Technical Physics, Vol. 45, No. 12, 2000, pp. 1513–1517. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 1–5.

Nonlinear Effects in Trapped Modes of Standing Waves on the Surface of Shallow Water

D. A. Indeĭtsev and E. V. Osipova

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

e-mail: oav@phase.ipme.ru

Received January 25, 1999

Abstract—It was shown that traveling waves may coexist with standing waves in a planar infinitely long channel filled by ideal liquid with a free surface. The standing waves are localized near a dynamic inclusion—a massive die on an elastic base. The amplitude of the traveling waves may be turned to zero by appropriately selecting the vibration frequency of the die. The standing waves arise because the vibration eigenfrequencies have a mixed spectrum; that is, the discrete and continuous spectra superpose. Nonlinear effects were observed for the first time when standing waves form in shallow water. In particular, a relationship between the die weight necessary to excite trapped modes, die dimensions, and vibration frequency was derived. It was shown that the nonlinear effects cause double-frequency traveling waves with amplitudes of the next order of smallness. These traveling waves vanish if the die geometry is properly chosen, as for the waves of the zeroth order. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Resonant properties of systems with at least one infinitely long boundary have recently attracted considerable interest [1–6]. This problem is of both practical value (for example, resonant behavior of mobile loads subjected to seismic forces) and theoretical significance (determination of discrete values of singular operators with a mixed spectrum). To date, the resonant properties of such systems have been studied only in the linear approximation [1, 2, 5, 6], apparently because of the complexity of the related equations. Many issues still remain unresolved as to how the nonlinearity influences the existence of trapped vibrational modes and their properties, how the resonant characteristics depend on vibration amplitudes, whether nonlinear waves appear, etc. To tackle these and other questions, we studied the nonlinear resonant properties of the simplest yet real system with a mixed spectrum—a planar channel filled by an ideal incompressible liquid and having a mobile massive die as an inclusion at the bottom. Nonlinearity is caused by the free surface of the liquid.

GRAVITY WAVES ON THE SURFACE OF SHALLOW WATER IN A PLANAR CHANNEL WITH AN INCLUSION

Let a die of mass M and width $2a$ execute vertical vibrations with a frequency ω_a while in continuous contact with an incompressible potential liquid that fills a planar infinitely long channel of depth H (the die is attached to a stationary base via an ideal spring of rigidity κ) (Fig. 1). Then, the coordinate of the upper edge of the die is given by

$$\xi = \xi_0 \sin\omega_a t = \varepsilon H \sin\omega_a t, \qquad (1)$$

where $t \geq 0$ is time and ε is a small parameter defined as the ratio of the die vibration amplitude ξ_0 to the channel depth H.

The die generates vibrational motion in the liquid that is characterized by the velocity field potential φ, which is related to the space coordinates x and y and time t through the equation [7]

$$\frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} = 0. \qquad (2)$$

At the bottom of the channel, the boundary condi-

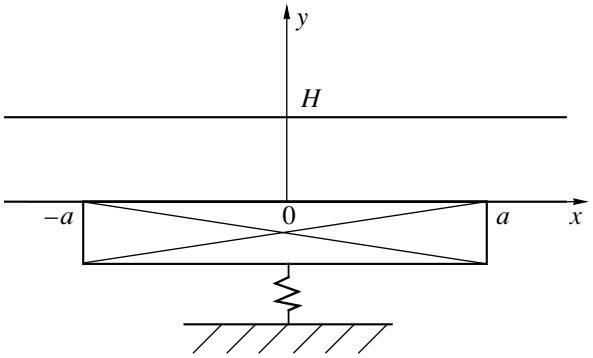


Fig. 1. Die of width $2a$ executing small-amplitude vibrations at the bottom of a liquid-filled infinitely long channel of depth H.

1063-7842/00/4512-1513$20.00 © 2000 MAIK "Nauka/Interperiodica"

tion is the equality of the liquid and die velocities at the die–liquid interface:

$$\left.\frac{\partial\varphi}{\partial y}\right|_{y=0} = \varepsilon H\omega_a\cos\omega_a t\,\theta\left(\frac{x}{a}\right), \tag{3}$$

$$\theta(x) \equiv \begin{cases} 1, & |x| \le 1, \\ 0, & |x| > 1. \end{cases} \tag{4}$$

Strictly speaking, condition (3) should be set at the boundary $y = \xi$, not at $y = 0$. However, since the given condition is linear and ε is small, this inaccuracy is physically insignificant. On the free surface ($y = h(x, t) + H$, where h is the gravity wave height at a point x at a time instant t), the boundary condition with the neglect of capillary effects has the form [7]

$$\frac{\partial\varphi}{\partial t} + \frac{1}{2}\left(\frac{\partial\varphi}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial\varphi}{\partial y}\right)^2 + gh = 0, \tag{5}$$

where g is the free-fall acceleration.

The value of h is determined from the equation [7]

$$\frac{\partial h}{\partial t} + \left(\frac{\partial\varphi}{\partial x}\right)_{y=h+H}\frac{\partial h}{\partial x} = \left(\frac{\partial\varphi}{\partial y}\right)_{y=h+H}, \tag{6}$$

$$h(x, 0) = 0. \tag{7}$$

In addition, the coordinate ξ of the upper edge of the die must satisfy the die vibration equation [5]

$$M\frac{d^2\xi}{dt^2} = -\kappa\xi + M_0 g\frac{\xi}{H} - \rho\int_{-a}^{a}\left(\frac{\partial\varphi}{\partial t}\right)_{y=0} dx, \tag{8}$$

where $M_0 = 2\rho aH$ is the mass of the liquid above the stationary die and ρ is the density of the liquid; here, it is taken into account that the force $M_0 g$ is counterbalanced by an external force acting on the die that provides the fulfillment of the condition $\xi = 0$ under equilibrium [5].

First consider the closed system of equations (2), (3), (5), and (6), which describes the propagation of nonlinear gravity waves over the channel surface (for a given type of die motion). Introduce new dimensionless variables

$$z = \frac{x}{a}, \quad \eta = \frac{y}{H}, \quad \tau = \frac{\sqrt{gH}}{a}t,$$
$$\psi = \frac{\varphi}{\varepsilon a\sqrt{gH}}, \quad u = \frac{h}{\varepsilon H} \tag{9}$$

and dimensionless parameters

$$\omega = \frac{\omega_a a}{\sqrt{gH}}, \quad \delta = \frac{H^2}{a^2}. \tag{10}$$

Then, Eqs. (2), (3), (5), and (6) will take the form

$$\psi_{\eta\eta} + \delta\psi_{zz} = 0,$$

$$\psi_\eta|_{\eta=0} = \delta\omega\cos\omega\tau\,\theta(z), \tag{11}$$

$$\left\| \begin{array}{l} \psi_\tau + \dfrac{\varepsilon}{2}\left(\psi_z^2 + \dfrac{1}{\delta}\psi_\eta^2\right) + u = 0 \\ u_\tau - (1/\delta)\psi_\eta + \varepsilon\psi_z u_z = 0 \end{array} \right\| \quad \text{where } \eta = 1 + \varepsilon u.$$

