Particle Trajectory in Matter* (Énergoatomizdat, Moscow, 1991).
13. M. E. Riley, C. J. MacCallum, and F. Biggs, At. Data Nucl. Data Tables **15**, 443 (1975).
14. V. P. Afanas'ev and D. Naujoks, Z. Phys. B **84**, 397 (1991).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1453–1457. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 109–113.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Coefficient of Ion Mobility Versus Electric Field Strength Dependence in Gases: Experimental Determination

I. A. Buryakov

Design and Technological Institute of Instrument Engineering for Geophysics and Ecology, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia

e-mail: majak@uiggm.nsc.ru

Received February 12, 2002

Abstract—A technique for finding the electric field dependence of the coefficient of ion mobility in gases is suggested. For ions of 2,4,6-trinitrotoluene, 2,4-dinitrotoluene, 1,3,5-trinitrobenzene, 1,3-dinitrobenzene, and dimethylmethylphosphonate, these dependences are taken in air by applying a variable periodic polarity-asymmetric specially shaped electric field. The accuracy of the technique suggested is estimated and compared with that of the conventional drift tube method. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The electric field (E) dependence of the coefficient of ion mobility (K) in gases in view of the gas molecule concentration (N) is of both scientific and applied interest. First, experimentally found dependences $K(E/N)$ may provide information on the molecule–ion interaction potential; help in estimating the ion mean energy as a function of the parameter E/N; and, with further advances in the ion motion theory, determine the cross sections of elastic, inelastic, and reactive collisions between particles [1, 2]. Second, data on the dependence $K(E/N)$ are necessary for analytical applications in using ion mobility increment spectrometers [3].

In experiments, the dependence $K(E/N)$ is usually found with drift tubes, which measure the ion drift velocity $V = KE$ with the coefficient K determined with an accuracy of 1–5% [4, 5]. Such an accuracy is insufficient for studying ions whose mobility varies with E/N within several percent. These are, as a rule, heavy ions formed by organic molecules or ionic complexes [6].

The recent method for finding the dependence $K(E/N)$ [7, 8] uses a variable periodic polarity-asymmetric electric field, under the action of which the drift velocity being measured varies in proportion to the increment of ion mobility coefficient $\alpha(E/N)$. On the assumption that the field is moderately high and the gas particle spatial distribution is isotropic, the dependence $K(E/N)$ can be represented in the form of a series involving only even powers of the parameter E/N [1]:

$$K(E/N) = K_0(1 + \alpha(E/N)) = K_0\left(1 + \sum_{n=1}^{\infty} \alpha_{2n}\left(\frac{E}{N}\right)^{2n}\right), \quad (1)$$

where K_0 in $cm^2/(V\ s)$ is the coefficient of mobility in a "weak" field, where the ion energy is close to the thermal energy of a gas ($E/N < 6$ Td, where 1 Td = 10^{-17} V cm^2), and α_{2n} are the coefficients preceding the powers of E/N in the expansion.

In [7, 8], the coefficients α_{2n} were determined for positive ions of tertiary alkylamines; in [9], for Cl^- ions. In all the works cited, the objects of investigation were ions for which the dependence $\alpha(E/N)$ is strong and which accordingly can be studied with conventional drift tubes.

The goal of this work is to study ions for which the dependence $\alpha(E/N)$ is weak. Of particular interest is the dependence $\alpha(E/N)$ in air for negative ions of organic substances: 2,4,6-trinitrotoluene (TNT), 2,4-dinitrotoluene (DNT), 1,3,5-trinitrobenzene (TNB), 1,3-dinitrobenzene (DNB), as well as for positive ions of dimethylmethylphosphonate (DMMP). Another goal was to find the coefficients α_{2n} for these ions and to estimate the accuracy in determining the dependences $\alpha(E/N)$.

METHOD FOR DETERMINING THE $\alpha(E/N)$ DEPENDENCE

Under the action of a variable periodic polarity-asymmetric electric field $E_d(t) = E_d f(t)$ (E_d is the drift field amplitude and $f(t)$ is shape of the field) satisfying the condition

$$\int_t^{t+T} f(t)dt = 0, \quad \frac{1}{T}\int_t^{t+T} f^{2n+1}(t)dt \equiv \langle f^{2n+1}\rangle \neq 0 \quad (2)$$

($n \geq 1$ is an integer), ions of particular sort in the gas execute fast oscillations with a period T, drifting along the electric field lines with a characteristic mean velocity V proportional to $\alpha(E/N)$ [10]. This drift can be

1063-7842/02/4711-1453$22.00

compensated for by a constant electric field E_c; then, $V = 0$. Substituting (1) into the equation

$$V = K(E_d(t) - E_c) = 0 \quad (3)$$

and using the approximation

$$(E_d - E_c)^n \approx E_d^n - nE_d^{n-1}E_c \quad \text{at} \quad |E_d| \gg |E_c|, \quad (4)$$

we find an expression for the compensating field [11]:

$$E_c \approx \frac{\left(E_d \sum_{n=1}^{\infty} \alpha_{2n} \left(\frac{E_d}{N}\right)^{2n} \langle f^{2n+1} \rangle\right)}{\left(1 + \sum_{n=1}^{\infty} (2n+1)\alpha_{2n} \left(\frac{E_d}{N}\right)^{2n} \langle f^{2n} \rangle\right)}. \quad (5)$$

From (5), it follows that E_c depends on the coefficients α_{2n} and on the drift field parameters $\langle f^{2n} \rangle$, $\langle f^{2n+1} \rangle$, and E_d. Calculating the values of $\langle f^{2n} \rangle$ and $\langle f^{2n+1} \rangle$ and experimentally finding the dependence of the compensating voltage on the amplitude of the polarity-asymmetric voltage $U_c(U_d)$, one can obtain the dependence $E_c(E_d)$ and find the coefficients α_{2n} by solving the inverse problem for Eq. (5).

EXPERIMENTAL

The setup used in our experiments is depicted in Fig. 1. Its mechanical part, developed by the Mine Safety Appliances Company [12], consists of heated ionization chamber *1* (^{63}Ni β source, $t_i = 150$°C); drift chamber *2*, formed by two coaxial cylindrical electrodes of radii $r_1 = 1.4$ cm and $r_2 = 1.8$ cm and purged by a carrier gas (purified dry air, water vapor concentration <50 ppm, flow rate $Q_g = 30$ cm^3/s, temperature $t_g = 21 \pm 1$°C, $N = (2.5 \pm 0.1) \times 10^{19}$ cm^{-3}); extracting electrode *3* (±200 V), which provides ion transport from the ionization chamber to the drift chamber; compensating voltage source *4*; generator *5* of variable periodic polarity-asymmetric voltage, which is connected to the coaxial electrodes; electrometric amplifier *6* (noise amplitude 1×10^{-14} A); and collector *7*.

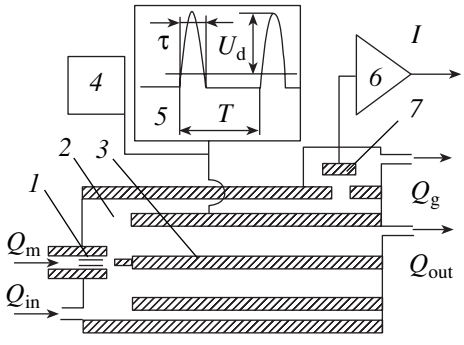


Fig. 1. Experimental setup: Q_m, vapor–air mixture flow; Q_{in}, incoming flow of pure gas; Q_g, carrier gas flow; Q_{out}, outward flow.

The parameters of the generator (designed in the Design and Technological Institute of Instrument Engineering for Geophysics and Ecology) are as follows:

$$f(t) = (\sin[\pi(t - mT)/\tau] - 2\tau/\pi T)/(1 - 2\tau/\pi T) \quad \text{for } mT \le t \le (mT + \tau);$$
$$f(t) = -(2\tau/\pi T)/(1 - 2\tau/\pi T) \quad \text{for } (mT + \tau) \le t \le (m+1)T \quad (6)$$

($m \ge 0$ is an integer), the voltage amplitude range $U_d = \pm(600\text{–}2900)$ V, the voltage period $T = 7.1$ μs, and the voltage pulse duration $\tau = 2.7$ μs. The compensating voltage U_c is measured with an accuracy of no worse than $\delta U_c = \pm 5\%$; the polarity-asymmetric voltage amplitude, with an accuracy of no worse than $\delta U_d = \pm 7\%$.

Vapor–air mixtures of the nitrocompounds were prepared by passing purified air (flow rate $Q_m = 1.2$ cm^3/s) through a quartz tube (diameter 0.3 cm, length 20 cm) with its inner wall covered by the reagent-grade compound to be studied. An air–DMMP mixture was obtained by passing the purified air through a cuvette with a diffusion tube containing 97% pure DMMP (Aldrich cat., no. D16910). Then, the mixture was delivered to the ionization chamber.

STUDIED IONS

The physicochemical properties of the compounds considered, as well as the types of ion–molecular reactions and ions produced by β ionization under atmospheric pressure, is well known. Specifically, in the presence of nitrogroups, the methyl group of DNT and TNT exhibits acidic properties, while negative reacting ions $(H_2O)_k O_2^-$ ($k \ge 0$ is an integer) behave as Brönsted bases. Therefore, negative ions $(M\text{–H})^-$ of these compounds (M is a molecule of the compound, H is a hydrogen atom) are formed through ion–molecular reactions of proton detachment [13, 14]. DNB and TNB molecules have a high electron affinity, as a result of which these compounds produce largely M^- ions via electron-exchange reactions [15, 16]. Because of a high proton affinity of DMMP, its positive ions $(M + H)^+$ are produced in proton-exchange reactions [17].

DEPENDENCE OF THE COMPENSATING VOLTAGE U_c ON THE AMPLITUDE OF THE POLARITY-ASYMMETRIC VOLTAGE U_d

To obtain the $U_c(U_d)$ curve experimentally, one should first record ion current (I) vs. compensating voltage U_c spectra for various U_d. Such a spectrum for DMMP $(M + H)^+$ ions for different U_d is given in Fig. 2. Current peaks for these ions are observed when rela-

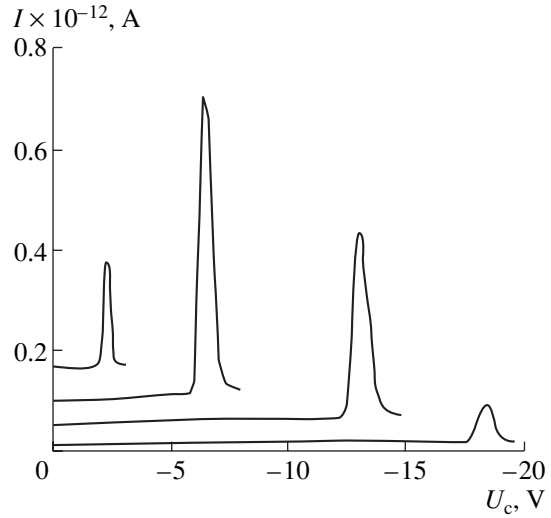


Fig. 2. Spectra of DMMP $(M + H)^+$ ions for a drift voltage U_d = 1070, 1620, 2140, and 2560 V (from left to right).

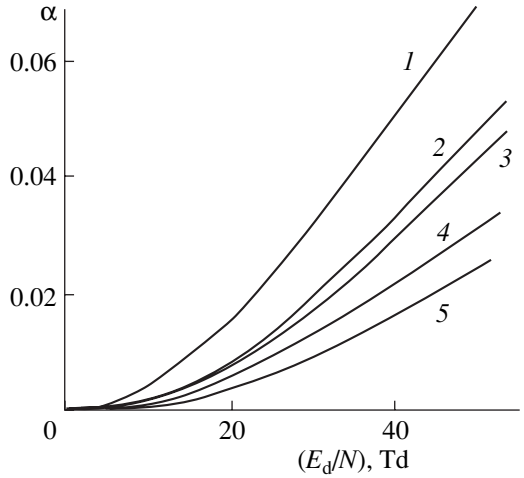


Fig. 3. Increment of the ion mobility coefficient vs. E_d/N: (*1*) DMMP, (*2*) DNB, (*3*) DNT, (*4*) TNB, and (*5*) TNT.

tionship (5) is met. For a given U_d, one fixes U_c at which the peak is observed. In this example for DMMP $(M + H)^+$ ions, the absolute value of the compensating voltage, $|U_c|$, grows with the drift voltage amplitude. The positive value of U_d and the negative value of U_c for the positive ions indicate that $\alpha(E/N)$ has a positive sign and the coefficient of mobility increases with U_d. For the negative ions, $\alpha(E/N) > 0$ when $U_d < 0$ and $U_c > 0$.

Table 1 summarizes experimentally found U_c for ions of the compounds under study for different U_d.

CALCULATION OF COEFFICIENTS α_{2n} AND RECOVERY OF DEPENDENCES $\alpha(E)$

The coefficients $\langle f^{2n} \rangle$ and $\langle f^{2n+1} \rangle$ for the generator [see (6)] were found by numerical calculation at τ/T = 0.38 (Table 2). The coefficients α_2, α_4, α_6, and α_8 (Table 3) were calculated by solving the inverse problem for Eq. (5) with regard for the experimental data (Table 1) expressed in units of field strength (with the equation $U = E(r_2 - r_1)$ and the coefficients $\langle f^{2n} \rangle$ and $\langle f^{2n+1} \rangle$ (Table 2). The coefficients α_{2n} were calculated by means of a polynomial approximation with the least

Table 1. Experimental values of U_c vs. U_d

U_d, V	E/N, Td	U_c, V, DNB	U_c, V, DNT	U_c, V, TNB	U_c, V, TNT	U_d, V	E/N, Td	U_c, V, DMMP
−530	10.6	0.12	0.1	–	–	530	10.6	−0.2
−680	13.6	0.3	0.3	0.14	0.12	680	13.6	−0.5
−830	16.6	0.6	0.6	0.34	0.3	850	17	−1
−1020	20.4	1.2	1.1	0.7	0.64	1070	21.4	−2.1
−1160	23.2	1.8	1.7	1.2	0.9	1350	27	−3.9
−1350	27	2.8	2.7	1.9	1.5	1620	32.4	−6.4
−1570	31.4	4.4	4.2	3	2.4	1900	38	−9.5
−1730	34.6	5.7	5.5	3.9	3.1	2140	42.8	−13
−1900	38	7.3	6.9	5	3.9	2320	46.4	−15.8
−2060	41.2	9	8.5	6.1	4.7	2560	51.2	−19.7
−2200	44	10.8	9.7	7.3	5.5	2840	56.8	−24.5
−2330	46.6	12.3	11.3	8.4	6.2			
−2500	50	14.6	13.1	9.6	7.6			
−2670	53.4	16.9	15.4	11.4	–			

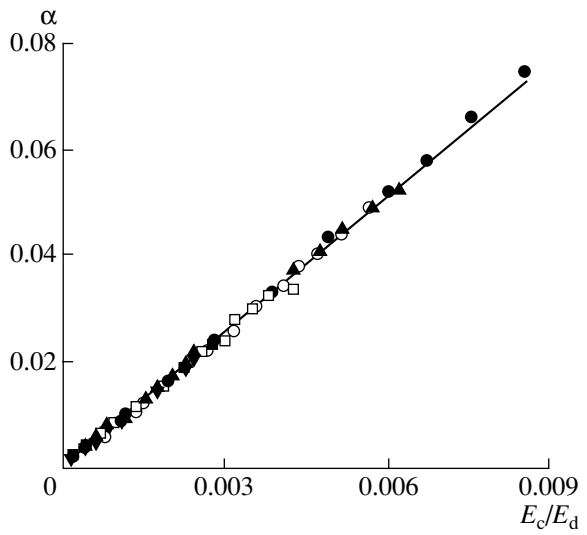


Fig. 4. Calculated values of the coefficient α vs. E_c/E_d (or U_c/U_d). (●) DMMP, (▲) DNB, (○) DNT, (□) TNB, and (◆) TNT and the regression line $\alpha = (K - K_0)/K_0 = 8.354E_c/E_d$ with an rms deviation $\sigma_s = 0.00074$.

squares method (polynomial regression). Since the coefficient α_2 defines the initial slope of the dependence of the ion mobility K on the parameter $(E/N)^2$ [1], it was calculated in the interval $10.6 \le E/N \le 21.4$ Td. When finding the coefficient α_4, we used the value of α_2 found. The coefficient α_4 was calculated in the interval $E/N \le 32.4$ Td. The coefficients α_6 and α_8 were determined using the previously found coefficients α_2 and α_4 and the experimental data for the entire range of E/N. Such a partition made it possible to minimize the rms deviations σ_E of data points from theoretical curves $U_c(U_d)$ recovered using the coefficients α_{2n}. The numerical values of σ_E are listed in the last column of Table 3.

Figure 3 shows dependences $\alpha(E/N)$ calculated by Eq. (1) with the coefficients from Table 3. In the entire range of E/N studied, $\alpha(E/N)$ grows in the order DMMP > DNB > DNT > TNB > TNT.

ERRORS IN DETERMINING THE DEPENDENCE $\alpha(E)$

In Fig. 4, the calculated coefficients α for all the ions are plotted against experimental values of E_c/E_d (or U_c/U_d). The solid curves are the regression line. The total number of measurements is $R = 64$. It is seen that the data points fit well a straight line with a slope $S = 8.354 = 1/0.1197$. According to (5) and Table 2, we can put $S \approx 1/\langle f^3 \rangle$; therefore, for the ions studied and for $E/N = 10$–50 Td, Eq. (5) can be represented in the form

$$\alpha = \frac{SE_c}{E_d} \approx \frac{U_c}{\langle f^3 \rangle U_d}. \qquad (7)$$

From (7), it follows that the total error σ_α in determining the increment of the ion mobility coefficient depends on the relative errors in measuring the voltages U_c and U_d, as well as on the rms deviation of data points from the regression line σ_s ($\sigma_s = \pm 0.00074$). To find the relative error of the linear approximation, it is necessary to divide σ_s by α; then, the total error σ_α is calculated from the expression

$$\sigma_\alpha = \sqrt{\left(\frac{\sigma_s}{\alpha} \times 100\%\right)^2 + \delta U_d^2 + \delta U_c^2}, \qquad (8)$$

in other words, the error σ_α involved in α depends on the absolute value of α. For example, the relative errors in determining the increments $\alpha = 0.005$ and 0.01 are, respectively,

$$\sigma_{\alpha = 0.005} = \sqrt{\left(\frac{0.074}{0.005}\right)^2 + 7^2 + 5^2} \approx 17\%,$$

Table 2. Coefficients $\langle f^{2n} \rangle$ and $\langle f^{2n+1} \rangle$ for generator (6)

Generator	$\langle f^2 \rangle$	$\langle f^3 \rangle$	$\langle f^4 \rangle$	$\langle f^5 \rangle$	$\langle f^6 \rangle$	$\langle f^7 \rangle$	$\langle f^8 \rangle$	$\langle f^9 \rangle$
$\frac{\tau}{T} = 0.38$	0.229	0.119	0.130	0.110	0.104	0.096	0.090	0.085

Table 3. Coefficients α_{2n} for the ions studied

Compound	$\alpha_2 \times 10^{-5}$, Td^{-2}	$\alpha_4 \times 10^{-9}$, Td^{-4}	$\alpha_6 \times 10^{-12}$, Td^{-6}	$\alpha_8 \times 10^{-16}$, Td^{-8}	σ_E, V
DNB	2.39	0.47	–1.24	1.36	0.11
DNT	2.17	3.05	–3.51	6.47	0.11
TNB	1.48	1.53	–0.96	0.06	0.07
TNT	1.18	1.61	–1.65	2.79	0.13
DMMP	3.55	–3.11	–0.84	1.83	0.08

$$\sigma_{\alpha = 0.01} = \sqrt{\left(\frac{0.074}{0.01}\right)^2 + 7^2 + 5^2} \approx 11\%.$$

As was noted above, the error in measuring the coefficient of mobility K with conventional drift tubes equals 1% in the best case, which corresponds to the absolute value $\sigma_{K=2} \approx \pm 0.01$ cm^2/(V s). For DMMP $(M\text{–H})^+$ ions, $K_0 = 1.95$ cm^2/(V s) [17]. Let, for example, the increment $\alpha = (K - K_0)/K_0 = (1.97 - 1.95)/1.95 \approx 0.01$; then, the error in drift-tube measurements of such an increment will be ±50%, i.e., five times greater than in measurements with a variable periodic polarity-asymmetric electric field.

CONCLUSION

It is shown that the use of a variable periodic polarity-asymmetric electric field for determining the increment of ion mobility coefficients that only slightly depend on the field strength improves the measurement accuracy several times as compared with conventional drift tubes.

REFERENCES

1. E. A. Mason and E. W. McDaniel, *Transport Properties of Ions in Gas* (Wiley, New York, 1988).
2. A. A. Radtsig and B. M. Smirnov, in *Chemistry of Plasma*, Ed. by B. M. Smirnov (Énergoatomizdat, Moscow, 1984), Vol. 1, pp. 170–200.
3. I. A. Buryakov, E. V. Krylov, A. L. Makas', *et al.*, Zh. Anal. Khim. **48**, 156 (1993).
4. E. W. McDaniel and E. A. Mason, *The Mobility and Diffusion of Ions in Gases* (Wiley, New York, 1973; Mir, Moscow, 1976).
5. H. Bohringer, D. W. Fahey, W. Lindinger, *et al.*, Int. J. Mass Spectrom. Ion Processes **81**, 45 (1987).
6. B. M. Smirnov, *Complex Ions* (Nauka, Moscow, 1983).
7. I. A. Buryakov, E. V. Krylov, E. G. Nazarov, *et al.*, Int. J. Mass Spectrom. Ion Processes **128**, 143 (1993).
8. I. A. Buryakov, E. V. Krylov, A. L. Makas', *et al.*, Pis'ma Zh. Tekh. Fiz. **17** (2), 60 (1991) [Sov. Tech. Phys. Lett. **17**, 446 (1991)].
9. L. A. Viehland, R. Guevremont, R. W. Purves, *et al.*, Int. J. Mass Spectrom. **197**, 123 (2000).
10. M. P. Gorshkov, USSR Inventor's Certificate No. 966583, Int.C1. G 01 N 27/62 (1982).
11. I. A. Buryakov, Int. J. IMS **4** (2), 112 (2001).
12. B. L. Carnahan and A. S. Tarasov, US Patent No. 5,420,424 (1995).
13. G. E. Spangler, J. P. Carrico, and D. N. Campbell, J. Test. Eval. **13**, 234 (1985).
14. K. A. Daum, D. A. Atkinson, and R. G. Ewing, Int. J. IMS **4** (2), 179 (2001).
15. P. Kebarle, Int. J. Mass Spectrom. **200**, 313 (2000).
16. K. A. Daum, D. A. Atkinson, R. G. Ewing, *et al.*, Talanta **54**, 299 (2001).
17. S. N. Ketkar, J. G. Dulak, and W. L. Fite, Anal. Chem. **63**, 924 (1991).

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1458–1460. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 114–116.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

On the Experimental Verification of the Semiconductor Model of Detonation

K. F. Grebenkin, A. L. Zherebtsov, A. L. Kutepov, and V. V. Popova

All-Russia Research Institute of Technical Physics, Russian Federal Research Center, Snezhinsk, Chelyabinsk oblast, 456770 Russia

e-mail: greb@nine.ch70.chel.su

Received May 14, 2001

Abstract—An experiment for verifying the semiconductor model of detonation is described. Within this model, the electron thermal conductivity is the decisive parameter in initiating the detonation of 1,3,5-triamino-2,4,6-trinitrobenzene (TATB), a low-sensitivity explosive. The idea of the experiment is to detect an increase in the electron concentration in the conduction band and to estimate the electron thermal conductivity from the electrical conductivity of TATB compressed by shock waves of subthreshold intensity. *© 2002 MAIK "Nauka/Interperiodica".*

It is known [1] that chemical reactions attendant on the shock-wave initiation of the detonation of condensed heterogeneous explosives start at so-called "hot spots"—microsites of combustion that appear because of a local temperature increase when the shock wave passes through pores, cracks, and other inhomogeneities. The volume fraction of microsites is small [2], and the explosive comes into reaction gradually as a combustion wave propagates from them.

According to the recent semiconductor model of initiating detonation in low-sensitivity TATB, a combustion wave propagates from microsites owing to electron heat conduction. The basic parameter of the model is the energy gap ε_g of a crystalline explosive [3, 4], which specifies the electron concentration in the conduction band and, hence, the electron thermal conductivity of the shock-compressed explosive.

The available information about the energy gap of crystalline TATB is incomplete and contradictory even under normal conditions. In [5], ε_g was estimated at 10 eV by the Hartree–Fock method. The estimation of the electron density by the method of density functional gave ε_g = 2–4 eV [6] and showed that ε_g tends to decrease with increasing pressure. The recently obtained experimental value of the TATB energy gap was 6.6 eV [7]; however, these measurements were performed on thin films (of thickness less than 1 μm) and the validity of applying these values to explosive single crystals with a thickness of tens to hundreds of micrometers is doubtful.

High-pressure measurements of the TATB energy gap have not been carried out. For many molecular crystals, the energy gap decreases with increasing pressure (see, e.g., [8]); it appears that the same takes place in crystalline explosives [6, 8].

In this work, we describe an experiment aimed at verifying the hypothesis that is central to the semiconductor model of detonation [3, 4]. According to this hypothesis, the TATB energy gap is ε_g = 1.5–2.0 eV at pressures of 10–20 GPa. If this hypothesis is valid, the electron concentration in the conduction band must be high after the passage of a shock wave. The idea of the experiment is to detect a possible increase in the electron concentration in the conduction band by measuring the electrical conductivity of TATB compressed by shock waves of subthreshold intensity. From the electrical conductivity measurements, one can estimate the electron thermal conductivity of the shock-compressed explosive with the Wiedemann–Franz law [9], because the electron thermal conductivity is the basic parameter responsible for detonation initiation in this model.

In fact, the delay $\Delta\tau$ of detonation initiation after the passage of a shock wave depends on the wave propagation velocity D from hot spots: $\Delta\tau \approx \delta/D$, where δ = 10–100 μm is the characteristic spacing between hot spots, which is on the order of the size of explosive microcrystals. For typical delays of initiation of 0.1–1.0 μs, the combustion wave propagation velocity from hot spots is estimated at D = 10–100 m/s. The wave velocity depends on the thermal diffusivity χ of the shock-compressed explosive and reaction rate τ_r of the explosive in the combustion wave [10]: $D \approx (\chi/\tau_r)^{1/2}$.

Thus, to find the explosive thermal diffusivity corresponding to the combustion wave velocities expected, one must know the reaction time of the explosive in the combustion zone, which propagates from microsites at high pressures and temperatures (3–5 GPa and 3000 K for usual explosives, such as hexogen or octogen, and 10–20 GPa and 2500 K for low-sensitivity TATB [1]). Note that experimental data for explosive reaction

1063-7842/02/4711-1458$22.00

times under such conditions are absent and the extrapolation of explosive reaction times measured at lower pressures and temperatures is unjustified and may introduce large errors.

In this work, we estimated the explosive reaction time by the method of classical molecular dynamics. Such an approach is valid, since the ambient temperature under these conditions exceeds or is comparable to typical vibration energy quanta of explosive molecules. Calculations were performed for nitromethane—the simplest explosive in the class of nitrocompounds. Reaction times for other explosives were evaluated by varying the energy of C–N bond breaking. Molecular dynamics calculations were carried out with the Tinker program [11] with potentials taken from [12], and valence bonds were simulated with the Morse potential. Two values of the C–N bond breaking energy were taken: 60 and 40 kcal/mol. The first value is for nitromethane and low-sensitivity TATB; the second, for usual explosives like octogen and hexogen [13]. Atomic charges were selected so as to approximate the potential distribution near a nitromethane molecule as closely as possible. This distribution was obtained for an isolated nitromethane molecule by means of quantum calculations using the GAMESS program [14] in the Hartree–Fock approximation with the N21-3G basis. In the basic series of calculations, we simulated the motion of 128 molecules in a cube with periodic boundary conditions. The doubling of the number of molecules affects the result insignificantly. The molecular dynamics calculations (NPT ensemble) were performed with a time step of 0.1 fs.

The method was checked by performing calculations for points on the shock adiabat of nitromethane. A good fit with the experimental D–u relationship [15] was obtained.

When estimating the explosive reaction time, we assumed (see, e.g., [16]) that the limiting stage controlling the overall reaction time is C–N bond breaking, since this bond is the weakest in molecules of explosives based on nitrocompounds. From the molecular dynamics calculations, we established the time dependence of the number of dissociated molecules (with broken C–N bonds). The bonds were assumed to be broken when the spacing between the C and N atoms reached 3 Å (C–N bond doubled length).

Figure 1 shows the calculated time dependence of the number of dissociated molecules for the two values of the dissociation energy. For usual explosives like hexogen and octogen, the characteristic time of decomposition in the combustion wave is 1 ps; for low-sensitivity TATB, 10 ps. Figure 2 demonstrates the calculated reaction time vs. pressure and temperature for the low-sensitivity explosive.

Based on the reaction time estimated, one can evaluate the thermal diffusivity of shock-compressed TATB: $\chi \approx D^2\tau_r = 10^{-9}$–$10^{-7}$ m^2/s. On the assumption that the energy in the combustion wave is transferred by

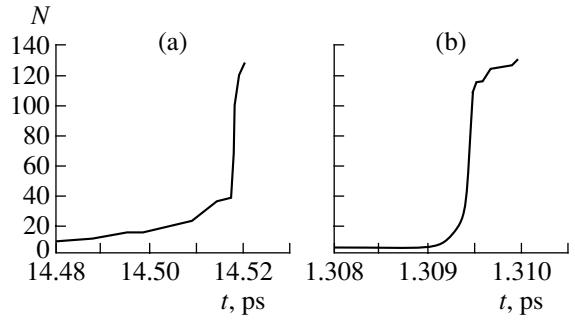


Fig. 1. Time dependence of the number of dissociated molecules: (a) low-sensitivity explosive (the dissociation energy 60 kcal/mol), P = 10 GPa, T = 2500 K and (b) standard explosive (dissociation energy 40 kcal/mol), P = 3 GPa, T = 3000 K.

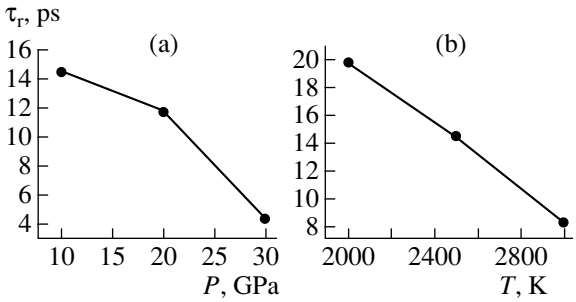


Fig. 2. Reaction time vs. (a) pressure (T = 2500 K) and (b) temperature (P = 10 GPa) for the low-sensitivity explosive (dissociation energy 60 kcal/mol).

means of electron heat conduction, the above interval of the thermal diffusivity corresponds, according to the Wiedemann–Franz law [9], to the electrical conductivity of shock-compressed TATB in the interval 10^2–10^4 $(\Omega\ \mathrm{m})^{-1}$ at pressures of 10–15 GPa, which are close to the TATB detonation threshold.

Thus, according to the semiconductor model, the TATB electrical conductivity is expected to sharply increase up to a value typical of semiconductors under conditions of shock wave compression. The experimental corroboration of this effect is of great importance for gaining a better insight into fundamental mechanisms of detonation initiation in low-sensitivity explosives.

Calculated atomic charges in a nitromethane molecule

Atom	Charge (in terms of unit charge)
C	–0.3746
N	0.6898
O	–0.3872
H	0.1527

REFERENCES

1. G. I. Kanel', S. V. Razorenov, A. V. Utkin, and V. E. Fortov, *Impact–Wave Phenomena in Condens Media* (Yanus-K, Moscow, 1996).
2. J. Sharma, J. W. Forbes, C. S. Coffey, and T. P. Liddiard, J. Phys. Chem. **91**, 5139 (1987).
3. K. F. Grebenkin, Pis'ma Zh. Tekh. Fiz. **24** (20), 1 (1998) [Tech. Phys. Lett. **24**, 789 (1998)].
4. K. F. Grebenkin, in *Proceedings of the International Conference "V Zababakhin Scientific Lecturing"* (RFYaTs-VNIITF, Snezhinsk, 1999), Part 1, p. 189.
5. A. B. Kunz, Phys. Rev. B **53**, 9733 (1996).
6. K. F. Grebenkin and A. L. Kutepov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 1212 (2000) [Semiconductors **34**, 1161 (2000)].
7. S. Kakar, A. J. Neison, R. Treusch, *et al.*, Phys. Rev. B **62**, 15666 (2000).
8. M. M. Kuklja, E. V. Stefanovich, and B. A. Kunz, J. Chem. Phys. **112**, 3417 (2000).
9. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976; Nauka, Moscow, 1978).
10. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
11. *TINKER Software Tools for Molecular Design. User's Guide* (2000).
12. S. Mayo, B. Olafson, and W. Goddard, J. Phys. Chem. **94**, 8897 (1990).
13. C. J. Wu and L. E. Fried, *A Comparative Study of the Bond Dissociation Energy for High Explosive Molecules*, Lawrence Livermore National Laboratory Preprint UCRL-JC-127877 (1977).
14. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, *et al.*, J. Comput. Chem. **14**, 1347 (1993).
15. R. Chaiken, in *Proceedings of the USA Symposium H.D.P., Paris, 1978.*
16. M. Kamlet, in *Proceedings of the Sixth Symposium (International) on Detonation, Coronado, 1976.*

Translated by V. Isaakyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1461–1462. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 117–118.

BRIEF COMMUNICATIONS

Generation of Intense Shock Waves by an Electric Discharge in Gases

K. V. Korytchenko, Yu. Ya. Volkolupov, M. A. Krasnogolovets, M. A. Ostrizhnoĭ, and V. I. Chumakov

Kharkov State Technical University of Radio Engineering, Kharkov, 61166 Ukraine

Received March 13, 2002

Abstract—A device for generating intense shock waves at low energies of an electric discharge in a gas is suggested and tested. © *2002 MAIK "Nauka/Interperiodica".*

The generation of intense shock waves by an electric discharge is a challenge because the energy distribution of gas molecules over the degrees of freedom is severely nonuniform. Experimental data show that more than 95% of discharge energy evolves in the form of vibratory energy of gas molecules, while a shock wave is generated by the kinetic energy. Moreover, a part of the discharge energy is spent on the dissociation and ionization of gas molecules.

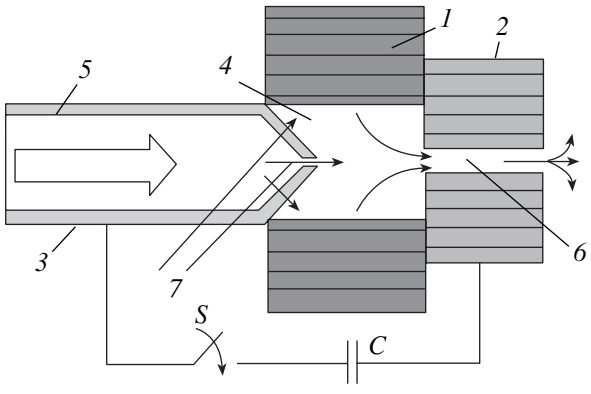
For example, if it is necessary to produce a shock wave in air under normal conditions with a pressure drop at the front of 10 MPa, the local pressure induced by an electric discharge should also be no less than 10 MPa. As follows from the laws of classical thermodynamics, the kinetic temperature of gas molecules in the discharge zone must be no less than 30 000 K. However, due to the almost complete dissociation of air molecules, the necessary pressure can be reached at a kinetic temperature of about 15 000 K, with the air specific heat increasing more than fourfold. Even though the thermodynamic equilibrium distribution of the discharge energy over air molecules took place, the necessary discharge energy would increase several times. Because of the short time of the shock wave formation (comparable to the molecule mean free time), the growth of the local pressure is provided by cutting the discharge duration. An increase in the discharge energy and a simultaneous decrease in the discharge duration leads to an increase in the discharge electron temperature. This, in its turn, results in a more intense energy absorption by the molecules. The internal (vibratory and electron) molecular energy rises, which is accompanied by dissociation and ionization. Eventially, the discharge energy must be raised still further.