We are interested in the behavior of the gravity waves in the simplest case of shallow water, where dispersion effects are negligible, that is, at $\delta \longrightarrow 0$. By shallow water is meant the situation when the depth is much smaller than the characteristic length of a wave solution. As will be shown below, this characteristic length in our case is on the order of a; hence, the term "shallow water" will be used if $\delta \ll 1$. Following the general theory [7–9], we expand the function ψ in a power series in η:

$$\psi(z, \eta, \tau) = \psi_0(z, \tau) + \psi_1(z, \tau)\eta + \frac{1}{2}\psi_2(z, \tau)\eta^2 + \ldots. \tag{12}$$

By substituting (12) into the first two equations of set (11), it is easy to express ψ_i ($i = 1, 2, 3, \ldots$) in terms of ψ_0. Eventually, we have

$$\psi(z, \eta, \tau) = \psi_0(z, \tau) + \delta\omega\cos\omega\tau\,\theta(z)\eta - \frac{\delta}{2}\psi_{0zz}\eta^2 + O(\delta^2). \tag{13}$$

This series obviously converges, since $\eta \sim 1$, $\delta \ll 1$, and each of the next terms is of a higher power of δ [9]. Substituting (13) into the last two equations of set (11) yields the equation for the height u of the wave and that for vertical movement velocity in it, $v \equiv \psi_{0z}$:

$$\begin{aligned} &u_\tau - \omega\cos\omega\tau\,\theta(z) \\ &+ v_z(1 + \varepsilon u) - (\delta/6)v_{zzz} + \varepsilon v u_z = 0, \\ &v_\tau + u_z - (1/2)\delta v_{zzz} + \varepsilon v v_z = 0. \end{aligned} \tag{14}$$

In (14), only the terms linear in ε and δ are left; those of higher degrees in ε and δ are rejected. Then, on rearrangement, we obtain

$$\begin{aligned} &u_{zz} - u_{\tau\tau} - \omega^2\sin\omega\tau\,\theta(z) \\ &-\frac{\varepsilon}{2}(v^2)_{zz} - \varepsilon(uv)_{z\tau} - \frac{\delta}{3}v_{zzz\tau} = 0, \\ &v_\tau + u_z + \varepsilon v v_z - \frac{\delta}{2}v_{zzz} = 0. \end{aligned} \tag{15}$$

According to the general perturbation theory, we will seek the solution to set (15) as series in ε [10]:

$$u = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \ldots, \tag{16}$$

$$v = v_0 + \varepsilon v_1 + \varepsilon^2 v_2 + \ldots, \tag{17}$$

$$\omega = \omega_0 + \varepsilon^2\omega_2 + \ldots. \tag{18}$$

Renormalization of frequency (18) is necessary in order to eliminate the secular terms in (16) and (17) (they are absent in the first order in ε). Our goal is to solve set (15) in the first order in ε [terms of higher orders in ε are insignificant, since set (15) considers only first-order (in ε) nonlinear effects). To do this, we substitute (16)–(18) into (15) and equate the coefficients multiplying the terms with the same powers of ε:

$$\begin{gathered} u_{0zz} - u_{0\tau\tau} - \omega_0^2 \sin\omega_0\tau\theta(z) - (\delta/3)v_{0zzz\tau} = 0, \\ v_{0\tau} + u_{0z} - (\delta/2)v_{0zzz} = 0, \end{gathered} \tag{19}$$

$$\begin{gathered} u_{1zz} - u_{1\tau\tau} + (1/2)(v_0^2)_{zz} - (u_0 v_0)_{z\tau} - (\delta/3)v_{1zzz\tau} = 0, \\ v_{1\tau} + u_{1z} + v_0 v_{0z} - (\delta/2)v_{1zzz} = 0, \end{gathered} \tag{20}$$

$$\begin{aligned} &u_{2zz} - u_{2\tau\tau} - 2\omega_0\omega_2 \sin\omega_0\tau\theta(z) - \omega_0^2 \cos\omega_2\tau\theta(z) \\ &+ (v_0 v_1)_{zz} - (u_0 v_1 + u_1 v_0)_{zz} - (\delta/3)v_{2zzz\tau} = 0, \end{aligned} \tag{21}$$

$$v_{2\tau} + u_{2z} + (v_0 v_1)_z - (\delta/2)v_{2zzz} - \delta\omega_0^2 \cos\omega_2\tau\theta = 0.$$

To elucidate the physical meaning of the trapped vibrational modes and find their basic properties, let us consider linear approximation (19) in the shallow water limit ($\delta \longrightarrow 0$). Set (19) is split into two independent equations, with the gravity wave amplitude $f \equiv u_0$ having the form

$$f_{zz} - f_{\tau\tau} - \omega_0^2 \sin\omega_0\tau\theta(z) = 0. \tag{22}$$

Boundary conditions for (22) are easily obtained from (3) and (7):

$$f(z, 0) = 0; \quad f_\tau(z, 0) = \omega_0\theta(z). \tag{23}$$

The solution to (22) in view of (23) has the form

$$\begin{aligned} u(z, \tau) &= \theta(z)\sin\omega_0\tau \\ &+ \frac{1}{2}\mathrm{sgn}(z-1)\sin[\omega_0(\tau - |z-1|)]\Phi(\tau - |z-1|) \\ &- \frac{1}{2}\mathrm{sgn}(z+1)\sin[\omega_0(\tau - |z+1|)]\Phi(\tau - |z+1|), \end{aligned} \tag{24}$$

where Φ is the Heaviside function.

This formula describes the origination and propagation of linear gravity waves in shallow water under the action of a vibrating die. It discloses the physical meaning of the trapped vibrational modes. In fact, formula (24) can be recast as follows:

$$f = \begin{cases} \sin\omega_0\tau, & \tau \le 1 - |z|, \\ \sin\omega_0\tau - \dfrac{1}{2}\sin\omega_0(|z| + \tau - 1), & \\ & 1 - |z| < \tau \le |z| + 1, \\ (1 - \cos\omega_0\cos\omega_0 z)\sin\omega_0\tau & \\ + \sin\omega_0\cos\omega_0 z\cos\omega_0\tau, & \\ & \tau > |z| + 1 \end{cases} \tag{25}$$

above the die and

$$f = \begin{cases} 0, & \tau \le |z| - 1, \\ \dfrac{1}{2}\sin\omega_0(\tau - (|z| - 1)), & |z| - 1 < \tau \le |z| + 1, \\ \sin\omega_0\cos\omega_0(|z| - \tau), & \tau > |z| + 1 \end{cases} \tag{26}$$

outside the die. It follows that a vibrating die generates three types of gravity waves: standing waves over a die that go in the opposite directions within the domain $\tau - 1 \le |z| \le \tau + 1$ and traveling waves with an amplitude $\sin\omega_0$. They all result from the interference of waves at a given point that have come from each point of the vibrating die surface at a given time instant. The traveling waves continuously transfer the energy of the vibrating die to infinity; hence, to maintain the vibration amplitude, an energy proportional to $\sin\omega_0$ must be delivered to the die. Therefore, Eq. (8) will have a solution like $\sin\omega t$ only if condition $\sin\omega_0 = 0$, or $\omega_0 = \omega_k$, where

$$\omega_k = k\pi \ (k = 1, 2, 3, \ldots) \tag{27}$$

is met (this result follows from the mixed spectrum of the problem, i.e., the superposition of discrete and continuous spectra [5]). In this case, at $\tau > 2$, $f(z, \tau) = f_k(z, \tau)$, where

$$\begin{aligned} f_k(z, \tau) &= (1 + (-1)^{k-1}\cos k\pi z)\sin k\pi\tau\theta(z) \\ &+ \frac{(-1)^{k-1}}{2}\sin[k\pi(|z| - \tau)]\theta(|z| - \tau) \end{aligned} \tag{28}$$

is the amplitude of a linear gravity wave that corresponds to the trapped mode with the number k. The wave with $k = 2$ for $\tau = 3.4$ is shown in Fig. 2.

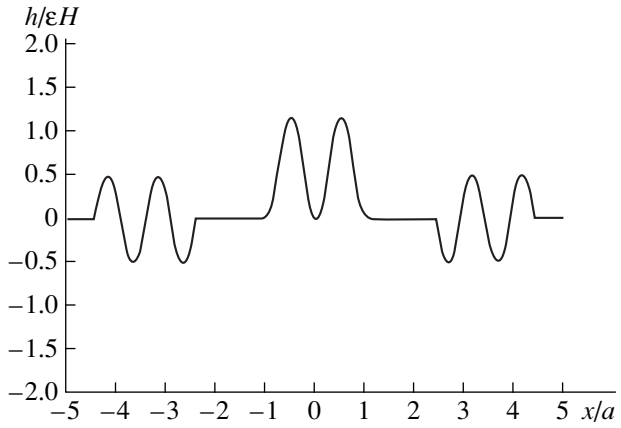


Fig. 2. Waves excited by the die vibrating with the frequency of the trapped mode with $k = 2$ at the time instant $\tau = 3.4$.

NONLINEAR EFFECTS

Now, let us see how the nonlinear effects influence the frequencies of the trapped modes and their associated gravity waves. If we immediately put $\delta = 0$ in initial set (15) and consider the perturbation theory in terms of ε, even first-order secular terms like $\varepsilon\tau\sin\omega_0\tau$ appear. They cannot be eliminated either by renormalization or by the method of nonlinear scales [10]. The reason for this is the absence of dispersion, which counterbalances nonlinearity in this case. Consequently solving sets (19)–(21), one easily obtains that the first-order secular terms disappear and new, second-order terms like $\varepsilon^2\tau\sin\omega_0\tau$ arise. In fact, the solution to (19) contains standing and traveling waves, the amplitude of the latter being equal to $\sin(\omega_0/\sqrt{1-\delta\omega_0^2/3})$. By taking the die vibration frequency in the form

$$\omega_0 = \omega_k \equiv \frac{k\pi}{\sqrt{1+\frac{\delta}{3}(k\pi)^2}}, \tag{29}$$

the traveling waves can be eliminated. In the shallow water approximation, the remaining standing waves are described by

$$u_0 \sim [1+(-1)^{k-1}\cos k\pi x]\theta(x)\sin\omega_k t. \tag{30}$$

The solution to (20) also contains standing waves (their specific form is of no significance) and double-frequency traveling waves

$$u_1 \sim \frac{3}{8}\frac{\sin(\delta(k\pi)^3)}{\delta(k\pi)^2}\sin[2k\pi(|x|-t)]. \tag{31}$$

If we take H or a such that

$$\delta = \frac{H^2}{a^2} = \frac{m}{k^3\pi^2}, \tag{32}$$

where m is a positive integer, the amplitude of these traveling waves will also vanish.

The solution to set (21) contains, together with traveling waves, a standing wave with growing amplitude, that is, the secular term

$$u_2 \sim \left(\frac{9}{8\delta} - 2\omega_0\omega_2\right)t\cos\omega_0 t X(x)\theta(x). \tag{33}$$

Here, $X(x)$ is the standing wave profile, which is of no significance in this case. This nonuniform suitability of expansion (16) is readily eliminated by renormalizing frequency (18) [10]. Taking

$$\omega_2 = \frac{9}{16\omega_0\delta}, \tag{34}$$

we provide uniformity in the first approximation. Below, we give the final expression for the frequencies of the trapped vibrational modes and their associated gravity wave heights in the dimensional form (in the first approximation in ε and in the shallow water limit):

$$\omega_k = k\pi\frac{\sqrt{gH}}{a}\frac{1+\left(\frac{3\xi_0 a}{4k\pi H^2}\right)^2}{\sqrt{1+\frac{1}{3}\left(\frac{k\pi H}{a}\right)^2}}, \tag{35}$$

$$\begin{aligned} h(x,t) &= \xi_0\left[1+(-1)^{k-1}\cos\frac{k\pi x}{a}\right]\theta\left(\frac{x}{a}\right)\sin\omega_k t \\ &+ \frac{(-1)^{k-1}}{2}\xi_0\sin\left(k\pi\frac{|x|}{a}-\omega_k t\right)\theta\left(\frac{|x|}{a}-\frac{\omega_k}{k\pi}t\right) \\ &+ \frac{3\xi_0^2 a^2}{8(k\pi)^2H^3}\sin\left(\frac{H^2}{a^2}(k\pi)^3\right)\sin\left[2\left(k\pi\frac{|x|}{a}-\omega_k t\right)\right] \\ &\times\left[\theta\left(\frac{|x|}{a}\right)-1\right], \quad t > \frac{2a}{\sqrt{gH}}. \end{aligned} \tag{36}$$

Standing waves of amplitude ε are omitted here for simplicity. Let us now find the die mass M necessary for the die to execute harmonic vibrations with a trapped frequency ω_k. It follows from (36) that, with condition (32) satisfied, i.e., in the absence of traveling waves, expression

$$M_0 g\frac{\xi}{H} - \rho\int_{-a}^{a}\left(\frac{\partial\varphi}{\partial t}\right)_{y=0}dx = 0 \tag{37}$$

is fulfilled at any amplitude ξ_0; hence, from (8), we have

$$M\frac{d^2\xi}{dt^2} = -\kappa\xi. \tag{38}$$

Consequently, in the absence of traveling waves, the mass of the die must satisfy the condition $M = M_k$, where

$$M_k = \frac{a}{(k\pi)^2}\frac{\kappa}{gH}\left[1+\left(\frac{3\xi_0 a}{4k\pi H^2}\right)^2\right]^{-2}. \tag{39}$$

Simultaneously, nonlinear and dispersion effects should be small:

$$\frac{\xi_0}{H} \ll 1, \quad \left(\frac{3\xi_0 a}{4k\pi H^2}\right)^2 \ll 1, \quad \left(\frac{H}{a}\right)^2 \ll 1. \tag{40}$$

Specifically, at $H = 1$ m, $a = 10$ m, $\kappa = 10^5$ N/m, and $\xi_0 = 10^{-1}$ m, we have $M_2 = 2.5\times10^3$ kg and $\omega_2 = 2$ s^{-1} for $k = 2$. If condition (32) does not hold, the vibration amplitude will remain constant if a force on the order of $Mg\xi_0/H$ is applied to the die, since traveling waves carry away the vibration energy of the die.