In the case of discharge in an open space, only a small portion of the energy (up to 3%) is spent on shock wave production. Moreover, as the volume power of the energy release increases, the efficiency of discharge energy transformation into a shock wave decreases after a small rise.

Thus, to produce intense shock waves using an electric discharge, one should resolve the contradiction between the need for a considerable energy release for a short time and the need for the effective transformation of the discharge energy into the kinetic energy of gas molecules.

Let us consider a method of generating intense shock waves at small energies of an electric discharge. It relies on the formation of a guided shock wave and uses the dynamic properties of gas flows. This method was implemented with the device schematically shown in the figure.

The device operates as follows. Cooled air flows under pressure from an external source through inner channel *5* of conductor *3* and fills cavity *4* of insulator *1*, passing hole *7*. The geometric sizes of outlet *6* and inlet *5* with hole *7* are chosen in such a way that the pressure in cavity *4* of insulator *1* is higher than the ambient pressure. The air pressure drops to the ambient pressure in and behind channel *6*. The maximum velocity of the gas in channel *6* cannot exceed the critical velocity of sound C_{*2}. The velocity of disturbance



Device for generating shock waves by electric discharge in gas: (*1*) insulator; (*2*) conductor; (*3*) leading-in conductor; (*4*) cavity; (*5*) entrance channel; (*6*) outlet channel; and (*7*) holes.

1063-7842/02/4711-1461$22.00 © 2002 MAIK "Nauka/Interperiodica"

Parameters behind the shock wave front in air under normal conditions before the front at $P_0 = 10^5$ Pa and $T_0 = 293$ K

Temperature T, K	Velocity D, m/s	Pressure P_1, 10^5 Pa	Density ρ_1/ρ_0
3840	2935	100	6.76

propagation in the gas will vary from C_0 to C_{*2} (C_0 is the velocity of sound in a stagnant flow) [1].

Once the capacitive energy storage has been discharged through switch *S* in cavity *4*, a portion of the discharge energy is directly converted into kinetic energy. As a result, the pressure in cavity *4* grows and a compression wave starts propagating along channel *6*. The inlet pressure prevents the gas motion toward channel *5*. Note that the rate of vibratory-to-kinetic energy conversion is directly proportional to the gas pressure. Therefore, as the pressure in the cavity grows further because of the deactivation, compression waves with an increasing amplitude will propagate along the channel in the partially heated gas with the resulting formation of a shock wave. This process is similar to the formation of a shock wave in detonation tubes [2].

The parameters of our device were analyzed based on the theory of vibratory deactivation of gas molecules [3] with regard for conditions for shock wave formation in detonation tubes and gas properties.

Our design allows one to increase the time of shock wave formation and, hence, to utilize the discharge energy more efficiently by extending the discharge duration in the high-pressure zone.

The table lists the parameters of the device that provide the formation of a shock wave with intensity and under conditions corresponding to the above example.

To generate shock waves in detonation tubes, the laminar gas flow should be provided in the channels. As follows from experimental data, with disturbances at the inlet removed, the laminar flow can be sustained at Reynolds numbers up to $\mathrm{Re} = 10^5$, where $\mathrm{Re} = U_{\max} d/2\nu$, $U_{\max}$ is the liquid axial velocity, d is the tube diameter, and ν is the gas kinematic viscosity (in our case, $U_{\max}$ corresponds to the velocity D of the shock wave leading edge and d, to the diameter of the outlet channel). The maximum temperature T_1 in the insulator cavity can be determined by equating the maximum velocity of the heated air flow to the critical velocity of sound C_{*2}:

$$T_1 = \frac{C_{*2}^2 M(\gamma + 1)}{2\gamma R} = \frac{D^2 M(\gamma + 1)}{2\gamma R}.$$

Here, M is the gas molar weight and R is the universal gas constant.

The discharge duration is found from the time of shock wave formation t:

$$t = \frac{l}{C_{01}} - \frac{l}{C_{02}},$$

where l is the outlet channel length and C_{01} and C_{02} are the velocities of sound in the stagnant flow under initial conditions and in the discharge-heated gas, respectively.

For example, if $d = 0.001$ m and $l = 0.0015$ m, we find that $R \approx 10^4$, $T_1 = 11100$ K, and $t = 4 \times 10^{-5}$ s. For comparison, the time of shock wave formation in air under normal initial conditions is no less than 10^{-8} s.

The device considered can be used as an initiator of detonation waves of various frequency (with a maximum frequency of no less than 500 Hz) in earlier developed supersonic gas-dynamic pulsed lasers or pulsed ramjets [4]. Moreover, this method allows one to modify the design of shock and detonation tubes and put associated devices to scientific and technological use.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
2. F. A. Baum, K. L. Stanyukovich, and B. I. Shekhter, *Physics of Explosion* (Fizmatgiz, Moscow, 1959).
3. Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Fizmatgiz, Moscow, 1963; Academic, New York, 1966).
4. Ukr. Patent No. 2,000,084,769 (August 10, 2000).

Translated by M. Astrov

Technical Physics, Vol. 47, No. 11, 2002, pp. 1463–1464. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 119–120.

BRIEF COMMUNICATIONS

Specific Features of Electrolytic–Plasma Quenching

Yu. N. Tyurin* and A. D. Pogrebnjak**

** Paton Institute of Electric Welding, National Academy of Sciences of Ukraine, Kiev, Ukraine*

*** Sumy Institute of Surface Modification, Sumy, 40030 Ukraine*

Received October 9, 2000; in final form, January 30, 2002

Abstract—Processes in an electrolyte and a plasma layer between the liquid electrode (anode) and the solid surface of a workpiece are analyzed. A method of controlling the surface-heating power density by periodically varying the time of turning on an elevated voltage and the voltage value is grounded. This method allows one to control the rates of heating and cooling the workpiece in the range from 20 to 500°C/s and, correspondingly, to heat surface layers to a depth of 0.1–10 mm. © *2002 MAIK "Nauka/Interperiodica".*

The electrolytic–plasma method of quenching has been known for more than 50 years [1]. A workpiece is immersed into an electrolyte, which is usually an aqueous salt solution. A plasma forms when an electric current passes through the electrolyte near the cathode, which is the workpiece [1–5]. The electrical energy is converted to heat mainly at the cathode (workpiece) surface. It was found experimentally that the energy is spent largely on heating the surface and forming the plasma layer; the loss is only about 15% [6]. Electric discharges appear in the plasma layer [1–5]. Discharge-induced pressure variations cause the surface of the liquid anode to vibrate. As a result, the gap between the surfaces of the liquid electrode and the solid cathode varies. The electric field strength in the electrolyte is not high (up to 80–200 V/m) but can reach 10^4–10^6 V/m in the plasma layer (Fig. 1). The voltage in a 300-mm-thick electrolyte layer was experimentally found to be lower than 25 V, while in a plasma layer 0.1–3.0 mm thick, it reaches 300 V.

Depending on the voltage across the anode–cathode gap, five characteristic modes of heating the surfaces of the liquid electrode and the solid cathode can be distinguished [4].

In the first mode, the voltage is U = 60–120 V. The electrolyte (liquid electrode) surrounds the surface of the cold cathode. The temperature of the cathode surface is lower than the boiling temperature of the electrolyte. The electrical energy is spent mainly on heating the electrolyte. In the second mode (U = 80–160 V), the temperature of the cathode surface exceeds the boiling temperature of the electrolyte and a vapor–gas layer forms at the cathode. In the third mode (U = 120–200 V), the vapor–gas layer adjacent to the solid surface is locally overheated owing to electric discharges. The vibration of the electrolyte surface provides a local increase in the electric field strength to 10^6 V/m and, thus, the formation of microdischarges alternating with cooling the cathode surface by the electrolyte. In the fourth mode (U = 180–260 V), a stable plasma layer of a variable cross section forms between the cathode and the electrolyte within 2–10 s and a stable plasma glow is observed at the cathode.

A further increase in the voltage to 240–320 V provides the virtually instantaneous formation of an intensely glowing plasma layer (in 0.1–0.5 s). The rate of surface heating can reach 500°C/s.

Earlier [7], we proposed combining various modes of heating the liquid electrode and a metallic workpiece. Figure 2 shows the measurement of the surface temperature at the voltages between the electrodes corresponding to the five heating modes. When a voltage of 320 V is turned on (curve *1*), the sample surface is rapidly heated (fifth mode). The rate of heating the surface reaches 150°C/s. The current density equals 8–9 A/cm^2, and the surface layer begins melting within 10–15 s. The periodic transition from the fifth to the third heating mode or periodic switching from the high (320 V) to low (200 V) voltage provides steady-state heating at a mean rate of up to 50°C/s. The mean current density is 6–7 A/cm^2. The heating time can be set

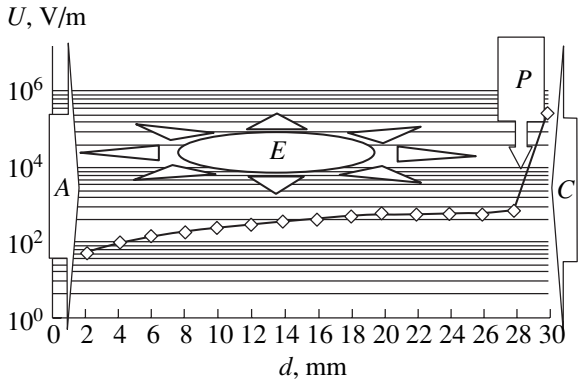


Fig. 1. Variation of the electric field strength between the cathode (*C*) (workpiece surface) and metallic anode (*A*) in the electrolyte (*E*) and plasma layer (*P*).

1063-7842/02/4711-1463$22.00 © 2002 MAIK "Nauka/Interperiodica"

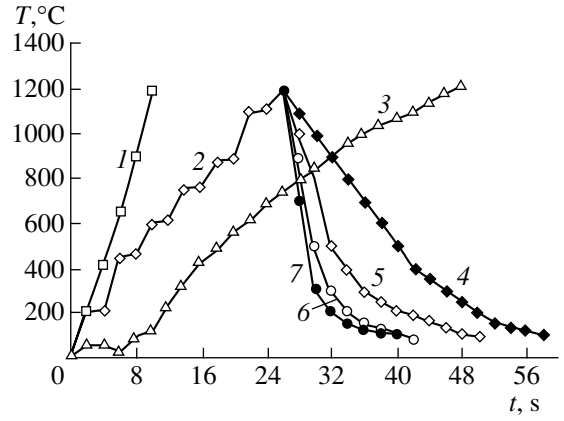


Fig. 2. Cathode surface temperature as a function of (*1–3*) heating time and (*4–7*) cooling time at a voltage of (*1*) 320; (*2*) 200, 320, and 30; and (*3*) 220 V, as well as at a reduced voltage of (*4*) 60, (*5*) 40, (*6*) 20, and (*7*) 0 V upon cooling.

such that rather thick (up to 10 mm) layers can be quenched without melting the surface.

At 200 V, the heating process is initially unstable. The instability time reaches 12 s. The current density changes jumpwise from 0 to 50 A/cm^2. It drops to the nominal value (J = 5–6 A/cm^2) only after the formation of the vapor–gas layer with a relatively low conductivity (σ = 0.5–0.7 Ω^{-1} m^{-1}. Increasing and decreasing the voltage (i.e., alternating high and low power densities), one can heat the workpiece with a mean rate of 30–60°C/s (Fig. 2, curve *2*).

The periodic variation of the voltage changes the mean power density of heating from 1×10^3 to 1×10^4 W/cm^2, which makes it possible to control electrolytic–plasma heating and the formation of 0.1- to 1.0-mm thick quenched layers. Cooling rates can be varied over wide limits by turning on the electric current upon cooling the workpiece surface (Fig. 2, curves *4–7*).

The possibility of controlling the mean power density allows us to use electrolytic–plasma treatment in cleaning, melting, and soldering processes [1–4, 8, 9].

ACKNOWLEDGMENTS

This work was supported in part by the Ukrainian Research Center (project no. 1472) and the Ministry of Education and Science of Ukraine (project no. 2M/0145-2001).

REFERENCES

1. Ya. Z. Yasnogorodskiĭ, *Automatic Heating in Electrolyte* (Oborongiz, Moscow, 1947).
2. A. L. Yerokhin, X. Nie, A. Leyland, *et al.*, Surf. Coat. Technol. **122**, 73 (1999).
3. V. N. Nikitin, K. I. Eretnov, and A. V. Artem'ev, Élektron. Obrab. Mater., No. 2, 35 (1983).
4. D. I. Slovetskiĭ, S. D. Terent'ev, and V. G. Plekhanov, Teplofiz. Vys. Temp. **24**, 353 (1986).
5. V. I. Chernenko, L. A. Snezhko, I. I. Papanova, and K. I. Litovchenko, *Theory and Technology of Anodic Processes under High Voltage* (Naukova Dumka, Kiev, 1995).
6. G. A. Ostroumov, *Interaction of Electric and Hydrodynamic Fields* (Nauka, Moscow, 1979).
7. Yu. N. Tyurin and A. D. Pogrebnjak, Surf. Coat. Technol., Nos. 142–144, 293 (2001).
8. Yu. P. Raizer, *Gas Discharge Physics* (Fizmatgiz, Moscow, 1992; Springer-Verlag, Berlin, 1991).
9. K. N. Eretnev and S. V. Lebedev, *Heating and Cleaning of Metal Surface in Electrolytes and Their Application* (Lipetsk, 1997).

Translated by K. Shakhlevich

Technical Physics, Vol. 47, No. 11, 2002, pp. 1465–1468. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 121–124.

BRIEF COMMUNICATIONS

Dynamics of a Magnetic Flux Jump in a Composite Superconductor

S. L. Kruglov and V. I. Shcherbakov

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

e-mail: sherby@isssph.kiae.ru

Received April 9, 2002

Abstract—An attempt is made at direct experimental verification of the theory of thermomagnetic instability in composite superconductors under conditions of external magnetic field or transport current variations. The development of thermomagnetic instability in the form of a magnetic flux jump is experimentally studied in a bulk low-temperature composite niobium–tin superconductor. The liquid-helium-cooled sample representing a compressed tape helix (helicoid) is placed in an external magnetic field orthogonal to the turn plane and varying with a constant rate. For the first time, both the magnetic induction inside the sample and its temperature are simultaneously detected in experiments. The sample overheat preceding the magnetic flux jump is measured to be 0.23 + 0.02 K. This value is found to be independent of the rate of the external magnetic field variation and the value of the jump itself and coincides, within the experimental accuracy, with the temperature parameter of electric field buildup involved in the general exponential *I–V* characteristic of the composite superconductor, which depends on temperature and magnetic induction. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Thermomagnetic instability in the form of a magnetic flux jump in composite superconductors is related to the interaction of thermal and electromagnetic fluctuations (small perturbations) in superconducting elements incorporated into a normal metal matrix. As a rule, such a jump quenches superconductivity. Theoretically, the stability of superconductivity depends substantially on the relationship between the characteristic times of thermal and electromagnetic processes in a superconductor. The characteristic times involved are magnetic,

$$\tau_m = \frac{\mu_0 d^2}{\rho_\perp}; \quad (1)$$

thermal,

$$\tau_\lambda = \frac{Cd^2}{\lambda}; \quad (2)$$

and heat-exchange,

$$\tau_h = \frac{Cd}{h}. \quad (3)$$

Here, d is the characteristic transverse size of the composite superconductor, $\rho_\perp$ is the transverse resistivity, λ is the thermal conductivity, C is the specific heat per unit volume, and h is the coefficient of heat transfer to a coolant.

Of the most practical interest is the case

$$\tau_m \gg \tau_h \gg \tau_\lambda. \quad (4)$$

Relationships (4) is typical of most composite superconducting structures and meets the conditions of their cooling in superconducting magnetic systems. It means that, in the "dynamic approximation"

$$\tau = \frac{\tau_m}{\tau_\lambda} \longrightarrow \infty, \quad (5)$$

the growth of small temperature and electric field perturbations takes place with the magnetic flux remaining unchanged ("frozen" magnetic flux). This makes it possible to simplify significantly the elaboration of a criterion for the stability of the superconducting state of a composite superconductor. Using the empirical exponential *I–V* characteristic of a superconductor [1, 2]

$$E = E_c \exp\left(-\frac{T_c - T}{T_0} + \frac{B}{B_0} + \frac{j}{j_0}\right) \quad (6)$$

(where E is the electric field intensity in the superconductor; E_c is a constant; T_c is the superconductor critical temperature; T, B, and j are temperature, transverse external magnetic field induction, and current density, respectively; and T_0, B_0, and j_0 are the parameters characterizing a rise in the electric field with temperature, magnetic field, and current density, respectively), Mints and Rakhmanov [3] formulated a criterion for the stability of superconductivity in a composite superconductor. This criterion can be written as

$$\int_s j_c E dS \le hPT_0, \quad (7)$$

1063-7842/02/4711-1465$22.00

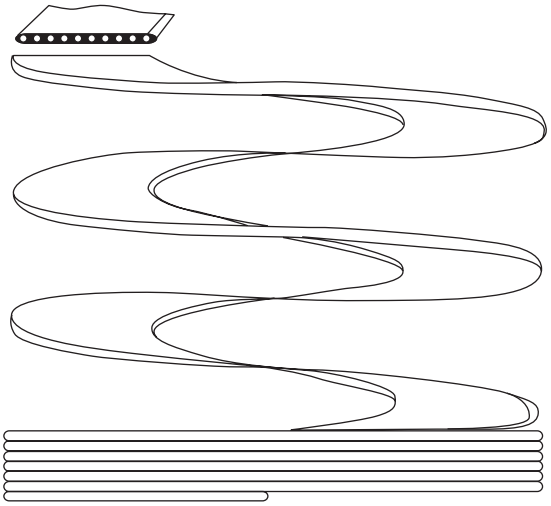


Fig. 1. Superconducting helicoid.