Thus, a dynamic inclusion of a finite size causes trapped vibrational modes to appear. Waves propagat-

ing from each point of the inclusion surface cancel out away from the inclusion. Because of the nonlinear effects, both the vibration frequency and the mass of the die depend on the vibration amplitude (the shape of the vibrations is amplitude-independent). Gravity waves are described by set (15) in the first-order approximation in ε and δ. If, however, only the zeroth approximation is of interest (that is, if only the terms linear in ε or δ that are responsible for nonuniformity are left), set (15) is greatly simplified and takes the form

$$u_{zz} - u_{\tau\tau} - \omega^2 \sin\omega\tau\theta(z) + \frac{3}{2}\varepsilon(u^2)_{zz} + \frac{\delta}{3}u_{zz\tau\tau} = 0, \quad (41)$$

which is the Boussinesque approximation [7, 9]. Therefore, Eq. (41) with initial condition (23) is the most simplified equation of those describing trapped vibrational modes in the nonlinear case. The vibration parameters are set by formulas (35), (36), and (39).

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (grant no. 99-01-00693).

REFERENCES

1. V. A. Babeshko, I. I. Vorovich, and I. F. Obraztsov, Izv. Akad. Nauk SSSR, Ser. Mekh. Tverd. Tela, No. 3, 74 (1990).
2. D. V. Evans, C. M. Linton, and F. Ursell, Q. J. Mech. Appl. Math. **46**, 253 (1993).
3. P. McIver and M. McIver, Q. J. Mech. Appl. Math. **50**, 165 (1997).
4. N. Kuznetsov and P. McIver, Q. J. Mech. Appl. Math. **50**, 565 (1997).
5. D. A. Indeĭtsev and E. V. Osipova, Zh. Tekh. Fiz. **66** (8), 124 (1996) [Tech. Phys. **41**, 811 (1996)].
6. D. Indejtchev, A. Abramjan, and V. Andreev, J. Acoust. Soc. Am. **95**, 3007 (1994).
7. G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974; Mir, Moscow, 1977).
8. L. N. Sretenskiĭ, *The Theory of Wave Motion in Liquid* (Nauka, Moscow, 1977).
9. L. Debnath, *Nonlinear Water Waves* (Academic, London, 1994).
10. A.-H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973; Mir, Moscow, 1976).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 12, 2000, pp. 1518–1523. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 6–11.

Effective Calculation of Energy Loss and Scattering Angle at Electron–Target Material Inelastic Interaction

E. Yu. Flegontova*, L. A. Bakaleĭnikov*, K. Yu. Pogrebitskiĭ*, Hwack-Joo Lee, Yang-Koo Cho**, Hyun-Min Park**, and Yong-Won Song****

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

*** Korean Research Institute of Standards and Science, Taejn, Republic of Korea*

Received March 1, 2000

Abstract—A Monte Carlo-based method for computing electron–target inelastic interaction is presented. It uses the double-differential inelastic scattering cross section. A resource-saving algorithm that simulates the interaction using the double-differential cross section was elaborated. The cross section was calculated from data on the optical permittivity of the target. © *2000 MAIK "Nauka/Interperiodica".*

The search for new-generation materials (including quantum-dimensional structures with an active region size on the order of several tens of atoms) requires object parameters to be measured at least with a nanometer resolution. In this respect, of great importance are methods that either employ an electron probe or detect electrons as a basic signal. Among these are electron probe microanalysis, Auger electron spectroscopy, X-ray photoelectron spectroscopy, and X-ray-induced electron emission XIEES spectrometry. These techniques allow one to obtain chemical compositions, dimensions of inhomogeneities (thicknesses of layers), atomic structure, electronic configuration, and other material properties with a subatomic resolution. To quantitatively determine desired parameters (that is, to simulate basic effects), information on the electron distribution function is needed. This function is obtained from a kinetic equation, solvable by the Monte Carlo procedure, which is among the simplest and most universal numerical methods. Present-day computing facilities make it possible to employ first-order scattering models, where each electron–target interaction is described through the differential cross sections of different processes. In this case, the optimum way of realizing the Monte Carlo procedure by software and even the very possibility of its realization depend on how information on the differential cross sections is represented.

In this paper, we suggest a method for computing electron–target inelastic interaction that is based on the use of the double-differential inelastic scattering cross section and improves the efficiency of the Monte Carlo procedure.

It is known [1] that electron–target inelastic interaction can be described in terms of permittivity $\tilde{\varepsilon}(q, \omega)$, which is related to the double-differential (DD) inelastic scattering cross section by the relationship

$$\frac{d^2\sigma}{d(\hbar\omega)d(\hbar q)} = \frac{1}{\pi a_0 q E}\mathrm{Im}\left(-\frac{1}{\tilde{\varepsilon}(q, \omega)}\right), \qquad (1)$$

where E is the electron energy with reference to the conduction band bottom, $\hbar\omega$ and $\hbar q$ are changes in the electron energy and momentum due to inelastic scattering, and $a_0 = \hbar^2/me^2$ is the Bohr radius. Analytical forms of the permittivity $\tilde{\varepsilon}(q, \omega)$ are known only for several simple cases, for instance when the approximation of noninteracting electron gas is true. Otherwise, one should employ experimental data. However, the well-known EELS experiments were performed in the low-q region, and optical experiments provide only the values of $\tilde{\varepsilon}(0, \omega)$. Thus, the function $\mathrm{Im}(-1/\tilde{\varepsilon}(q, \omega))$ is specified only in the vicinity of the straight line $q = 0$. To extend the domain of definition of $\mathrm{Im}(-1/\tilde{\varepsilon}(0, \omega))$ over the entire plane, we can take advantage of the extrapolation method from [2–4]. In this method, the function $\mathrm{Im}(-1/\tilde{\varepsilon}(q, \omega))$ is represented as the integral over the frequencies of probable plasmon excitations; the relationship between ω, q, and the plasmon frequency ω_p is then given by the dispersion relation

$$\omega = \omega_q(q, \omega_p).$$

After extrapolating $\mathrm{Im}(-1/\tilde{\varepsilon}(0, \omega))$, the expression for the DD inelastic scattering cross section takes the form [4]

$$\frac{d^2\sigma}{d(\hbar\omega)d(\hbar q)} = \frac{1}{\pi a_0 q E}\frac{\omega_0}{\omega}\mathrm{Im}\left(-\frac{1}{\varepsilon(\hbar\omega_0)}\right), \qquad (2)$$

where $\varepsilon(\hbar\omega) = \tilde{\varepsilon}(0, \omega)$ is the optical permittivity and ω_0

is a positive solution of the equation

$$\omega = \omega_p(q, \omega_0).$$

Integrating the DD cross section presented by Eq. (2) over all possible values of the transferred momentum (that is, over all possible angles of inelastic scattering) yields the energy-loss dependence of the differential cross section:

$$\frac{d\sigma}{d\hbar\omega} = \int_0^{\hbar q_{\max}} \frac{d^2\sigma}{d(\hbar\omega)d(\hbar q)} d\hbar q. \quad (3)$$

Here, $\hbar q_{\max}$ is the maximum value of the transferred momentum. The specific form of the function $\omega_0(\omega, q)$ and, therefore relationships (2) and (3) for the differential cross section, are defined by choosing a dispersion relation. In collisions with great energy and momentum transfer, an electron behaves as a classical object; therefore, $\omega_q \longrightarrow \hbar q^2/2m$ if $q \longrightarrow \infty$. Dispersion relations that obey this condition can be taken in different forms. In [5, 6], the dispersion relation was approximated by the expression [7]

$$\omega_q^2 = \omega_0^2 + \frac{1}{2}v_F^2(\omega_0)q^2 + (\hbar q^2/2m)^2, \quad (4)$$

where v_F is the Fermi velocity.

Taking the dispersion relation in the form [8, 9]

$$\omega_q = \omega_p + \frac{\hbar q^2}{2m} \quad (5)$$

leads to the simplest expression for the solution ω_0 in Eq. (2):

$$\omega_0 = \omega - \frac{\hbar q^2}{2m}. \quad (6)$$

The use of (6) significantly alleviates the calculation of the quantities that control electron–target inelastic interaction. It was shown [3] that the difference in the mean free paths obtained with the use of expressions (4) and (5) is insignificant at $E > 100$ eV. Substituting (2) and (6) into (3) yields

$$\frac{d\sigma}{d\hbar\omega} = \frac{1}{4\pi^2 a_0 E}\int_0^{2\hbar\omega - E + \sqrt{E(E+(-\hbar)\omega)}} \frac{x}{(\hbar\omega - x)\hbar\omega}\mathrm{Im}\left(\frac{-1}{\varepsilon(x)}\right)dx. \quad (7)$$

Using Eq. (7), we calculated the energy-loss-depended differential cross sections for Al, Au, and Ge in the energy range of 5 eV < E < 30 keV. From the obtained relationships, some conclusions regarding the properties of Eq. (7) can be drawn. For instance, it was found that, at electron energies of several keV or more, Eq. (7) can be approximated by

$$\frac{d\sigma}{d\hbar\omega} \approx C(E)W(\hbar\omega), \quad (8)$$

that is, within this energy range, the energy-loss dependence of the differential cross section for a given target is described by the universal function $W(\hbar\omega)$, which is free of the incident electron energy E. With energies E and energy losses $\hbar\omega$ much greater than the ionization potentials of all electron shells, differential cross section (7) asymptotically approaches that of free–free electron interaction [10]; hence, the functions $W(\hbar\omega)$ and $C(E)$ have the asymptotics:

$$W(\hbar\omega) \longrightarrow 1/(\hbar\omega)^2, \quad C(E) \longrightarrow C_0/E. \quad (9)$$

Here, C_0 is a constant depending on the target composition. When $\hbar\omega$ is about several tens of electron volts, the function $W(\hbar\omega)$ is rather complex, reflecting the form of $\mathrm{Im}(-1/\varepsilon(\hbar\omega))$. Nevertheless, it can be approximated by an analytical expression that involves a wealth of parameters defining the position and character of basic singularities. Such a description of the function $W(\hbar\omega)$ is more adequate than its definition through a data array and contains more information on its structure. The analytical expression approximating the function $W(\hbar\omega)$ has the form

$$W(x) = x^{-2-\rho}\prod_{k=1}^{n}\left[1 + \left(\frac{x}{d_k}\right)\right]^{r_k} W_0(x), \quad \rho = \sum_{k=1}^{n} r_k, \quad (10)$$

$$W_0(x) = \prod_{i=1}^{m}\left[\arctan\left(\alpha_i\left(\frac{x}{w_i} - 1\right)\right) + \frac{\pi}{2}\right]^{\gamma_i} \times \left[\arctan\left(\beta_i\left(\frac{w_i}{x} - 1\right)\right) + \frac{\pi}{2}\right]^{\gamma_i}. \quad (11)$$

The first two factors in (10) define the general form of the curve $W(x)$ and ensure its correct asymptotic behavior at large values of x, and either of the factors in product (11) specify the position and shape of one of the peaks. The values of the parameters r_k, d_k, α_i, β_i γ_i, w_i, n, and m obtained for Al, Au, and Ge are listed in the table. The plots of the differential cross sections calculated with the use of formula (7) and approximated by relationship (10) are presented in Figs. 1–3 for energies of 2, 10, and 30 keV.

Upon Monte Carlo numerical simulation of the electron trajectory, the electron energy loss and a change in the direction of its motion due to inelastic interaction with a target can be found by solving two transcendental equations.

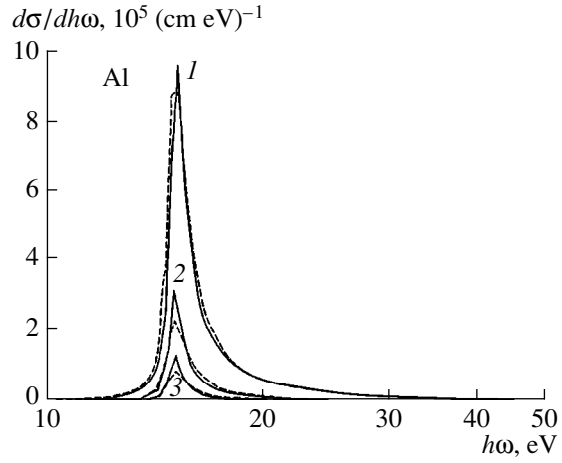


Fig. 1. Differential energy loss cross section vs. energy approximated by the analytical dependence for the Al target. The initial electron energy E = (*1*) 2, (*2*) 10, and (*3*) 30 keV. Solid lines, calculation; dashed lines, approximation.

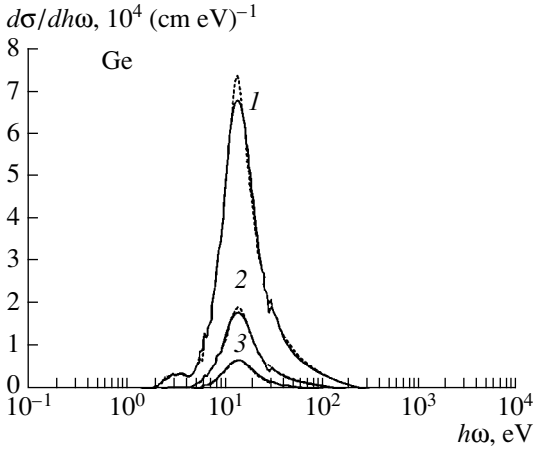


Fig. 2. The same as in Fig. 1 for Ge.

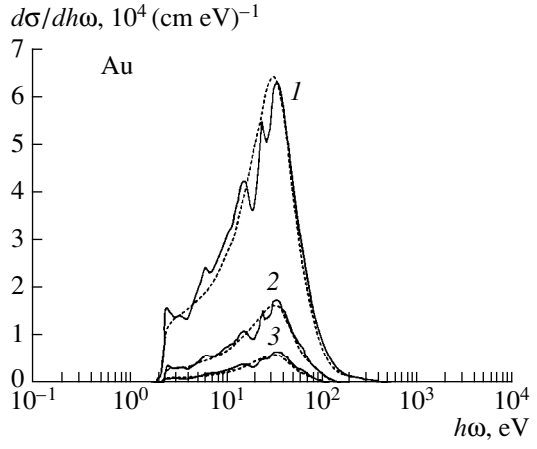


Fig. 3. The same as in Fig. 1 for Au.

The equation for energy loss has the form

$$R = \frac{\int_0^{Q\hbar\omega_{\max}} \frac{d\sigma}{d\hbar\omega} d\hbar\omega}{\int_0^{\hbar\omega_{\max}} \frac{d\sigma}{d\hbar\omega} d\hbar\omega} = \frac{G_1(E, Q)}{G_1(E, 1)}, \quad (12)$$

where Q is the relative energy loss due to collision; R is a random number uniformly distributed in the [0, 1] range, and $\hbar\omega_{\max}$ is the greatest possible energy loss given by

$$\hbar\omega_{\max} = \min(E, E - E_F),$$

where E_F is the Fermi level.