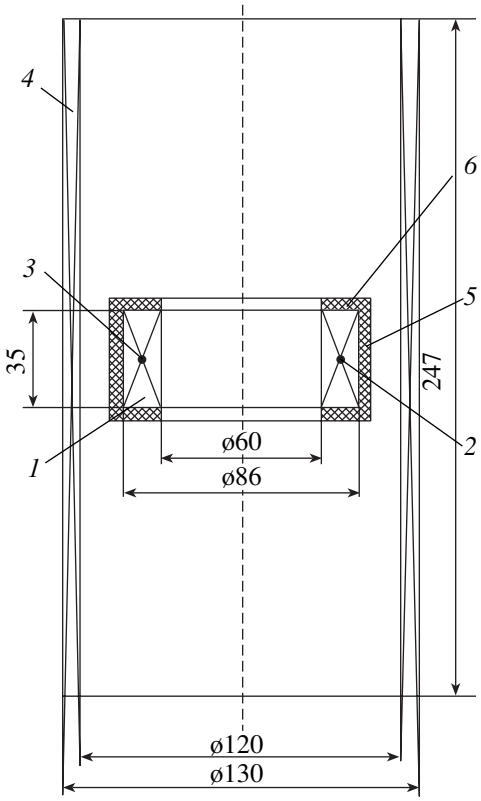


Fig. 2. Longitudinal cross section of the experimental setup: (*1*) sample (superconducting niobium–tin helicoid); (*2*) germanium thermometer; (*3*) Hall-effect sensor; (*4*) superconducting solenoid generating external field; (*5*) heater; and (*6*) thermal insulation.

where E is the longitudinal electric field intensity depending on the rate of variation of external parameters, j_c is the critical current density depending on magnetic field and temperature, S is the cross-sectional area of a composite superconductor, and P is the cooling perimeter of the composite superconductor.

The physical meaning of this criterion is apparent. The superconducting state looses stability when the overheat due to magnetization reversal losses caused by the variation of external parameters (magnetic field, transport current, and temperature) exceeds the value of the temperature parameter of electric field buildup. This fact was verified experimentally. Alternative attempts to verify the criterion [4, 5] were indirect: they involved the construction of experimental curves $\Delta B_j(dB/dt)$ (jump vs. rate of external magnetic field variation) with an adjustable heat-transfer coefficient.

As a sample, we employed a superconducting niobium–tin helicoid (a compressed tape helix [6–8]) (Fig. 1). In the presence of an external magnetic field, the open helicoid (without the transport current) behaves like a bulk composite superconductor, shielding its internal volume by surface currents closing at the ends of the helix.

EXPERIMENTAL SETUP

The sample consisted of 20 flat turns of inner diameter 60 mm, outer diameter 86 mm, and height 35 mm (Fig. 1). Each of the turns contained 14 parallel composite niobium–tin superconductors with a diameter of 0.85 mm joined together in the helical plane by copper plating. The characteristics of an individual superconductor were as follows: production method, bronzing; the fraction of the external copper sheath, 25%; the number of niobium–tin filaments, 14641; the filament diameter, 3 μm; and the rated current density in the field $B = 6$ T, 1.1×10^9 A/m^2.

After two-sided copper plating, the thickness of the flat turn was 1.75 mm. The sample was placed in the working volume of a superconducting solenoid generating a magnetic field of induction up to 4 T. The rate of field rise reached 0.15 T/s. The cross-sectional view of the setup is shown in Fig. 2. The turns of the helicoid were tightly compressed. A heater made of 0.1-mm-thick stainless steel foil, which is used to break the shielding superconducting circuit if necessary, and a heat-insulating layer were glued to the outer cylindrical surface of the sample. The top and bottom surfaces of the helicoid were also covered by a heat-insulating layer. The coolant was in direct contact only with the inner cylindrical surface of the sample. A Hall-effect sensor, measuring the magnetic field, and a germanium thermometer were placed at the central points of the transverse cross section of the sample. The actual thermal, magnetic, and heat-exchange characteristic times of the sample were $\tau_\lambda = 1.5 \times 10^{-3}$ s, $\tau_h = 10^{-2}$ s, and $\tau_m = 3$ s, respectively.

EXPERIMENT

Figure 3 shows the typical evolution of the magnetic flux jump. In the time interval between $t = 0$ and $t =$ 30 s, the surface superconducting currents induced on

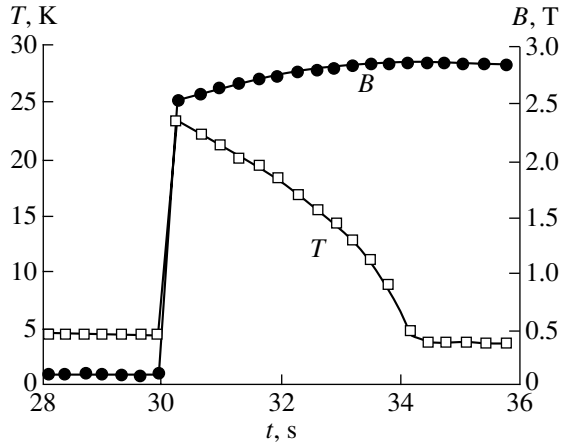


Fig. 3. Internal magnetic field at the central point of the cross section of the sample and the sample temperature versus time. The external magnetic field starts growing ($dB/dt = 8.36 \times 10^{-2}$ T/s) at the zero time moment.

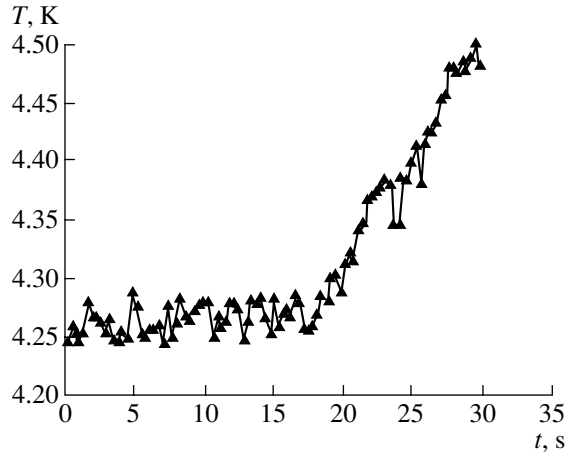


Fig. 4. Sample temperature versus time on an enlarged scale. The time interval extends from the time instant the external magnetic field starts growing ($dB/dt = 8.36 \times 10^{-2}$ T/s) to the instant of the magnetic flux jump.

the cylindrical surfaces of the sample completely shield the external magnetic field. The currents close at the ends of the sample through the copper matrix (the characteristic time of magnetic field diffusion from the ends of the sample to its center (6 h) significantly exceeds the time of measurement). The temperature of the sample is slightly higher than that of the coolant owing to electrical losses. The magnetic flux jump into the sample takes place at $t \approx 30$ s (the measurements were performed every 0.3 s). The energy stored in the shielding circuit is sufficient to quench superconductivity and heat the sample to 23 K. After cooling to the critical temperature (13.6 K), the sample can again shield the still increasing magnetic field. The next shielding period starts from $t \approx 33.5$ s and lasts up to the next jump.

Figure 4 shows temperature variations in the sample on an enlarged scale. The absolute measurement error is no greater than 0.02 K. The sample overheat relative to the coolant temperature at the jump instant ($t \approx 30$ s) is $\Delta T = 0.24$ K. Similar measurements were carried out for four values of the rate of change of the external magnetic field.

Figure 5 plots jump field versus the rate of change of the external field. The temperature parameter of field rise T_0 can be found from the I–V characteristic of a single superconductor of the helicoid. This characteristic usually yields the parameter δ^2 given by

$$\delta^2 = \frac{j_0}{j_c}. \tag{8}$$

This quantity is the relative variance of the normal distribution that characterizes the nonlinearity of the I–V characteristic at low electric fields ($\leq 10^{-4}$ V/m).

The parameters of field buildup are interrelated [2]:

$$j_0 = T_0\left|\frac{\partial j_c}{\partial T}\right|. \tag{9}$$

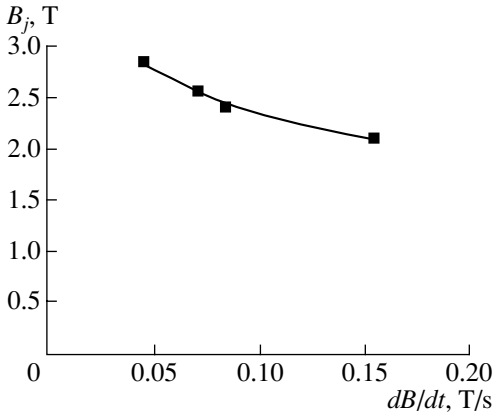


Fig. 5. Magnetic induction of the jump versus the rate of external magnetic field variation.

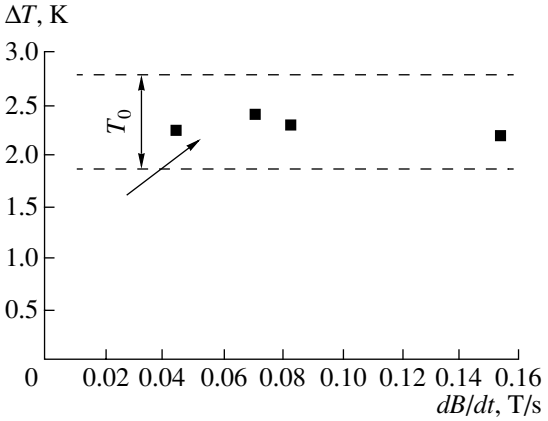


Fig. 6. Sample overheat at the instant of stability breakdown versus the rate of magnetic field variation.

Therefore, knowing δ^2, one can determine T_0 from

$$T_0 = \delta^2(T_c - T_b), \tag{10}$$

where T_b is the helium bath temperature during the measurements of j_c and j_0.

Using the empirical dependence of the critical temperature of the superconductors studied on magnetic induction,

$$T_c(B) = 15.39 - 0.615B, \quad (11)$$

we arrive at the expression for T_0:

$$T_0(B) = \delta^2(15.35 - 0.615B - T_b). \quad (12)$$

For our superconductors, δ^2 ranges from 2 to 3% (for convenience, the actual value of δ^2 is multiplied by 100%). Substituting the jump fields ΔB from Fig. 5 into expression (12) (instead of B), we find the spread of T_0:

$$T_0 = 0.19 - 0.28 \text{ K}.$$

Figure 6 demonstrates the sample overheat ΔT at the instants of magnetic field jumps versus the rate of external magnetic field variation dB/dt. The arrow shows the range of T_0. The independence of ΔT on dB/dt lets us conclude that ΔT and T_0 coincide in experiments.

DISCUSSION

The use of niobium–tin samples allowed us to measure the overheat of a composite superconductor that precedes magnetic flux jumps. Note by way of example that the overheat in niobium–titanium superconductors is less by one order of magnitude (0.02–0.03 K) owing to the significantly lower nonlinearity of the I–V characteristic ($\delta^2 = 0.1$–0.2%). This substantially complicates the measurements.

The experimental facts that the overheat of a bulk composite superconductor preceding a magnetic flux jump is independent of the rate of external magnetic field variation and that the overheat coincides with the temperature parameter of field buildup validate criterion (7) for the stability of superconductivity under varying conditions.

ACKNOWLEDGMENTS

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

REFERENCES

1. M. Polak, J. Hlashik, and L. Krempasky, Cryogenics **13**, 702 (1973).
2. G. L. Dorofejev, A. B. Imenitov, and E. Yu. Klimenko, Cryogenics **20**, 307 (1980).
3. R. G. Mints and A. L. Rakhmanov, J. Phys. D **15**, 2297 (1982).
4. E. Yu. Klimenko, N. N. Martovetskiĭ, and S. I. Novikov, Dokl. Akad. Nauk SSSR **261**, 1350 (1981) [Sov. Phys. Dokl. **26**, 1180 (1981)].
5. S. L. Kruglov and V. R. Romanovskiĭ, Pis'ma Zh. Tekh. Fiz. **20** (22), 89 (1994) [Tech. Phys. Lett. **20**, 929 (1994)].
6. V. E. Keĭlin, I. A. Kovalev, S. L. Kruglov, *et al.*, Dokl. Akad. Nauk SSSR **303**, 1366 (1988) [Sov. Phys. Dokl. **33**, 959 (1988)].
7. I. A. Kovalev and S. L. Kruglov, Zh. Tekh. Fiz. **65** (11), 150 (1995) [Tech. Phys. **40**, 1170 (1995)].
8. S. L. Kruglov, Zh. Tekh. Fiz. **67** (2), 110 (1997) [Tech. Phys. **42**, 220 (1997)].

Translated by A. Chikishev

Technical Physics, Vol. 47, No. 11, 2002, pp. 1469–1472. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 125–128.

BRIEF COMMUNICATIONS

Dynamic Gas Anisotropy under Optomechanical Parametric Resonance

S. V. Siparov

Academy of Civil Aviation, St. Petersburg, 196210 Russia

e-mail: sergey@siparov.spb.su

Received March 26, 2002

Abstract—An approximate kinetic equation describing a gas of two- or three-level atoms in an external laser field is found under conditions of optomechanical parametric resonance, which causes atomic oscillations along the field wave vector. The approximation is applicable to a gas of cylindrical particles whose axes retain their spatial orientation, making the gas anisotropic. The closed set of equations for macroparameters allows for a numerical solution. Expressions for direction-dependent transfer coefficients such as diffusion, viscosity, and thermal conductivity are derived phenomenologically. A two-temperature medium is shown to arise if a fluid consists of several gases with the optomechanical parametric resonance conditions satisfied for one of them. © *2002 MAIK "Nauka/Interperiodica".*

INTRODUCTION

An anisotropic fluid, or a liquid crystal, is a well-known and thoroughly studied system [1]. A gas may also exhibit the anisotropy of, at least, its electrical properties if it consists of dipole molecules subjected to an electrostatic field. However, such an anisotropy does not influence noticeably the transfer characteristics of a gas. Below, we will discuss an anisotropic gaseous system arising under conditions of optomechanical parametric resonance (OMPR).

Theoretically, the OMPR effect discovered recently [2–4] shows up as a correlation between electromagnetic radiation and the dynamics of a two- or three-level gas atom. This effect has many intriguing applications [4–6]. In particular, if an atom is placed in the strong field of resonant laser radiation consisting of two modes with close frequencies Ω_1 and Ω_2 (bichromate) and the OMPR conditions are satisfied in view of the Doppler shift, the oscillating force acting on the atom is

$$F = \eta k \alpha_R H(a, \sigma, v) \cos[(\Omega_2 - \Omega_1)t + \psi(v)]. \quad (1)$$

Here, k is the wave vector of the external field; α_R, the Rabi frequency (E is the electric component of the external field and μ is the dipole moment of the atom); and $(\Omega_2 - \Omega_1)$, the detuning of the bichromate. The dimensionless parameter $H(a, \sigma, v)$ is on the order of unity (for other designations, see Appendix). In this case, the atom oscillates along the direction of laser beam propagation. The term "strong field" means that atomic transitions are fully defined by an external field and the occupancy oscillates with the Rabi frequency. In other words, $\alpha_R \gg \gamma$, where γ is the inverse lifetime of an excited state. Under usual laboratory conditions (e.g., for a titanium–sapphire laser with a power of 1 W), $\alpha_R = 10^8$–10^9 s^{-1} corresponds to optical atomic transitions, whereas one of the OMPR conditions is that the bichromate detuning $(\Omega_2 - \Omega_1)$ almost coincides with the Rabi frequency α_R. Another OMPR condition yields [4] the expression for the atomic oscillation amplitude: $A = \gamma/k\alpha_R$. Under the same laboratory conditions, the oscillation amplitude A is on the order of 10^{-9} m. Considering the influence of OMPR on a gas system as a whole, one can expect significant modifications in its properties.

KINETIC APPROACH

In order to describe the gas behavior under OMPR conditions in terms of the molecular kinetic theory, one has to modify the Boltzmann kinetic equation [7] with regard for the existence of a nongradient oscillating force given by Eq. (1). This yields

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial u_x}\left(\frac{F}{m} f\right) = J(f, f). \quad (2)$$

Here, f is the velocity distribution function; m, the atomic mass; and J, the collision integral. Let c_0 be the atom velocity providing the resonance between the radiation and an atomic transition in view of the Doppler shift. Then, only atoms with velocities distributed in the vicinity of c_0 within δc will oscillate along the x axis. For as many atoms as possible to be involved in the effect, one should choose c_0 as close as possible to the most likely thermal velocity of gas atoms (which is 10^3 m/s at room temperature) and appropriately tune the laser frequency. In [2], it was shown that δc is on the order of 10^2 m/s; therefore, one has to provide a temperature of several Kelvins for all the atoms to be set in oscillation. In this case, however, the force amplitude

1063-7842/02/4711-1469$22.00

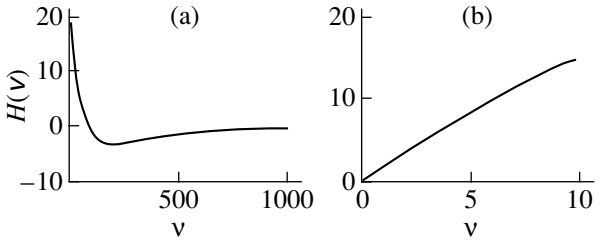


Fig. 1. Dimensionless force amplitude as a function of the atomic velocity at (a) near-room temperatures and (b) low temperatures.

depends on the velocity stronger than at room temperature (see Fig. 1) and the dependence becomes almost linear. Solving Eq. (2) in the general case poses great mathematical difficulties.

Let us take the approximate expression

$$F_1 = H\cos((\Omega_2 - \Omega_1)t + \psi) \times (\theta(u_x - c_0) - \theta(c_0 + \delta c - u_x)) \quad (3)$$

for the force F acting on the atoms, where H is a constant (see Fig. 1a) and $\theta(x)$ is the Heaviside theta function.

Then, an approximate solution to Eq. (2) can be found if the left of (2) is much smaller than the right. This restriction meets the case when the oscillation frequency of atoms far exceeds the frequency of atom collisions.

Under normal atmospheric pressure, the collision frequency of gas atoms is $10^9\ s^{-1}$. Since this value coincides with the frequency of atomic oscillations induced by OMPR at laser field intensities mentioned in the Introduction, the OMPR effect is impossible because of collisions. However, if either the gas is more rarefied or the field has a higher intensity, the situation will change. The oscillation frequency may substantially exceed the collision frequency, and the time between collisions will suffice to set the OMPR regime. Considering time intervals comparable to the transit time, one can notice that between collisions a rapidly oscillating atom can be considered as a cylindrical particle of height h equal to the doubled oscillation amplitude, $h = 2A$, and of diameter d equal to that of the atom. The cylinder axis is aligned with the wave vector of the field for any atomic motion. This allows one to consider the gas as anisotropic.