Using expression (7) and changing the order of integration, one can reduce the integrals in Eq. (12) to the form

$$G_1(E, Q) = \frac{1}{2\pi a_0 E} \int_0^{A(E, Q)} \mathrm{Im}\left(\frac{-1}{\varepsilon(x)}\right) \ln\left[\frac{E(Q)(E_1 + x)}{E_1(E(Q) + x)}\right] dx,$$

where

$$E(Q) = \min(Q\hbar\omega_{\max} - x, (\sqrt{E} + \sqrt{E - 2x})^2/4),$$

$$E_1 = (\sqrt{E} - \sqrt{E - 2x})^2/4,$$

$$A(E, Q) = \min(Q\hbar\omega_{\max}, E/2).$$

Note that, in the region where relationship (8) is valid, the function G_1 can be represented as

$$G_1(E, Q) = C(E)S(Q\hbar\omega_{\max}), \quad (13)$$

where $S(x)$ is the primitive function of $W(x)$.

The characterization of inelastic scattering with the use of Eq. (12) is inefficient, because the transcendental equation must be solved at each step of the numerical simulation of the trajectory. Therefore, there is reason in considering the function $Q(E, R)$, which is implicitly defined by relationship (12).

First, let us elucidate some properties of $Q(E, R)$ that follow from asymptotics (8) and (9) of the inelastic scattering differential cross section. When relationship (13) is true, the E-dependence of Q only changes the scale

$$Q = \frac{S^{-1}(RS(\hbar\omega_{\max}))}{\hbar\omega_{\max}}, \quad (14)$$

and is independent of $C(E)$ from Eq. (8). Using (8) and assuming that Q and Q_0 are taken from the region

Approximation parameters for the differential cross section of electron–target inelastic interaction

Al		Ge		Au	
$n = 1$; $m = 4$		$n = 2$; $m = 5$		$n = 1$; $m = 4$	
$d_1 = 14$; $r_1 = -3.3$		$d_1 = 0.5$; $d_2 = 18$		$d_1 = 62$; $r_1 = -2.95$	
$w_1 = 1.4$, $w_2 = 14.8$		$r_1 = 1.55$, $r_2 = -3.3$		$w_1 = 1$, $w_2 = 2.3$	
$w_3 = 72$, $w_4 = 160$		$w_1 = 0.68$, $w_2 = 2$		$w_3 = 50$, $w_4 = 140$	
$\alpha_1 = 155$, $\alpha_2 = 55$		$w_3 = 4$, $w_4 = 15$		$\alpha_1 = 1$, $\alpha_2 = 14$	
$\alpha_3 = 5$, $\alpha_4 = 2$		$w_5 = 200$		$\alpha_3 = 1.5$, $\alpha_4 = 9$	
$\beta_1 = 25$, $\beta_2 = 15$		$\alpha_1 = 1.7$, $\alpha_2 = 3$		$\beta_1 = 100$, $\beta_2 = 1$	
$\beta_3 = 50$, $\beta_4 = 4$		$\alpha_3 = 8$, $\alpha_4 = 3.1$		$\beta_3 = 2$, $\beta_4 = 5$	
$\gamma_1 = 0.4$, $\gamma_2 = 1.85$		$\alpha_5 = 2.5$		$\gamma_1 = 0.14$, $\gamma_2 = 1.6$	
$\gamma_3 = -0.5$, $\gamma_4 = 0.8$		$\beta_1 = 18$, $\beta_2 = 10$		$\gamma_3 = 3.8$, $\gamma_4 = -0.4$	
		$\beta_3 = 4$, $\beta_4 = 4.55$			
		$\beta_5 = 3$			
		$\gamma_1 = -4.2$, $\gamma_2 = -0.8$			
		$\gamma_3 = -1$, $\gamma_4 = 1.6$			
		$\gamma_5 = 0.6$			
E, eV	$C(E)$, (eV cm)$^{-1}$	E, eV	$C(E)$, (eV cm)$^{-1}$	E, eV	$C(E)$, (eV cm)$^{-1}$
2000	52000	2000	0.35	2000	55.5
4000	29000	4000	0.2	4000	31.6
6000	21000	6000	0.145	6000	22.5
8000	16000	8000	0.115	8000	17.6
10000	12900	10000	0.093	10000	14.5
12000	10700	12000	0.08	12000	12.4
14000	9400	14000	0.07	14000	10.9
16000	8400	16000	0.06	16000	9.6
20000	6700	20000	0.05	20000	7.9
25000	5400	25000	0.042	25000	6.5
30000	4500	30000	0.034	30000	5.5

where the asymptotics of the function W is true, we will obtain

$$G_1(E, Q) = G_1(E, Q_0) + \int_{Q_0\hbar\omega_{\max}}^{Q\hbar\omega_{\max}} \frac{C(E)}{(\hbar\omega)^2}\, d\hbar\omega, \quad (15)$$

hence,

$$Q(E, R) = \frac{C_1}{C_2 - R}. \quad (16)$$

Here,

$$C_1 = \frac{C(E)}{G_1(E, 1)\hbar\omega_{\max}},$$

$$C_2 = \frac{G_1(E, Q_0)}{G_1(E, 1)} + \frac{C(E)}{G_1(E, 1)Q_0\hbar\omega_{\max}}.$$

It should be noted that relationship (16) is valid for $R > G_1(E, Q_0)/G_1(E, 1)$, because the integral in Eq. (15) is positive. We succeeded in approximating the function $Q(E, R)$ by the five-parameter analytical expression

$$Q(E, R) = F(a(E), b(E), g(E), p(E), m(E), R),$$

where a, b, g, and p are piecewise continuous functions of E and the parameter m is a piecewise constant function that defines the general shape of the curve $F(R)$ (hereafter the parameters a, b, g, p, and m are omitted in the list of arguments).

The function $F(R)$ is given by

$$F(R) = Z(R)/Z(1),$$

where

$$Z(R) = (G(R) + H(R) - H(0))\arctan(5000cR),$$

$$G(R) = b\left[1 + \left(\frac{cR}{g}\right)^m\right]^{-\frac{p}{m}}(cR)^p,$$

$$H(R) = \left[\left(\frac{1}{1-g}\right)^{|m|} + \left(\frac{1}{1-cR}\right)^{|m|}\right]^{\frac{1}{|m|}},$$

$$c = \frac{1}{1+a}. \qquad (17)$$

It is easily seen that, as $cR \longrightarrow 1$, the term $H(R)$ is of first importance in the expression for $Z(R)$; in this case, $F(R)$ tends to the form of (16). If m is positive, this corresponds to $R > g/c$: at $R < g/c$, the function $F(R)$ is a power dependence. The larger m, the more abrupt the transition from one portion of $F(R)$ to another. At $m < 0$, the function $E(R)$ approaches a constant for low values of $R < g/c$; at $R > g/c$, both terms $H(R)$ and $G(R)$ contribute to the function $Z(R)$. When selecting the form of function (17), we take into account Eq. (14), from which it follows that a scale factor should appear in the approximating function. The quantity c is the scale factor in expression (17). A comparison of expressions (14) and (17) shows that, at initial electron energies $E > E_0$ (where E_0 is about several keV), energy dependence of the approximating function is expressed by the function $a(E)$; the remaining approximation parameters remain constant within this energy range.