Let us introduce macroscopic parameters

$$n = \int f\,du; \quad \bar{u} = \frac{1}{n}\int u f\,du;$$
$$\langle\varepsilon\rangle = \frac{1}{n}\int\frac{1}{2}(u-\bar{u})^2 f\,du; \quad \langle\varepsilon\rangle = \frac{3k_B T}{2}. \quad (4)$$

Then, in view of (2), we find

$$\frac{\partial n}{\partial t} = 0; \quad \frac{\partial u_y}{\partial t} = \frac{\partial u_z}{\partial t} = 0;$$
$$\frac{\partial u_x}{\partial t} = n_\delta\frac{H}{m}\cos((\Omega_2 - \Omega_1)t + \psi); \quad (5)$$
$$n\frac{\partial\langle\varepsilon\rangle}{\partial t} = q_\delta\frac{H}{m}\cos((\Omega_2 - \Omega_1)t + \psi),$$

where

$$n_\delta = \int_{-\infty}^{\infty} du_y \int_{-\infty}^{\infty} du_z \int_{c_0}^{c_0+\delta c} f\,du_x;$$
$$q_\delta = \int_{-\infty}^{\infty} du_y \int_{-\infty}^{\infty} du_z \int_{c_0}^{c_0+\delta c} (u_x - \bar{u}_x) f\,du_x. \quad (6)$$

If the left of kinetic equation (2) is small, the first-order approximation for the velocity distribution function is locally Maxwellian:

$$f = n\left(\frac{m}{2\pi k_B T}\right)^{3/2}\exp\left[-\frac{m}{2k_B T}(u-\bar{u})^2\right]. \quad (7)$$

Then, n_δ and q_δ are given by

$$n_\delta = \frac{n}{2}\left\{\Phi\left[\sqrt{\frac{2k_B T}{m}}(c_0 - \bar{u}_x + \delta c)\right] - \Phi\left[\sqrt{\frac{2k_B T}{m}}(c_0 - \bar{u}_x)\right]\right\}, \quad (8)$$

$$q_\delta = n\sqrt{\frac{k_B T}{2\pi m}}\left\{\exp\left[-\frac{m}{2k_B T}c_0^2\right] - \exp\left[-\frac{m}{2k_B T}(c_0 + \delta c)^2\right]\right\}, \quad (9)$$

where Φ is the probability integral.

Thus, the set of equations describing the gas properties near c_0 is closed and can be solved numerically. The gas properties beyond the velocity interval δc, where no forces act on the atoms, are described in the ordinary way. An evident drawback of the theory is the presence of singularities at c_0 and $c_0 + \delta c$ because of the theta function on the left side. At these velocities, the conditions of the approximation are violated and the approximation becomes inapplicable.

TRANSPORT COEFFICIENTS

The transport coefficients of a dynamically anisotropic gas depend on the direction. The well-known phe-

nomenological derivation of the transport coefficients yields the following expressions for the diffusion coefficient D, viscosity coefficient η, and thermal conductivity λ:

$$D = \frac{1}{3}\bar{u}l; \quad \eta = \frac{1}{3}\bar{u}lmn; \quad \lambda = \frac{1}{2}\bar{u}lk_{B}n. \qquad (10)$$

Here, $\bar{u}$ is the mean velocity of gas particles; l, the free path; m, the particle mass; n, the atomic concentration; and k_B, the Boltzmann constant. The free path l is of interest for further consideration. If particles are solid spheres, the free path l is given by

$$l = \frac{1}{\sqrt{2}\pi d^2 n}, \qquad (11)$$

where d is the particle diameter.

Under normal conditions, the mean free path is on the order of 10^{-7} m; therefore, it increases in a rarefied gas. Consider a cylindrical particle of diameter d that is equal to the atomic diameter and of height $h = 2A$ as an approximation of an oscillating atom. The free path of such a particle in the direction of laser beam propagation will alter insignificantly, since the amplitude A is less than 1% of l. However, this is untrue for two directions orthogonal to the beam direction. As follows from the well-known derivation of (11), for these directions the cross-sectional area πd^2 in (11) should be replaced by hd—the quantity characterizing a cylinder. This means that the free path l in these directions is $h/\pi d$ times shorter than in the direction of laser beam propagation. In view of (10), the diffusion coefficient, viscosity coefficient, and thermal conductivity in the directions orthogonal to the beam are also $h/\pi d$ times smaller than those along the beam direction. For the external field intensity mentioned above, $h/\pi d \approx (2 \times 10^{-9}/2\pi \times 10^{-10}) \approx 3$. This seems to be sufficient for this effect to be observed experimentally. In order to simplify the experiment and observe diffusion instead of self-diffusion, a mixture of atoms may be used such that the OMPR conditions are satisfied for atoms of one kind. Their diffusion coefficients are easier to measure than the self-diffusion coefficient. This will make it possible to reveal gas anisotropy, since the diffusion coefficients will differ in various directions.

THERMODYNAMIC PROPERTIES OF A DYNAMICALLY ANISOTROPIC GAS

The thermodynamic properties of a dynamically anisotropic gas will also change. In the case of OMPR, the term "temperature" acquires a new meaning similar to that of the kinetic temperature of an interstellar gas where atoms move in one direction with a high velocity. In our case, vibratory motion in one certain direction corresponding to two extra degrees of freedom takes place instead of rectilinear motion. The energy of the extra degrees of freedom $\langle\varepsilon\rangle_v = (1/2)mA^2(\Omega_2 - \Omega_1)^2$ corresponds to a temperature increment δT defined by the laser power, so that $\langle\varepsilon\rangle_v = (2/2)k_B\delta T$. This implies that the gas medium, which initially was in equilibrium with the environment, serves as a heat source. Thus, the energy of electromagnetic radiation turns into the mechanical energy of oscillating atoms, i.e., into the thermal energy of the gas. The heat flux toward a wall will be specified by particle collisions near the wall. The temperature difference between the heat source and the environment depends on the energy of atomic oscillations, i.e., on the energy injected into the system by the laser. Such an approach also implies the anisotropy of the gas pressure. Thus, in a dynamically anisotropic gas subjected to OMPR conditions, the propagation of sound has tensor properties.

If a gas is a mixture mentioned in the previous section, it is easy to check that a two-temperature medium arises, since only one component of the mixture is under OMPR conditions. In this case, the temperature difference $\delta T = \langle\varepsilon\rangle_v/k_B$ is small.[1]

The heat sources are isotropically distributed over the gas volume, while the thermal conductivity of the medium is anisotropic. This means that the heat fluxes through the lateral and front surfaces of a container will differ. At the same time, the wall temperature T is the same. The expression for the heat flux q from a heated body has the form

$$q = \alpha(T - T_0), \qquad (12)$$

where T_0 is the environmental temperature and α is the heat exchange coefficient usually found experimentally.

It is worth noting that gas anisotropy will result in various values of the heat exchange coefficient on the lateral (α_σ) and front (α_ϕ) walls of the container. The difference in the time it takes for similar bodies touching the lateral and front walls to be heated to the same temperatures could indicate gas anisotropy.

CONCLUSION

Our phenomenological derivation of the expressions for the transfer coefficients is evidently approximate. A rigorous mathematical derivation based on *ab initio* principles, i.e., on the Boltzmann kinetic equation, is a more complex mathematical problem currently being solved. In order to solve it, the Maxwellian atom velocity distribution function should be modified. The temperature effect should be taken into account upon comparing theoretical and experimental results. At room temperature, only some of the atoms will make a contribution to the anisotropy effect. The effect will be enhanced at low temperatures such that all the atoms meet the OMPR conditions. This can be achieved by using atomic traps. It is clear that both theoretical and

[1] The difference δT in the effective temperatures can be increased twice or thrice by using extra laser beams orthogonal to the initial one. They also must meet the OMPR conditions.

experimental results must correlate with the simple estimates presented here. It is the author's opinion that these estimates suffice to attract the attention of researchers to the problem.

ACKNOWLEDGMENTS

The author is grateful to Yu.G. Chesnokov for helpful discussions.

APPENDIX

The designations used are as follows:

$$H(a, \sigma, \nu) = \frac{a(\sigma^2 + 2\Gamma)}{G^5(J-\sigma)K\sqrt{2}}\left[1 + \frac{G-\sigma}{\delta} + \frac{2\sigma(1-\Gamma)}{\delta(\sigma^2+2\Gamma)}\right] \times \sqrt{1 + \Gamma(1+\sigma^2)^2 + \nu^2},$$

$$K = -\nu^2(\sigma^2 + 2\Gamma) - \sigma^2 - 2\Gamma^3(1+\sigma^2)^2 - 2\Gamma\left[1 + \sigma^2(1+\sigma^2) + a^2G^2\frac{(1+\sigma^2)}{(J-\sigma)^2}\right] - \Gamma^2(1+\sigma^2)[4 + \sigma^2(1+\sigma^2)],$$

$$\psi(\nu) = \arccos\frac{1 + \Gamma(1+\sigma^2)}{\sqrt{1 + \Gamma(1+\sigma^2)^2 + \nu^2}} + \pi,$$

$$\varepsilon a = \frac{\alpha_{R,2}}{\alpha_{R,1}}, \quad \varepsilon\sqrt{2} = \frac{\gamma}{\alpha_{R,1}}, \quad \sigma = \delta + \kappa u,$$

$$\delta\sqrt{2} = \frac{\Omega_2 - \Omega_1}{\alpha_{R,1}}, \quad \kappa\sqrt{2} = \frac{k}{\alpha_{R,1}},$$

$$\varepsilon\sqrt{2}\Gamma = \frac{\gamma_{1,2}}{\alpha_{R,1}}, \quad G = \sqrt{\sigma^2 + 2}.$$

Here, $\alpha_{R,1}$ and $\alpha_{R,2}$ are the Rabi frequencies for a strong bichromate, $\gamma_{1,2}$ is the transverse relaxation constant, u is the velocity of an atom, and $\nu = O(1)$ is a dimensionless parameter characterizing the tuning accuracy of parametric resonance [2].

REFERENCES

1. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1993).
2. A. Ya. Kazakov and S. V. Siparov, Opt. Spektrosk. **83**, 961 (1997) [Opt. Spectrosc. **83**, 888 (1997)].
3. S. Siparov, Phys. Rev. A **55**, 3704 (1997).
4. S. Siparov, J. Phys. B **31**, 415 (1998).
5. S. Siparov, J. Phys. B **34**, 2881 (2001).
6. S. Siparov, Space-time & Substance **2**, 44 (2001).
7. J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport in Gases* (North-Holland, Amsterdam, 1972; Mir, Moscow, 1976).

Translated by M. Fofanov

Technical Physics, Vol. 47, No. 11, 2002, pp. 1473–1476. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 129–132.

BRIEF COMMUNICATIONS

Emission from a Sulfur-Hexafluoride Volume Discharge Induced by Attachment Instability

A. K. Shuaibov and I. V. Shevera

Uzhhorod National University, vul. Pidgirna 46, Uzhhorod, 88000 Ukraine

e-mail: ishev@univ.uzhgorod.ua

Received April 9, 2002

Abstract—Characteristics of a low-pressure volume discharge in SF_6 are investigated. The discharge with a spherical anode–plane cathode electrode system was induced by attachment instability. It is found that applying a dc voltage ($U_{ch} \le 1.3$ kV) to the anode in the absence of a confining dielectric wall results in the ignition of a repetitive discharge ($f = 0.1$–120 kHz). The spatial, frequency, and current–voltage characteristics of the volume discharge; plasma emission in the spectral range 200–700 nm; and the waveforms of the discharge voltage, current, and plasma emission are investigated. It is shown that the plasma under study exists in the form of a domain (autosoliton) and that the volume discharge is self-induced because, during attachment instability, the plasma itself acts as a switch enabling repetitive pulsed operation. The results obtained can be applied to studying the physics and technology of electric-discharge chemical HF (DF) lasers based on nonchain reactions and to developing high-aperture low-pressure repetitive Ar-, Kr-, and Xe-fluoride lamps with low-corrosive working media (low-pressure mixtures of Ar, Kr, and Xe with SF_6).

INTRODUCTION

Studies of the characteristics of volume discharges in pure sulfur hexafluoride and its mixtures with some gases (SF_6/H_2, SF_6/C_2H_6, etc.), as well as physical processes in such discharges, have recently attracted great interest. This interest stems from the high output parameters of electric-discharge HF (DF) lasers that are based on nonchain reactions and in which the main component of the working medium is sulfur hexafluoride [1–3]. Recent developments in these lasers initiated a series of studies of the characteristics of repetitive volume discharges in mixtures containing sulfur hexafluoride [4, 5]. Repetitive discharges in the mixtures of SF_6 with noble gases, gaseous hydrocarbons, and hydrogen are ignited by using pulsed high-voltage generators with either capacitive or inductive energy storage banks. In these generators, spark gaps or hydrogen thyratrons are usually used as switches. These switches are characterized by limited service life and large dimensions and, sometimes (e.g., spark gaps), have an autonomous system for the working gas circulation. In view of these disadvantages, a search for new methods for igniting repetitive volume discharges in sulfur hexafluoride and its mixtures with other molecular gases is of great interest now. SF_6 can also be used in the working mixtures of electric-discharge excimer lamps [6], which, in contrast to excimer lamps based on molecular fluorine, have a longer gas fill lifetime. The working mixtures of such lamps are low-corrosive, which allows operation in the gas circulation regime, simplifies the technology of manufacturing the corresponding radiation sources, and extends their range of applicability.

Preliminary studies of volume discharges in chlorine and CFC-12 showed that applying a dc voltage to the anode resulted in the ignition of a repetitive discharge ($f = 0.1$–50 kHz) in the interelectrode gap. As the electron affinity of the electronegative molecule increased, the amplitude of the discharge current pulse increased from 0.1–0.5 to 3–5 A at a current pulse duration of ≤50 ns. Hence, in order to increase the current amplitude and control the other parameters of a self-induced discharge, it is of interest to study even more electronegative SF_6-based working mixtures. Previously, the attachment instability was observed in the generation zone of a multielectrode corona discharge in the mixtures of noble gases with fluorine at pressures of $P = 100$–300 kPa [7].

In this paper, we study the characteristics of a sulfur-hexafluoride repetitive volume discharge induced by attachment instability in a low-density plasma.

EXPERIMENTAL TECHNIQUE

The experimental facility consisted of an electrode system, a discharge chamber, a dc power supply, and a system for recording the electrical and optical characteristics of a volume discharge. The electrode system consisted of a massive 7-cm-diameter hemispherical duralumin anode and 9-cm-diameter plane cathode (Fig. 1). The radius of curvature of the anode working surface was 3 cm. The interelectrode distance was 2.8 cm. A dc positive voltage ($U_{ch} \le 1.3$ kV) was applied to the anode through a 20-kΩ ballast resistor.

1063-7842/02/4711-1473$22.00

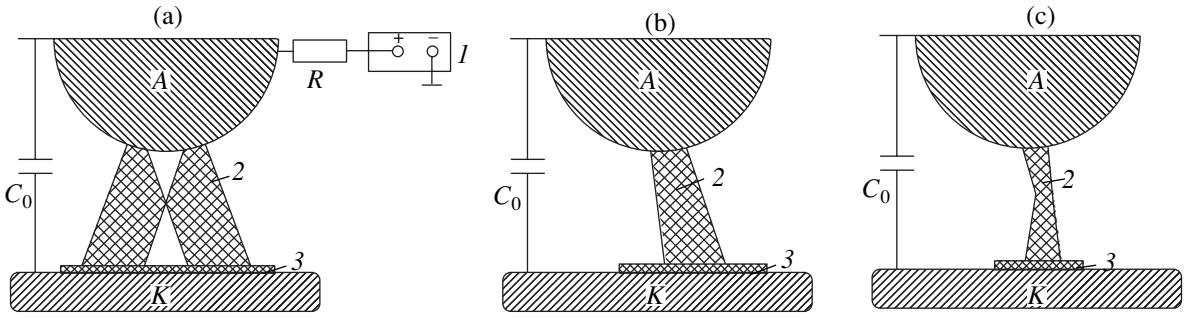


Fig. 1. Electrode system and the cross-sections of a volume discharge plasma for $P(SF_6)$ = (a) 40, (b) 200, and (c) 600 Pa: (*A*) anode, (*K*) cathode, (*R*) ballast resistance, (*1*) high-voltage dc power supply, (*2*) positive column, and (*3*) negative cathode glow. The capacitance of the pulsed capacitor is C_0 = 200 pF.

The pressure of technical-grade sulfur hexafluoride was varied in the range 400–600 Pa. The average discharge current was in the range 2–50 mA. The electrode system was mounted on a glass textolite flange and was placed in a 10-l-volume discharge chamber. The chamber with the electrode system was pumped down to a residual air pressure of 3–5 Pa. The chamber was hermetically connected via an LiF window to a vacuum monochromator with a spectral resolution of 0.7 nm. The monochromator and an FEU-142 photomultiplier (also with an LiF window) were relatively calibrated in the spectral range 130–350 nm. We also recorded the total emission from the plasma with the help of a FOTON photomultiplier and a broadband S1-99 oscilloscope. The FOTON photomultiplier was placed opposite the entrance slit of the vacuum monochromator so that the plasma under study was located halfway between them. The pulsed plasma emission in the spectral range 200–700 nm was recorded with a time resolution of 7–10 ns. The repetition rate of the radiation pulses was monitored using an S1-99 oscilloscope and a Ch 3-57 frequency meter. The ac component of the voltage across the anode–cathode gap was recorded using a low-inductive capacitive divider, and a low-inductive current shunt (r = 1–5 Ω) was used to monitor the current pulses. To investigate how the capacitance shunting the discharge gap affects the amplitude, duration, and repetition rate of the discharge pulses, a set of low-inductive capacitors with the total capacitance C_1 = 680 pF were introduced in the circuit in addition to the self-capacitance of the electrode system, the stray wiring capacitance ($C_0' \le 20$ pF), and the input capacitance of the voltage divider (C_0 = 200 pF).

SPATIAL AND ELECTRICAL CHARACTERISTICS

The steady-state distribution of the plasma emission intensity in the interelectrode gap was monitored by taking photos of the discharge. Figure 1 shows the pictures of the sulfur hexafluoride plasma emission obtained by processing the photos with a scanner and PC. At a low sulfur hexafluoride pressure ($P \le 40$ Pa), the discharge occupied most of the cross-sectional area of the interelectrode gap, except the central region, where the E/N parameter was maximum (Fig. 1a). The diameter of the negative cathode glow plasma was 1.5–2.0 times larger than the maximum diameter of the positive column plasma. As $P(SF_6)$ increased to 200 Pa, the volume of the positive column decreased about two times. As the sulfur-hexafluoride pressure increased further, the volume of the positive column decreased and its shape changed significantly (Fig. 1c).

The current–voltage characteristics of the volume discharge are shown in Fig. 2. The steady-state current–voltage characteristic of the discharge (Fig. 2a) is similar to that of a dc longitudinal subnormal glow discharge [8, 9]. As the current I_{ch} increases from 2 to 50 mA, the average voltage drop across the gap decreases from 1.3 to 1.05 kV. For the same experimental conditions, the dependence between the instant values of the voltage and current (the dynamic current–voltage characteristic) is completely different (Fig. 2b). In the initial stage, when the domain plasma just begins to form, the discharge current increases to 29 A with increasing discharge voltage. Then, the current decreases, which results in a negative slope of the current–voltage characteristic. Such a behavior of the dynamic current–voltage characteristic is typical of a plasma with developed attachment instability [10].

OPTICAL CHARACTERISTICS

Study of the emission spectra from a sulfur-hexafluoride volume discharge showed that UV emission consists mainly of the impurity molecular bands: the second positive system; the third Kaplan system of $N_2(A^3\Sigma_u^+–E)$; and the forth positive system of molecular nitrogen $N_2(B^3\Pi_g–D^3\Sigma_u^+)$, OH($A$–$X$), and CN($B$–$X$). This is related to the presence of small amounts of air and water vapor (at a level of 1–2 Pa) in SF_6 and the residual air pressure in the discharge chamber ($P_1 \le$ 5 Pa). Because of the strong attachment of slow elec-

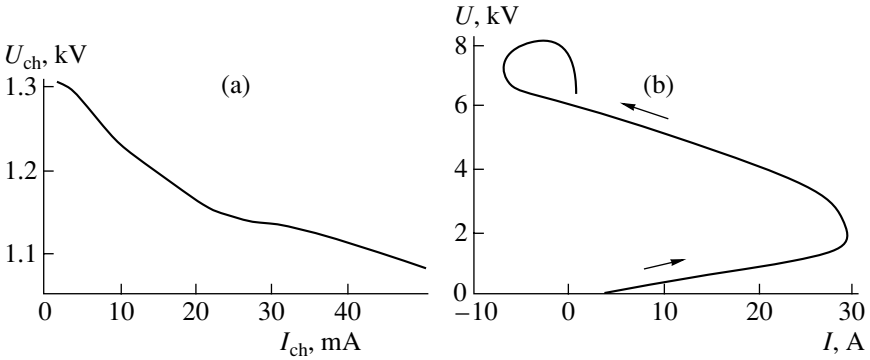


Fig. 2. Steady-state and dynamic current–voltage characteristics of the volume discharge for $P(SF_6)$ = 400 Pa and C_0 = 220 pF.

trons to SF_6 and SF_5 molecules (the SF_5^- radicals are efficiently produced in the dissociation of SF_6 molecules in the discharge), the high-lying states of nitrogen are excited most efficiently, which results in the prevalence of the $N_2(C–B)$ 337.1- and 357-nm bands (belonging to the second positive system of N_2) in the plasma emission.

The waveforms of the discharge voltage and current and the total emission intensity from a discharge in a sulfur hexafluoride–based mixture are shown in Fig. 3. It follows from Fig. 3 that the discharge under study is repetitive. At a dc voltage at the anode of 1.0–1.3 kV, the pulsed voltage amplitude attains 4 kV. The voltage pulse has a short ($\tau \le$ 50–100 ns) leading edge and an extended (up to 1.3 μs) trailing edge. The current pulse with a full width of 75–150 ns is formed during the leading edge of the voltage pulse. The duration of the emission pulse attains 400–500 ns. The emission intensity reaches its maximum in the near afterglow of the current pulse; this fact can be related to the recombination and stepwise mechanisms for the formation of the emitting plasma particles. The domain propagation velocity in the anode–cathode gap can be determined from the duration of the leading edge of the current pulse and the interelectrode distance [11]. Under conditions corresponding to Fig. 3, this velocity is 4×10^7 cm/s.

Let us consider the influence of the capacitance shunting the interelectrode gap on the pulsed characteristics of the volume discharge. Figure 4 presents the waveforms of the current and the total plasma emission intensity for different shunting capacitances at a constant value of $P(SF_6)$. The increase in the capacitance results in a significant increase in the current pulse amplitude (by a factor of greater than 3), twofold increase in the pulse duration, and decrease in the domain propagation velocity in the interelectrode gap from 6×10^7 to 3×10^7 cm/s. As the shunting capacitance increases, the emission pulse shifts in time from the leading edge of the current pulse toward its maximum (Fig. 4). Figures 3 and 4 present the most representative waveforms with the maximum pulse amplitudes; in the experiments, we recorded a rather wide set of U, I, and J_F waveforms (up to 10–15 waveforms synchronized in time), whose amplitudes obeyed certain dispersion relations, as in the case with the other types of autosolitons [12].

Figure 5 shows the typical dependences of the pulse repetition rate on the average discharge current for different values of $P(SF_6)$ and different shunting capacitances. Generally, the pulse repetition rate increases with the average discharge current. The pulse amplitude increases linearly with the average discharge current. The increase in $P(SF_6)$ at the minimum capacitance C_0^{min} results in a significant increase in the pulse repetition rate and the disappearance of the descending segment in the $f(I_{ch})$ curve (Fig. 5, curve *1*). The increase in C_0 from 20 to 900 pF does not change the behavior of the $f(I_{ch})$ curve; however, the current pulse repetition rate decreases by more than one order of magnitude. In this case, the shunting capacitance acts as a capacitive energy storage bank, whereas the vol-

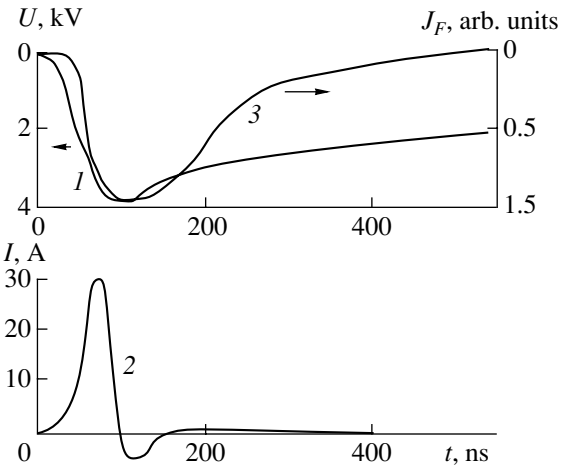


Fig. 3. Waveforms of (*1*) the discharge current, (*2*) voltage, and (*3*) total plasma emission intensity at $P(SF_6)$ = 200 Pa and C_0 = 220 pF.

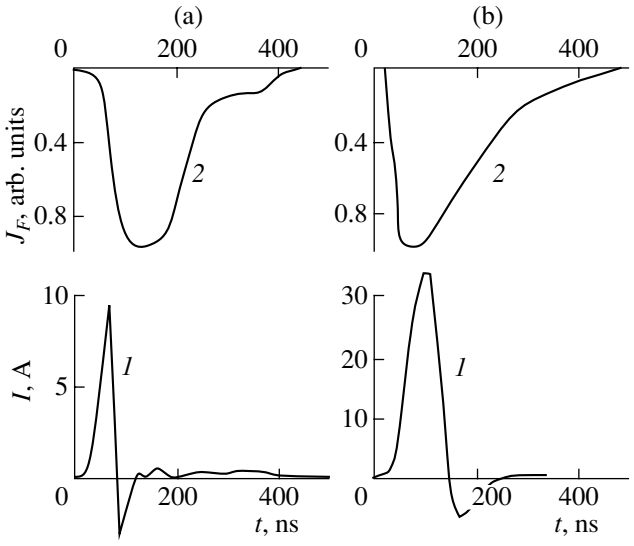


Fig. 4. Waveforms of (*1*) the current of and (*2*) total plasma emission intensity from a volume discharge for $P(SF_6)$ = 200 Pa and $C_0 \le$ (a) 20 and (b) 220 pF.

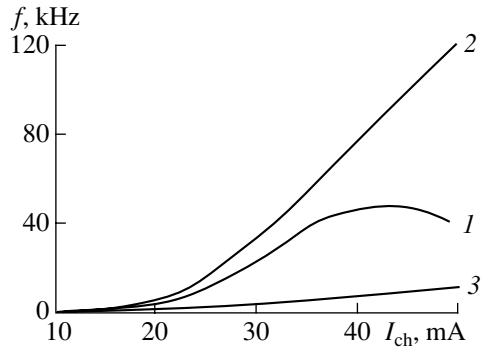


Fig. 5. Repetition rate of the volume discharge pulses vs. average discharge current for (*1*) $P(SF_6)$ = 200 Pa and $C_0 \le$ 20 pF, (*2*) $P(SF_6)$ = 600 Pa and $C_0 \le 20$ pF, and (*3*) $P(SF_6)$ = 200 Pa and C_0 = 900 pF.

ume discharge plasma itself serves as a pulsed voltage switch.

CONCLUSION

It has been shown that the plasma of a self-induced repetitive volume discharge in sulfur hexafluoride represents a domain (a set of autosolitons), whose formation is related to the onset of attachment instability. This is indicated by the presence of a segment with a negative slope in the dynamic current–voltage characteristic of the volume discharge. The plasma emits in the UV–VUV spectral range due to vibronic transitions of impurity molecules (N_2, OH, etc.). Shunting the discharge gap with an additional capacitance C_0 = 20–900 pF allows one to control the amplitude and duration of the discharge-current and emission pulses, as well as the domain propagation velocity in the interelectrode gap, over wide limits. Volume discharges in sulfur hexafluoride and its mixtures with noble gases and gaseous hydrocarbons can be used to develop high-repetition-rate switches, low-pressure excimer lamps, and active elements for high-repetition-rate HF lasers and amplifiers.

REFERENCES

1. V. P. Borisov, V. V. Burtsev, S. D. Velikanov, *et al.*, Kvantovaya Élektron. (Moscow) **30**, 225 (2000).
2. B. V. Lazhintsev, V. A. Nor-Arevyan, and V. D. Selemir, Kvantovaya Élektron. (Moscow) **30**, 7 (2000).
3. V. F. Tarasenko, V. M. Orlovskiĭ, and A. N. Panchenko, Kvantovaya Élektron. (Moscow) **31**, 1035 (2001).
4. Yu. I. Bychkov, S. L. Gorchakov, and A. G. Yastremskiĭ, Kvantovaya Élektron. (Moscow) **30**, 733 (2000).
5. V. V. Apollonov, A. A. Belevtsev, S. Yu. Kazantsev, *et al.*, Kvantovaya Élektron. (Moscow) **30**, 207 (2000).
6. A. K. Shuaibov and I. V. Shevera, Zh. Prikl. Spektrosk. **69**, 113 (2002).
7. A. K. Shuaibov, L. L. Shimon, and A. J. Minya, J. Phys. Stud. **4**, 291 (2000).
8. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).
9. V. P. Aleksandrov and A. P. Napartovich, Usp. Fiz. Nauk **163** (3), 1 (1993) [Phys. Usp. **36**, 107 (1993)].
10. V. S. Golubev and S. V. Pashkin, *Increased Pressure Glow Discharge* (Nauka, Moscow, 1990).
11. V. L. Aleksandrov, I. V. Kochetov, D. I. Mazalov, *et al.*, Fiz. Plazmy **18**, 1468 (1992) [Sov. J. Plasma Phys. **18**, 758 (1992)].
12. B. S. Kerner and V. V. Osipov, *Autosolitons: a New Approach to Problems of Self-Organization and Turbulence* (Nauka, Moscow, 1991; Kluwer, Dordrecht, 1994).

Translated by N. Ustinovskiĭ

Technical Physics, Vol. 47, No. 11, 2002, pp. 1477–1480. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 133–136.

BRIEF COMMUNICATIONS

Reversed Doppler Effect under Reflection from a Shock Electromagnetic Wave

A. M. Belyantsev and A. B. Kozyrev

Institute for Physics of Microstructures, Russian Academy of Sciences, Nizhni Novgorod, 603950 Russia

e-mail: kozyrev@ipm.sci-nnov.ru

Received April 18, 2002

Abstract—The possibility of observing the reversed Doppler effect in an electrodynamic system of coupled transmission lines with different dispersion types is demonstrated. © *2002 MAIK "Nauka/Interperiodica".*

Recently, the creation of artificial media that offer negative effective permittivity ε and permeability μ simultaneously has been reported [1]. Such artificial media are of interest primarily because they allow for the observation of reversed effects, in particular, the reversed Doppler effect in the microwave range when nonlinear "inclusions" are introduced into the medium in order to create the moving boundary of its parameters.

Below, we propose a nonlinear electrodynamic system where the reversed Doppler effect can be realized relatively simply for positive ε and μ of the filling medium. It has been shown [3] that a shock electromagnetic wave (SEMW) that travels in an electrodynamic system of coupled transmission lines with different types of dispersions can efficiently excite an rf backward wave. In this case, a video pulse propagating through the electrodynamic system that is filled with a nonlinear hysteretic medium remaining in the saturated state for a long time transforms into a radio pulse with a carrier frequency meeting the synchronism condition (i.e., when the wave phase velocity $v_p(\omega)$ equals the SEMW velocity v_s). For the backward waves, $v_g(\omega)v_p(\omega) < 0$ ($v_g(\omega)$ is the group velocity) and the SEMW-generated radio pulse can be turned around to pursue the SEMW front, for example, reflecting from the input of the electrodynamic system. If $|v_g(\omega)| > v_p(\omega) = v_s$, the backward wave (radio pulse) will catch up with the SEMW front and reflect from the receding interface between the linear and nonlinear media. The radio pulse reflected from the receding SEMW front will have a higher carrier frequency than the incident radio pulse (reversed Doppler effect).

For this scenario to occur, the electrodynamic system must allow for both the SEMW formation and the existence, in the saturated nonlinearity state, of backward waves with $v_p(\omega) < |v_g(w)|$ in a certain frequency range. Such an electrodynamic system can be implemented by coupling two transmission lines one of which is linear and has a anomalous dispersion (backward waves or a wave with a dominating zeroth spatial harmonic) and the other is nonlinear with a normal dispersion (forward wave) in the saturated state. If a ferrite is used as the nonlinear medium, the optimal normal-dispersion transmission line is, in our opinion, a coaxial (or a quasi-coaxial) transmission line in which a steep-front (several tenths of a nanosecond wide) SEMW can be formed. As the anomalous dispersion line, the system can use a shielded interdigital slow-wave structure [4, 5] where, with the parameters chosen appropriately, the fundamental spatial harmonic is backward and dominates and where the magnitude of the phase velocity of the fundamental wave can be less than that of the group velocity over a certain frequency range (Fig. 1). Such slow-wave structures are used mostly in wideband backward-wave tubes. The dispersion relation for the interdigital structure derived by the multiwire line method is written as

$$\tan^2\frac{kl\sqrt{\varepsilon}}{2} = \frac{\gamma_0 + 4\gamma_1\cos^2\varphi/2}{\gamma_0 + 4\gamma_1\sin^2\varphi/2}, \qquad (1)$$

where

$$\gamma_0 = 4\pi\varepsilon(d-D)(1/a - 1/b), \quad \gamma_1 = 4\pi\varepsilon p/D,$$

φ is the phase shift over the structure period d, l is the finger length, $k = \omega/c$, and ε is the permittivity.

These transmission lines can be coupled through a narrow longitudinal slot carved in the shields as shown in Fig. 1.

It is most convenient to approximate the dispersion characteristics of the interdigital system by a *CL* multisection circuit with inductive cross links. The normal-dispersion transmission line is convenient to approximate by a usual *LC* multisection circuit. When coupled through the narrow slot at the shield center, the transmission lines interact via the electric field alone. In the equivalent circuit of the electrodynamic system (Fig. 2), this coupling is included by the coupling capacitance C_*, which can easily be controlled by the

1063-7842/02/4711-1477$22.00

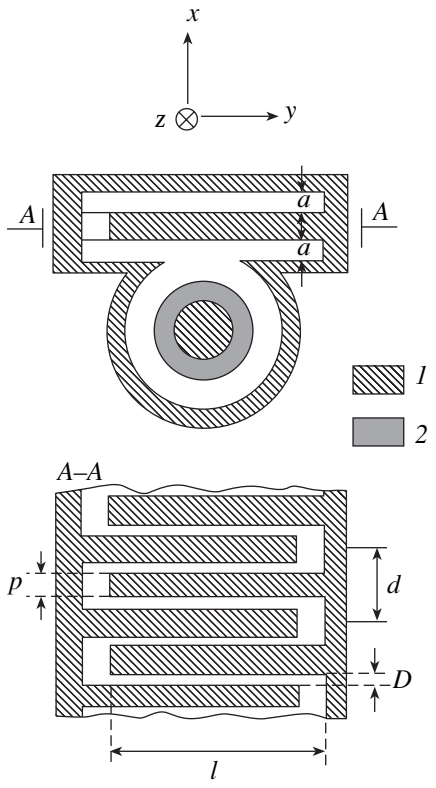


Fig. 1. System of coupled coaxial transmission line and interdigital slow-wave structure: (*1*) metal and (*2*) ferrite.

parameters of the system. It is thus convenient to simulate the wave evolution and propagation in this electrodynamic system in terms of the equivalent circuit shown in Fig. 2 [3]. Figure 3 plots the dispersion characteristics of the coupled transmission lines calculated using the equivalent circuit of Fig. 2 (its parameters are given in the figure caption). As follows from Fig. 3, the dispersion characteristics have two branches. The synchronism between the SEMW and the backward wave is possible on both the low-frequency and the high-frequency branches (Fig. 3) at the frequencies $\omega_{\text{synch }1}$ and $\omega_{\text{synch }2}$, respectively. However, the parameters of the coupled transmission lines and the SEMW amplitude are chosen such that the condition $v_p(\omega_{\text{synch }1}) = v_s < |v_g(\omega_{\text{synch }1})|$, under which the SEMW-generated radio pulse with the carrier frequency $\omega_{\text{synch }1}$ strikes the SEMW front after being reflected from the input of the electrodynamic system, is only satisfied for the low-frequency synchronous wave. Clearly, the frequencies ω_{sec} of the secondary waves generated at the interface between the linear and nonlinear media, which moves with the velocity $\mathbf{v} = \mathbf{v}_x$, are related to the incident wave frequency $\omega_{\text{in}} = \omega_{\text{synch }1}$ by the well-known Doppler formula [2]

$$\frac{\omega_{\text{sec}}}{\omega_{\text{in}}} = \frac{1 - (\mathbf{v}\cdot\mathbf{v}_{\text{in}})/v_{\text{in}}^2}{1 - (\mathbf{v}\cdot\mathbf{v}_{\text{sec}})/v_{\text{sec}}^2}, \qquad (2)$$

where $\mathbf{v}_{\text{in}}$ and $\mathbf{v}_{\text{sec}}$ are the phase velocities of the incident and secondary waves, respectively.