At small energies comparable to those of plasmon excitation, the shape of $Q(E, R)$ is governed by the complex structure of differential cross section (7) in the range of plasmon peaks and relationships (8) and (9) fail; therefore, the expression for the approximating function $F_0(R)$ differs from (17):

$$F_0(R) = 1 - Y(1-R)^{g(1-Y(1-R))},$$

$$Y(R) = R\left(a + \frac{1-a}{1-p}\frac{1-(1-R^b)^p - pR^b}{R^b}\right). \qquad (18)$$

The function $F_0(R)$ readily changes its form in accordance with the values of the parameters.

Thus, by approximating relationships (17) and (18), the energy losses per inelastic collision are calculated from the simple analytic formula. Complete information necessary to calculate the energy losses per inelastic collision is contained in the five one-dimensional data arrays $a(E)$, $b(E)$, $g(E)$, $p(E)$, and $m(E)$, and in the only data array $a(E)$ at $E > E_0$.

We are coming now to the calculation of the scattering angle in an inelastic interaction. A change in the direction of electron motion due to inelastic collision is determined from the transcendental equation

$$R = G_2(E, \hbar\omega, \theta)/G_2(E, \hbar\omega, \theta_{max}), \qquad (19)$$

where R is, as before, a random number uniformly distributed over the [0, 1] range and $G_2(E, \hbar\omega, \theta)$ is the integral

$$G_2(E, \hbar\omega, \theta) = \int_{B(E,\hbar\omega,\theta)}^{B(E,\hbar\omega,0)} \frac{x}{\hbar\omega - x}\mathrm{Im}\left(\frac{-1}{\varepsilon(x)}\right)dx,$$

$$B(E, \hbar\omega, \theta) = 2(\hbar\omega - E + \cos\theta\sqrt{E(E-\hbar\omega)}).$$

The greatest possible scattering angle θ_{max} is found from the equation

$$B(E, \hbar\omega, \theta_{max}) = 0.$$

Finding the scattering angle from relationship (19) requires knowledge of the three-dimensional array of the values of the integral $G_2(E, \hbar\omega, \theta)$. However, $G_2(E, \hbar\omega, \theta)$ depends on E and θ only through the limits of integration; therefore, the data array necessary to determine the scattering angle can be significantly shortened. In fact, entering the function

$$G(\hbar\omega, A) = \int_0^{D(\hbar\omega,A)} \frac{x}{\hbar\omega - x}\mathrm{Im}\left(\frac{-1}{\varepsilon(x)}\right)dx,$$

where

$$D(\hbar\omega, A) = B(A, \hbar\omega, 0) \qquad (20)$$
$$= 2(\hbar\omega - A + \sqrt{A(A-\hbar\omega)}),$$

and the parameter A is related to E and θ by the relationship

$$\hbar\omega - E + \cos\theta - \sqrt{E(E-\hbar\omega)} = \hbar\omega - A + \sqrt{A(A-\hbar\omega)},$$

one readily comes to

$$G_2(E, \hbar\omega, \theta) = G(\hbar\omega, A(E, 0)) - G(\hbar\omega, A(E, \theta)).$$

In numerical simulation of inelastic collisions, the values of $G(\hbar\omega, A)$ for arbitrary $\hbar\omega$ and A are obtained by interpolation. We can estimate how an interpolation error affects the determination accuracy of the scattering angle. Differentiating Eq. (20) with respect to θ gives

$$d\theta = \frac{dG}{G'_D\sqrt{E(E-\hbar\omega)}\sin\theta},$$

$$G'_D = \frac{D(\hbar\omega, A)}{\hbar\omega - D(\hbar\omega, A)}\mathrm{Im}\left(-\frac{1}{\varepsilon(D(\hbar\omega, A))}\right).$$

It follows that particular emphasis on the interpolation accuracy should be placed at high energies, where $G'_D \sim \mathrm{Im}(-1/\varepsilon(\hbar\omega)) \sim 10^{-6}$–$10^{-5}$ (cm eV)$^{-1}$. Energies of this range significantly exceed the excitation energies of plasmons and valence-band electrons; therefore, the

integrand varies smoothly everywhere, exclusive of a finite number of points corresponding to the excitation of inner electron shells. At these points, the integrand in Eq. (20) is discontinuous. To improve the accuracy of interpolation in the neighborhood of the discontinuity points, instead of using the function $G(\hbar\omega, A)$, we examine the integral of the smooth function

$$\tilde{G}(\hbar\omega, A) = \int_0^{D(\hbar\omega, A)} \frac{xP(x)}{\hbar\omega - x} dx,$$

where

$$P(x) = \begin{cases} \mathrm{Im}(-1/\varepsilon(x)), & x > x_n, \\ \mathrm{Im}(-1/\varepsilon(x)) + \sum_{i=k}^{n} U_i, & x_{k-1} < x < x_k, \end{cases}$$

x_k are the discontinuity points of the integrand, and

$$U_i = \mathrm{Im}(-1/\varepsilon(x_i + 0)) - \mathrm{Im}(-1/\varepsilon(x_i - 0)).$$

The values of $G(\hbar\omega, A)$ are easily obtained from $\tilde{G}(\hbar\omega, A)$:

$$G(\hbar\omega, A) = \tilde{G}(\hbar\omega, A) - \sum_{i=1}^{k} \int_0^{x_i} \frac{U_i x}{\hbar\omega - x} dx$$

$$- \int_0^{A} \left(\sum_{i=k+1}^{n} U_i \right) \frac{x}{\hbar\omega - x} dx,$$

$$x_k < D(\hbar\omega, A) < x_{k+1}.$$

This procedure reduces the errors in the determination of the angle of electron inelastic scattering in the neighborhood of an integrand jump from 0.2 to 0.05 rad.

Finally, we also estimated the relative errors in determining the average energy loss per unit length δ_β, which is of great importance in the scattering process. For the Al target, at energies of 0.1–15 keV, $\delta_\beta < 0.015$ and at energies of 15–30 keV, $\delta_\beta < 0.025$; for the Au target, at energies of 0.1–30 keV, $\delta_\beta < 0.015$; and for Ge, at 0.1–15 keV, $\delta_\beta < 0.02$ and, at 15–30 keV, $\delta_\beta < 0.04$. The differential scattering cross sections calculated with formula (7), as well as the tables of the parameters of the approximating function for the function $Q(E, R)$, are available from the web site http://www.ioffe.rssi.ru/ES.

Thus, we elaborated a resource-saving algorithm that makes it possible to simulate electron–target inelastic interaction with the use of the DD cross section. The cross section was calculated within the approach presented in [1–4] by using data on the optical permittivity of a target.

ACKNOWLEDGMENTS

This work was performed within the Program "Surface Atomic Structures" (project no. 99-1.18) supported by the Ministry of Science of the Russian Federation, and also within Agreement KRISS/PTI 98-01.