Using relationship (2), one can graphically calculate the frequencies of all the secondary waves. They are given by points of intersection with the straight line of slope v_s that passes through the point on the dispersion curve for the synchronous wave reflected from the input (Fig. 3). In our case, all the secondary waves are forward waves and satisfy the radiation condition, two of them having frequencies higher than that of the wave incident on the SEMW front (ω_1', $\omega_2' > \omega_{\text{synch }1}$).

In terms of the equivalent circuit of Fig. 2, wave processes in the electrodynamic system are described by Kirchhoff's usual difference equations [6]. If the magnetization of the ferrite material in the coaxial line is reversed rapidly and incoherently, the magnetization variation is described by the equation [6]

$$\frac{dM_{\text{av}n}}{dt} = -\frac{\alpha\gamma}{(1+\alpha^2)M}(M^2 - M_{\text{av}n}^2)H_n(I_n^{(1)}), \qquad (3)$$

where $I_n^{(1)}$, $M_{\text{av}n}$, and $H(I_n^{(1)})$ are the current in the ferrite, the averaged magnetization vector (parallel to the

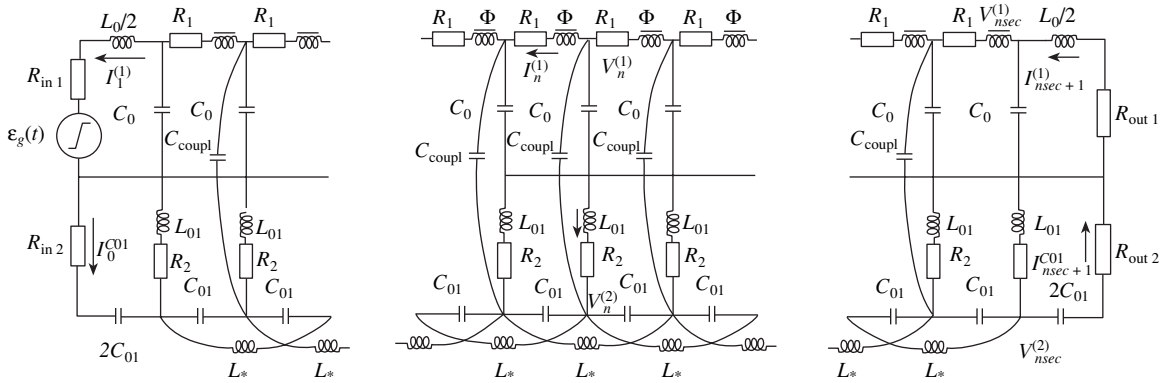


Fig. 2. Equivalent circuit of the electrodynamic system in Fig. 1.

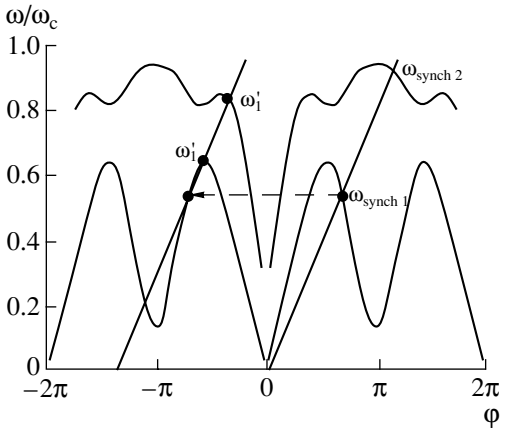


Fig. 3. Relative frequency ω/ω_c ($\omega_c = 2/(L_0C_0)^{1/2}$) versus relative wave number $\varphi = k/d$ (d is the structure period) for the normal-dispersion transmission line coupled to the interdigital slow-wave structure at $C_{01}/C_0 = 0.1$, $C_{coupl}/C_0 = 0.16$, $L_*/L_0 = 5$, and $L_{01}/L_0 = 22.5$.

magnetic field), and the magnetic field, respectively, in the nth cell of the nonlinear transmission line; M is the saturation magnetization of the ferrite; γ is the absolute value of the gyromagnetic ratio; and α is the dissipation factor.

The system of Kirchhoff's equations [6] and Eqs. (3) were solved numerically by the Runge–Kutta fourth-order method. As in [6], we assumed that, at the initial time moment ($t = 0$), the currents and voltages in all the cells are zero, a certain initial magnetization of the ferrite is specified, and a semi-infinite voltage step is applied to the input of the normal-dispersion transmission line.

Figure 4a plots the typical time behavior of the relative voltage at nodes of the normal-dispersion (upper plot) and anomalous-dispersion (lower plot) transmission lines. The results were obtained numerically with the input of the anomalous-dispersion line short-circuited for the high frequency. As follows from this numerical experiment, the backward wave, striking the SEMW front, partially "penetrates" the traveling interface between the linear and nonlinear media (the signal that passed through the SEMW front is seen more distinctly in the voltage waveform for the anomalous-dispersion line). The wave that penetrates the SEMW front decays rapidly. Figure 4b plots the spectrum of oscillations in the anomalous-dispersion line. Peaks related to the lower-frequency and higher-frequency synchronous waves ($\omega_{synch\,1}$ and $\omega_{synch\,2}$), as well as peaks related to the secondary waves with the frequencies ω'_1 and ω'_2, are clearly observed in the spectrum. Note that the centers of the secondary-wave spectral lines are in good agreement with those calculated from Eq. (2). The amplitudes of the secondary waves with ω'_1 and ω'_2 amount, respectively, to 20 and 10% of the backward wave amplitude incident on the SEMW front. Since in this electrodynamic system, reflections from the input, the SEMW, and forward secondary waves travel toward the output of the system, the output radio pulse will contain three carrier frequencies. The two higher frequencies appear due to the reversed Doppler effect.

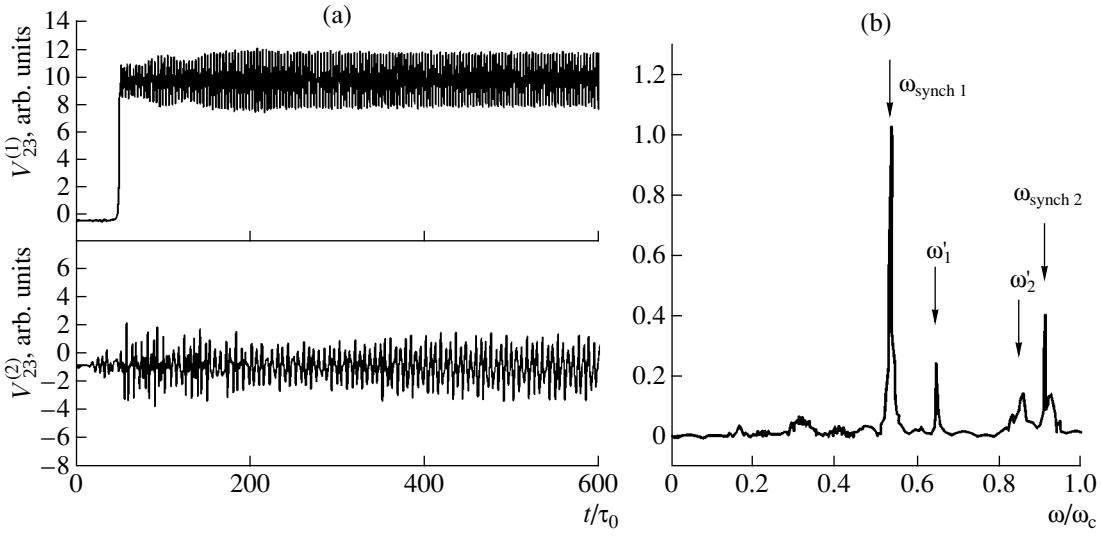


Fig. 4. (a) Voltage waveforms for the 23th cell of the negative- and anomalous-dispersion lines, $V_{23}^{(1)}/V_s$ and $V_{23}^{(2)}/V_s$, normalized to the SEMW amplitude and (b) voltage oscillation spectrum for the anomalous-dispersion line. The parameters of the nonlinear medium and the amplitude of the input video pulse are such that the SEMW velocity is $v_s/v_0 = 0.506$ ($v_0 = d/(L_0C_0)^{1/2}$), which corresponds to the slope of the straight line in Fig. 2 (here, $\tau_0 = (L_0C_0)^{1/2}$).

Our simulations have thus shown that the reversed Doppler effect can be observed at certain parameters of an interdigital slow-wave structure coupled to a quasi-coaxial transmission line. This effect can be used to directly convert a single-carrier radio pulse into a multicarrier radio pulse. The spectrum of this pulse can easily be controlled electronically (by varying the input video pulse amplitude or the bias magnetization of the ferrite material that fills the transmission line).

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 99-02-18046).

REFERENCES

1. D. R. Smith, W. J. Padilla, D. C. Vier, *et al.*, Phys. Rev. Lett. **84**, 4184 (2000).
2. L. A. Ostrovskiĭ and N. S. Stepanov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **14**, 488 (1971).
3. A. M. Belyantsev and A. B. Kozyrev, Zh. Tekh. Fiz. **71** (7), 79 (2001) [Tech. Phys. **46**, 864 (2001)].
4. R. A. Silin and V. P. Sazonov, *Slow-Down Systems* (Sov. Radio, Moscow, 1966).
5. *Cross-Field Microwaver Devices* (Academic, New York, 1965), Vols. 1, 2.
6. A. M. Belyantsev and A. B. Kozyrev, Zh. Tekh. Fiz. **68** (1), 89 (1998) [Tech. Phys. **43**, 80 (1998)].

Translated by A. Khzmalyan

Technical Physics, Vol. 47, No. 11, 2002, pp. 1481–1483. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 137–139.

BRIEF COMMUNICATIONS

Effect of Adsorption of Alkali Metal Atoms on the Work Function of Rutile

S. Yu. Davydov* and I. V. Noskov**

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

*** St. Petersburg State Electrotechnical University (LETI), ul. Prof. Popova 5, St. Petersburg, 197376 Russia*

Received May 14, 2002

Abstract—Calculation of the change of work function $\Delta\phi$ at the surface of TiO_2 due to the adsorption of Na, Ka, and Cs atoms has been carried out in the framework of a model taking into account the repulsion of adatoms and metallization of the adsorbed layer. The calculation results are in good agreement with experiment. © *2002 MAIK "Nauka/Interperiodica".*

Studies of adsorption at the surface of metal oxides has begun fairly recently [1]. At present, the best studied are submonolayers of metal oxides on the surface of rutile. The M–TiO_2 system therefore appears to be an obvious choice for building a theoretical model of adsorption on oxides, in which an alkali metal can be conveniently chosen as the adsorbate metal M, as it is the simplest adsorbate from the theoretical standpoint and, at the same time, its deposition onto a nonmetallic substrate shows all the characteristic features of electronic structure formation in an adsorption system [2]. In [3], we proposed a model giving a fair description of decrease $\Delta\phi$ of work function ϕ_0 at the (110) face of rutile due to the adsorption of cesium. Taken into account were both the dipole–dipole repulsion of adatoms and the broadening of their quasi-levels at large coverages $\Theta = N_a/N_{ML}$ (N_a is the concentration of adatoms on the substrate; N_{ML} is the concentration of adatoms in the monolayer) due to the tunneling of electrons between adjacent adatoms. Application of the model [3] to experimental data on the adsorption of K atoms on the (100) face [4] and Na atoms on the (110) face [5] of TiO_2 revealed that the model needed modification. The modified version is presented in this study.

Consider an atom with one electron in its outer *s*-shell. Following adsorption, the electron can transfer from the adatom to a Ti^{4+} ion. As a consequence, the filling number *n* of the *s*-orbital becomes different from unity and the adatom acquires a charge of $Z = 1 - n$. The dependence of *Z* on the coverage Θ can be calculated using the formula [6–10]:

$$Z(\Theta) = \frac{2}{\pi}\arctan\frac{\Omega - \xi\theta^{3/2}Z(\Theta)}{\Gamma(\Theta)}, \qquad (1)$$

$$\xi = 2e^2\lambda^2 N_{ML}^{3/2}A, \quad \Gamma = \Gamma_0(1 + \gamma\Theta).$$

Here, Ω is the energy of the adatom quasi-level relative to the Fermi level in the substrate; ξ is the dipole–dipole repulsion constant of the adatoms; 2λ is the distance between charges in a surface dipole; $A \sim 10$ is a dimensionless coefficient weakly dependent on the adatom lattice geometry; Γ_0 is the halfwidth of the level of an isolated adatom; and γ is a dimensionless coefficient taking into account the band broadening [3].

Change of work function $\Delta\phi$ is determined in the following way [6, 7]:

$$\Delta\phi(\Theta) = -\Phi\Theta Z, \quad \Phi = 4\pi e^2 N_{ML}\lambda. \qquad (2)$$

In order to determine the model parameters, we will use both theoretical estimates and experimental data. The value of λ will be determined in a self-consistent manner proceeding, first, from the initial slope of the experimental curve of work function versus coverage,

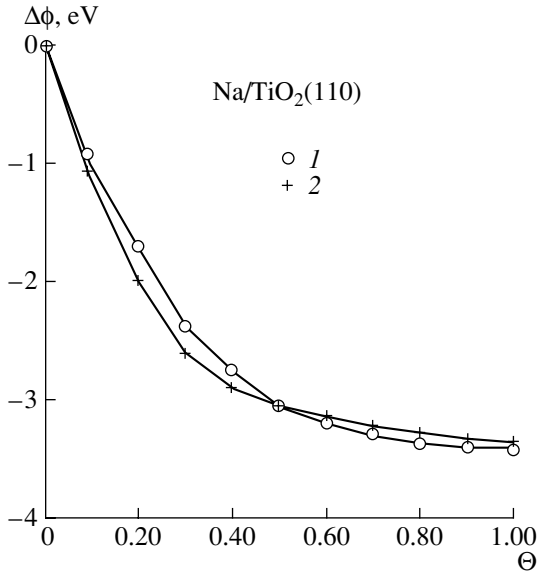


Fig. 1. Variation of the work function $\Delta\phi$ in the Na/TiO_2 (100) system with coverage Θ: (*1*) calculation results, (*2*) experimental data of [5].

1063-7842/02/4711-1481$22.00

Model parameters

Parameter	Na	K	Cs
N_{ML}, 10^{14} cm^{-2}	7.0	5.5	6.0
Γ_0, eV	1	1	1
Ω, eV	4.16	4.13	3.58
λ, Å	1.0	1.4	1.7
Φ, eV	12.7	15.25	18.0
ξ, eV	5.3	8.3	11.7
Z_0	0.85	0.92	0.83
Z_{ML}	0.16	0.24	0.18
γ	5.0	4.6	3.9

i.e., $(d\Delta\phi/d\Theta)_{\Theta \to 0}$, and, second, from the adatom charge Z. To determine the latter we assume

$$\Omega = \phi - I + \Delta, \quad \Delta = e^2/4\lambda, \tag{3}$$

where I is the ionization energy of the adsorbed atom [11], ϕ is the work function of the rutile surface, and Δ is the Coulomb shift of the adatom quasi-level due to interaction of its electron with those of the substrate [11].

It is by taking into account the shift Δ that modification of the model is performed. Then, using the experimental value of the decrease in the work function for monolayer coverage $\Delta\phi_{ML}$, we determine Z_{ML} from (2); in so doing, parameter γ

$$\gamma = \frac{\Omega - \xi Z_{ML}}{\Gamma_0 \tan\left(\frac{\pi}{2} Z_{ML}\right)} - 1. \tag{4}$$

Thus, the initial slope of the $\Delta\phi(\Theta)$ curve, and then its final slope $\Delta\phi_{ML}$, is fitted.

The parameter values calculated in this way are given in the table. In addition to sodium and potassium, parameters for cesium are also given, calculated using the experimental data of [12], which differ markedly from the values given in [13] that we used in [3]. The calculation results for $\Delta\phi(\Theta)$, shown in Figs. 1–3, and the experimental data agree remarkably well.

Let us analyze how the model parameters vary across the Na–K–Cs series and their numerical values. First of all, good correlation is notable between the values of parameter λ given in the table and ionic radii r_i in the Na–K–Cs series, which equal, respectively, 0.92, 1.33, and 1.86 Å [10]. Namely, the increase in the dipole charge separation 2λ in the Na–K–Cs series is accompanied by a corresponding increase in the energy parameters $\Phi(\sim\lambda)$ and $\xi(\sim\lambda^2)$ in the series. It should be emphasized that the values of these parameters appear quite reasonable (see, e.g., [14–16].

Parameter γ characterizes the broadening of the quasi-levels into the band and therefore should correlate with integrals of nearest neighbor transitions in the monolayer of alkali metal atoms. These integrals can be equated to matrix elements $V_{ss\sigma}$ of the Harrison model of binding orbitals [17], which are proportional to r_a^{-2} (r_a being the atomic radius of the alkali metal). For the Na–K–Cs series, the respective ratios of the $V_{ss\sigma}$ integrals are 0.40 : 0.18 : 0.14, corresponding to decreasing γ values transferring from sodium to cesium (see table).

The parameter that is most difficult to evaluate is Γ. In the simplest case of an atom adsorbed on the metal surface, Γ has a clear physical meaning: it is the half-

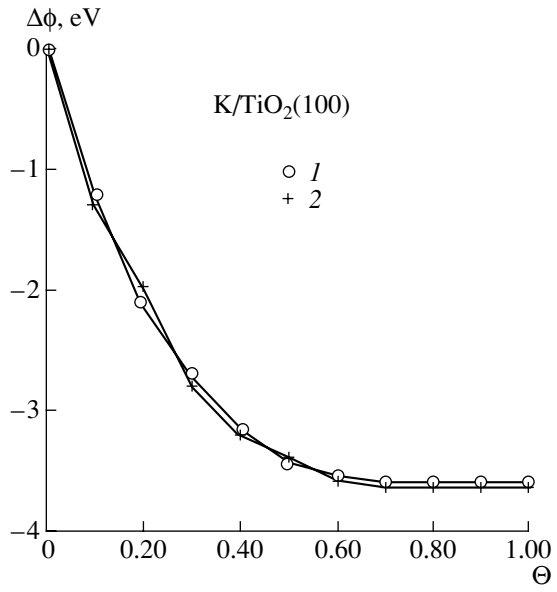


Fig. 2. Same as in Fig. 1 for the K/TiO_2 (100) system, using the experimental data of [4].

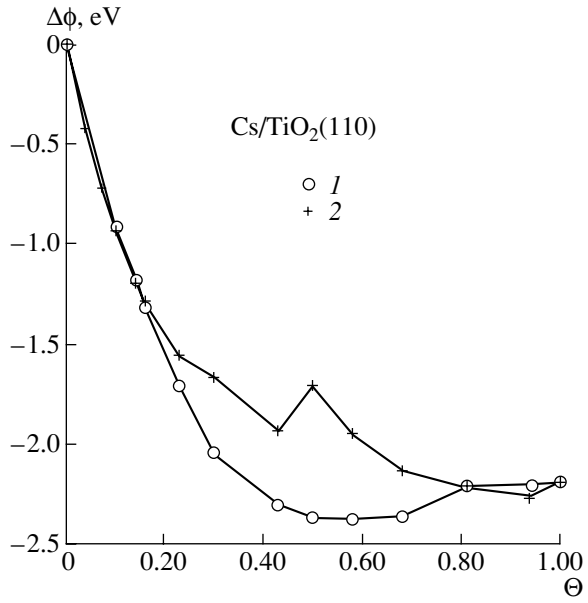


Fig. 3. Same as in Fig. 1 for the Cs/TiO_2 (110) system, using the experimental data of [12].

width of an impurity quasi-level whose finite lifetime hinges on the possibility of an electron tunneling between adatom orbitals and the states in the conduction band of the metal. However, even in this simple case, the scatter in the estimates of Γ is considerable. As to the model of this study, with its nonmetal substrate that has a complicated band structure, here, the parameter Γ takes into account all the processes determining the finite lifetime of the electron on the adatom, i.e., both tunneling and Auger processes (for discussion on such transitions see, e.g., [12]). It is evident that, in this case, estimates of Γ are all the more difficult. Therefore, in this work, a value of $\Gamma = 1$ eV is assumed for all adsorbates.

Thus, in the framework of the modified model, the variation of the work function of rutile caused by the adsorption of alkali metal atoms can be quite adequately described.

ACKNOWLEDGMENTS

This study was carried out in the framework of the Russian Federal Program "Surface Atomic Structures" (grant no. 4.5.99).

REFERENCES

1. V. E. Henrich and P. A. Cox, *The Surface Science of Metal Oxides* (Cambridge Univ. Press, Cambridge, 1994).
2. *Physics and Chemistry of Alkali Metal Adsorption*, Ed. by H. P. Bonzel, A. M. Bradshow, and G. Ertl (Elsevier, Amsterdam, 1989).
3. S. Yu. Davydov and I. V. Noskov, Pis'ma Zh. Tekh. Fiz. **27** (20), 1 (2001) [Tech. Phys. Lett. **27**, 844 (2001)].
4. R. Casanova, K. Prabhakaran, and G. Thornton, J. Phys.: Condens. Matter **3**, S91 (1991).
5. H. Onishi, T. Aruga, C. Egawa, and Y. Iwasawa, Surf. Sci. **199**, 54 (1988).
6. J. P. Muscat and D. M. Newns, J. Phys. C **7**, 2630 (1974).
7. S. Yu. Davydov, Fiz. Tverd. Tela (Leningrad) **19**, 3376 (1977) [Sov. Phys. Solid State **19**, 1971 (1977)].
8. S. Yu. Davydov, Fiz. Tverd. Tela (St. Petersburg) **41**, 1543 (1999) [Phys. Solid State **41**, 1413 (1999)].
9. S. Y. Davydov, Appl. Surf. Sci. **140**, 58 (1999).
10. *Handbook of Physical Quantities*, Ed. by I. S. Grigoriev and E. Z. Meilikhov (Énergoatomizdat, Moscow, 1991; CRC Press, Boca Raton, 1997).
11. J. W. Gadzuk, Phys. Rev. B **1**, 2110 (1970).
12. M. Brause, S. Skordas, and V. Kempter, Surf. Sci. **445**, 224 (2000).
13. A. W. Grant and C. T. Campbell, Phys. Rev. B **55**, 1844 (1997).
14. S. Yu. Davydov, Zh. Tekh. Fiz. **69** (5), 116 (1999) [Tech. Phys. **44**, 584 (1999)].
15. S. Yu. Davydov, Fiz. Tverd. Tela (St. Petersburg) **42**, 1129 (2000) [Phys. Solid State **42**, 1164 (2000)].
16. P. W. Anderson, Phys. Rev. **124**, 41 (1961).
17. W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* (Freeman, San Francisco, 1980; Mir, Moscow, 1983), Vol. 2.

Translated by B. Kalinin

Technical Physics, Vol. 47, No. 11, 2002, pp. 1484–1486. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 72, No. 11, 2002, pp. 140–142.

BRIEF COMMUNICATIONS

Low-Frequency Electromagnetic Radiation Generated by Electron Beams in Gases

G. R. Alimov*, A. T. Muminov*, T. M. Muminov*, U. S. Salikhbaev, A. N. Safarov***, and R. R. Usmanov***

** Research Institute of Applied Physics, National University of Uzbekistan, Tashkent, 700174 Uzbekistan*
*** Institute of Nuclear Physics, Academy of Sciences of Uzbekistan, Tashkent, 702132 Uzbekistan*
**** Samarkand State University, Samarkand, 703004 Uzbekistan*
e-mail: u_salikh@suninp.tashkent.su
Received May 16, 2002

Abstract—13-MeV electron beams are shown to generate low-frequency electromagnetic radiation in the gaseous medium of a 100-m-long dielectric electron pipe. © *2002 MAIK "Nauka/Interperiodica".*

We present the results of studying low-frequency (LF) secondary electromagnetic radiation (SER) generated by electron beams in a rarefied gaseous medium.

The experimental setup (Figs. 1, 2) comprised an MT-22S microtron, an electron beam transportation path with a 100-m-long dielectric electron pipe (glass, inner diameter 50 mm, intensity loss at the exit 15%, residual gas pressure 10^{-3} Pa), antenna feeders, and an on-line measuring module including an SK4-72 spectrum analyzer, a CAMAC system, and an IBM computer [1, 2]. The results were obtained for a pulsed electron beam (electron energy E_e = 13 MeV, average current I_e = 1–10 μA, pulse duration τ = 2.5 μs, and pulse repetition rate f = 388 Hz) in the final 15-m-long section of the dielectric electron pipe, where the residual gas pressure was varied from 1 to 10^5 Pa. A remote control device was used to vary the spatial orientation of the antennas and their position along (L = 0–15 m) and across (d = 0.4–1.1 m) the beam axis within the final section.

Our experiment demonstrated that the interaction of the electron beam with the gaseous medium in the near-field zone gives rise to low-frequency SER with frequencies f that are multiples of the repetition rate of beam pulses (Figs. 3, 4). It is shown that (i) the SER signal U grows linearly with beam current (i.e., with an increase in the number of ionized gas molecules; Fig. 5) and decreases linearly with increasing spacing d between the antenna and the beam axis (i.e., with a decrease in the SER flux intensity inducing the antenna signal; Fig. 5); (ii) an increase in the antenna offset L from the beginning of the section studied or in the residual pressure P in the electron pipe (i.e., the growth of the length or density of a plasma interacting with the electron beam) results in a monotonic rise in the U signal amplitude (Figs. 3, 6); and (iii) the U signal amplitude depends on the antenna orientation, indicating the poorly understood horizontal polarization of the SER electric component (inset in Fig. 2).

Note that the considerable sizes of the frame antenna and, accordingly, the high level of noise induced in it and the use of the 100-m-long biconical vibrator precluded the detection of individual spectral lines. Another point to be emphasized is that, along with the electron-beam-induced SER, the experimental spectrum also included the contribution from the antenna interaction with stray radiations due to electron beam scattering.

The results obtained are insufficient to draw any conclusions concerning the physical nature of the LF secondary electromagnetic radiation discovered.

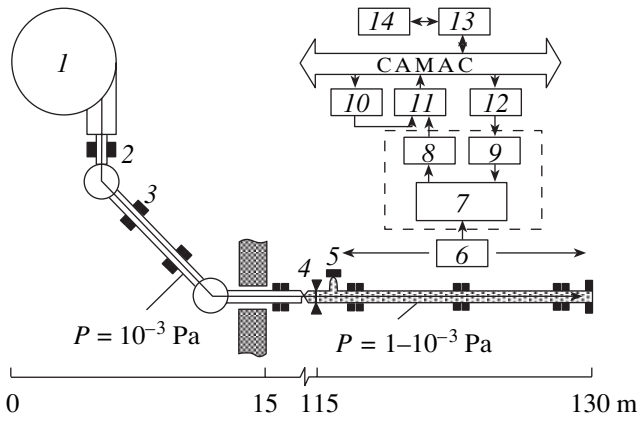


Fig. 1. Experimental setup: (*1*) MT-22S microtron, (*2*) quadrupole magnetic lens doublet, (*3*) rotary magnet, (*4*) vacuum partition in dielectric electron pipe, (*5*) leak, (*6*) movable antennas, (*7*) SK4-72 spectrum analyzer, (*8*) analog interface, (*9*) digital interface, (*10*) generator, (*11*) analog-to-digital convertor, (*12*) digital-to-analog convertor, (*13*) CAMAC crate controller, and (*14*) computer.

1063-7842/02/4711-1484$22.00 © 2002 MAIK "Nauka/Interperiodica"

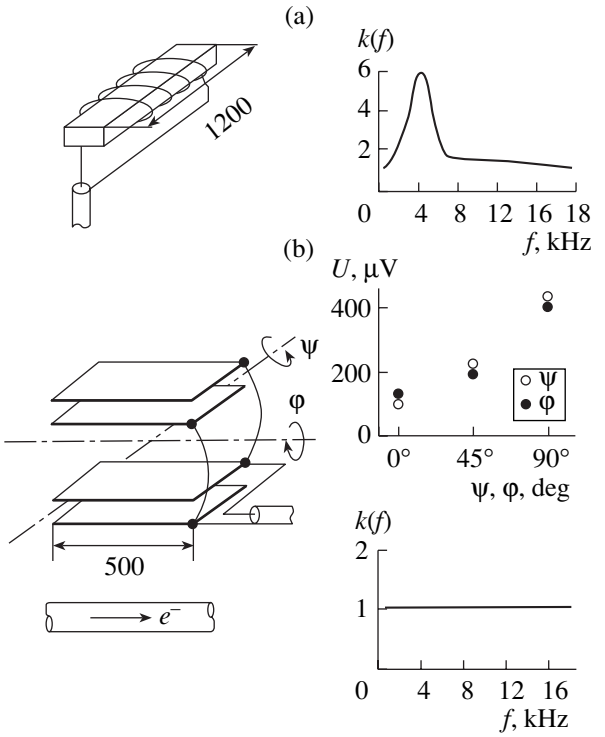


Fig. 2. (a) Ferrite and (b) electrical antennas. The insets show their amplitude–frequency characteristics ($k(f)$ is the frequency-dependent transmission coefficient) and the dependence of the SER signal amplitude U on the orientation angles φ and ψ for the electrical antenna. L = 10 m, d = 0.8 m, P = 6 Pa, and I = 3 μA.

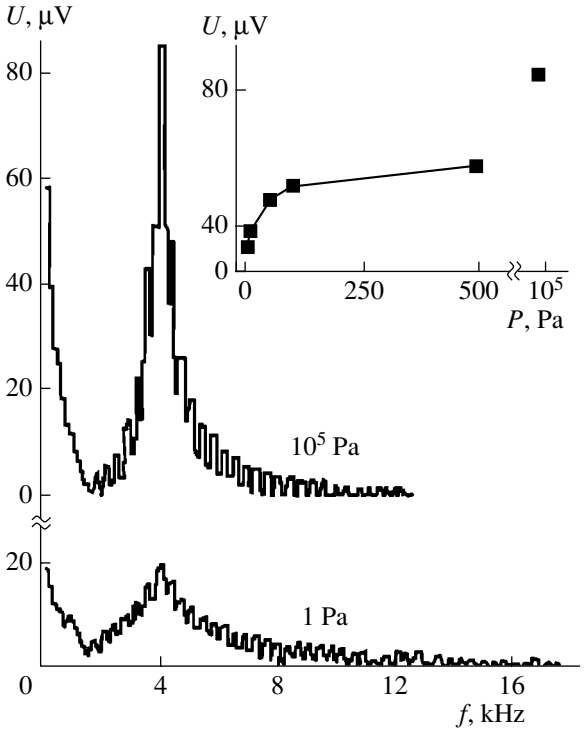


Fig. 3. SER spectrograms and the amplitudes U vs. pressure P for the ferrite antenna with L = 10 m, d = 0.8 m, and I = 3 μA.

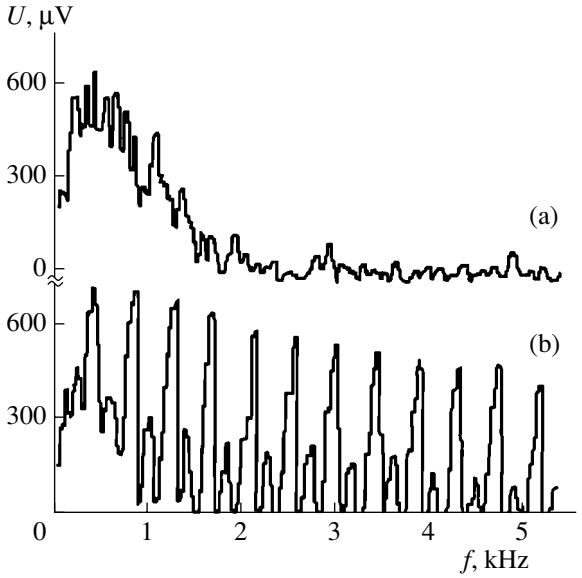


Fig. 4. Spectrograms of (a) background and (b) SER without background (both measured with the electrical antenna). L = 10 m, d = 0.8 m, P = 6 Pa, and I = 3 μA.

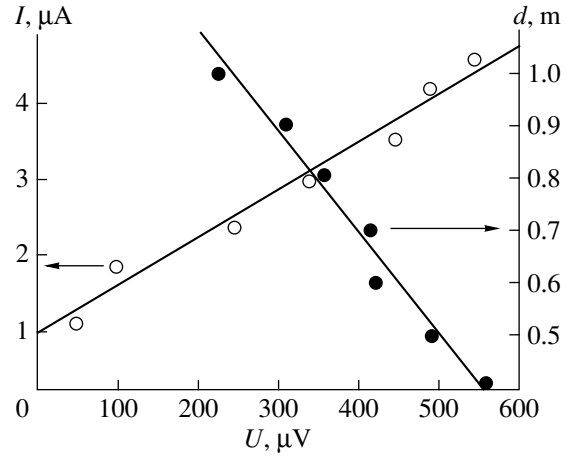


Fig. 5. SER amplitude U vs. beam current I (at d = 0.8 m) and spacing d (at I = 3 μA) from the electrical antenna to the beam axis (L = 10 m, P = 6 Pa).

(1) The SER observed cannot be explained by the excitation of the antenna feeders, since their amplitude–frequency characteristics have no resonance peaks in the range under study because of the low Q factor of the oscillatory circuit formed by the antenna feeders and the input circuit of the spectrum analyzer.

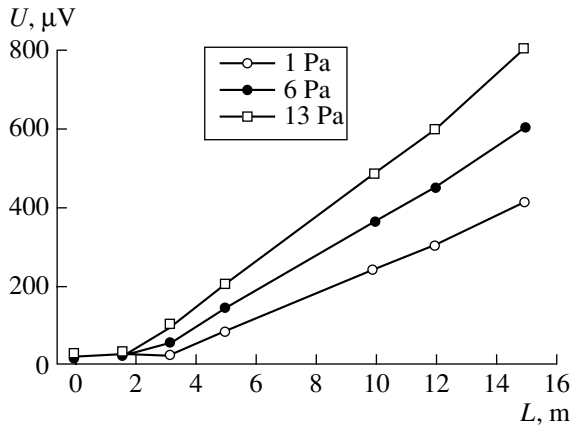


Fig. 6. SER amplitude U vs. spacing L between the electrical antenna and the beginning of the section under study (d = 0.8 m, I = 3 μA).

(2) The amplitudes of the SER harmonic components bear no resemblance to the amplitude spectrum of a periodic sequence of rf or video signals, in the Fourier expansion of which the harmonic amplitudes should noticeably decay.

(3) The SER observed is a radiation with a continuous frequency spectrum in the range from $f = 0$ to $f = 3c\gamma^3/2R$, where c is the velocity of light, $\gamma = E_e/mc^2$, and R is the radius of curvature of the electron trajectory.

Elucidating the origin of an SER generated by pulsed beams of charged particles in a gaseous medium evidently needs careful experiments with more intense beams of particles.

REFERENCES

1. G. R. Alimov, S. B. Baraev, S. A. Koval', *et al.*, Izv. Akad. Nauk, Ser. Fiz. **57** (4), 66 (1993).
2. G. R. Alimov, F. R. Kungurov, T. M. Muminov, *et al.*, Uzb. Fiz. Zh., No. 4, 78 (1997).

Translated by A. Sidorova