REFERENCES

1. D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963; Mir, Moscow, 1965).
2. D. R. Penn, Phys. Rev. B **35**, 482 (1987).
3. Z.-J. Ding and R. Shimizu, Surf. Sci. **222**, 313 (1989).
4. Z.-J. Ding, *Fundamental Studies on the Interactions of kV Electrons with Solids for Applications to Electron Spectroscopies*, PhD Thesis (Osaka Univ., 1990).
5. S. Tanuma, C. J. Powell, and D. R. Penn, Surf. Interface Anal. **11**, 577 (1988).
6. W. Brandt and J. Reinheimer, Phys. Rev. B **26**, 3104 (1982).
7. B. I. Lundqvist, Phys. Condens. Mater. **6**, 206 (1967).
8. C. M. Kwei and C. J. Tung, J. Phys. D **19**, 255 (1986).
9. S. Tougaard, Surf. Interface Anal. **11**, 453 (1988).
10. R. K. Peterkops, *Theory of Atom Ionization by Electron Impact* (Zinatne, Riga, 1975; Colorado Associated Univ. Press, Boulder, 1977).

Translated by B. Malyukov

Technical Physics, Vol. 45, No. 12, 2000, pp. 1524–1528. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 12–17.

ATOMS, SPECTRA, AND RADIATION

Geometrical Shape Resonances in Multicenter Systems with Icosahedral Symmetry: A Controllable Molecular Trap for Electrons

Yu. F. Migal' and V. S. Kovaleva

Donskoĭ State Technical University, Rostov-on-Don, 344010 Russia;
e-mail: root@sintez.rnd.su

Received June 3, 1999

Abstract—Specific shape resonances were studied in systems with the highest point symmetry. The resonances occur in narrow energy intervals and disappear when inner potentials increase. A model system of point scatterers that are placed in the vertexes of an icosahedron and at its center was analyzed. In this system, the geometric resonance lifetime is extremely sensitive to structure changes and may attain 10^{-10} s, i.e., exceed the rotation period of electrons in atoms by six orders of magnitude. The feasibility of molecular traps that capture elastically scattered electrons and bring them into the resonant state is discussed. The resident time of electrons in these traps can be controlled with external static fields. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

One intriguing feature of elastic scattering of electrons by molecules is distinct maxima of the cross section that are observed when incident electrons are trapped and brought into a one-electron quasi-stationary state. This effect has been called shape resonances (see, e.g., [1]). Recently, a new type of shape resonances that offer a number of specific properties has been found [2]. These shape resonances are observed in narrow energy intervals that depend on the geometry of a many-atom system. Unlike conventional (hybrid) resonances (see below), whose spectrum gradually becomes discrete with increasing potentials, geometric resonances disappear under such conditions. The latter occur only in high-symmetry systems, for example, in a system with eight point scatterers located at cube vertexes [2]. The higher the symmetry, the more pronounced the resonances. In view of the recent discovery of new high-symmetry compounds (specifically, carbon compounds with the icosahedral symmetry), the behavior of geometric resonances in systems with the icosahedral symmetry seems to be of interest. One could expect that the lifetime of the resonance states in such objects is a record.

Since geometric resonances occur in high-symmetry systems and only in narrow energy intervals, even minor structure modifications may cause their breakdown. This leads researchers to look for systems where the resonance conditions are created or violated by external effects. It is known that even the highest electric and magnetic fields that have been produced to date by advanced laboratory equipment are weaker than inner atomic fields; hence, external static fields cannot significantly affect the electron states in atoms. The superhigh sensitivity of geometric resonances to structure changes could be used in nanotechnology for controlling atomic-scale processes by means of external fields.

Below, we will try to construct a multicenter system with the lifetime of the resonance states as large as possible. In addition, the feasibility of molecular traps for elastically scattered electrons in which the electron resident time can be controlled with external fields is discussed. First, however, we will consider general issues concerning the problem of geometric shape resonances.

EXISTENCE CONDITIONS OF GEOMETRIC RESONANCES

It was shown [3] that the primary reason for the appearance of collective shape resonances in a multicenter system is destructive interference of waves originating at individual centers. As a result, an extended region arises around the system. In this region, the wave function of an electron is suppressed. This region can be simulated by a barrier surrounding the system. For highly symmetric small systems, this barrier is of a centrifugal type.

To quantitatively characterize the effect, consider a solution to the Schrödinger equation for a multicenter system each center of which is a wave source (many-source representation [4]):

$$\Psi_\lambda^+(k, \mathbf{r}) = \Sigma_{jL} B_{\lambda L}^{(j)+}(k) h_l^+(k|\mathbf{r} - \mathbf{r}_j|) Y_L(\widehat{\mathbf{r} - \mathbf{r}_j}). \quad (1)$$

Here, $k = \sqrt{E}$, $L = (l, m)$ is the combined quantum number, $\mathbf{r}_j$ are the radius vectors of the centers, h_l^+ are Hankel spherical functions, and Y_L are spherical harmonics.

1063-7842/00/4512-1524$20.00

The complex coefficients $B_{\lambda L}^{(j)+}$ define the amplitude and phase characteristics of the sources, and the subscript λ numbers orthonormal solutions that transform with respect to irreducible representations of the point symmetry group.

Note that expression (1) is valid only outside the system, where, in addition, the potential of electron–system interaction is assumed to be zero. This assumption affects quasi-localized states only slightly, since their density is, as a rule, high inside the system and quickly drops with distance from the system. It should also be emphasized that expression (1) is not an electron wave function, because it has singularities at the source sites. An ordinary wave function can be obtained if the imaginary part of (1), containing all the singularities, is rejected.

Let us expand expression (1) in terms of spherical harmonics about the center of point symmetry of the system:

$$\Psi_\lambda^+(k, \mathbf{r}) = \Sigma_L D_{\lambda L}^+(k) h_l^+(kr) Y_L(\Omega). \quad (2)$$

It is important that, at small k's, when the wavelength far exceeds the linear dimensions of the system, only the term with the least l (for a given irreducible representation), $l_{\min}$, can be left in (2). Then, the scattering process can be described in terms of spherically symmetric objects. Specifically, one can use the concept of a centrifugal barrier. In a system surrounded by a barrier with $l_{\min}$, standing waves with $E > 0$, which are virtually resonant states, can arise provided that the potential penetrates deeply into the barrier. This statement holds for both spherically symmetric objects and multicenter systems.

However, scattering by multicenter systems has a number of specific features. For example, the first coefficient $D_{\lambda L}^+(k)$ with $l = l_{\min}$ [see (2)] may be equal or close to zero in some energy interval ΔE. Then, the terms with the next value of l, $l = l_1$ (which is the least among the remaining values of l), will dominate in (2). In this case, the system will be surrounded by a centrifugal barrier with $l_1 > l_{\min}$, which is higher than that with $l_{\min}$. If the system potential is such that the system has a resonance state with $E_{\text{res}} \in \Delta E$, the lifetime of this resonance depends on the penetrability of the barrier with l_1 [3]. Now assume that we can increase the potential, thus decreasing the energy E_{res}. At a sufficiently high potential, the value E_{res} may go beyond the interval ΔE and the term with $l_{\min}$ will prevail in expansion (2). In this state, the system is surrounded by the lower barrier with $l_{\min}$; hence, the lifetime of the resonance decreases. If $l_{\min} = 0$, the barrier is absent and the resonance totally breaks down as the potential grows. Such resonances are not typical of spherically symmetric objects and are called geometric resonances [2] (the meaning of this name will be explained below).

The simplest illustration of a geometric resonance is that in a system of eight identical point scatterers (PS) placed at the vertexes of a cube (the O_h group of symmetry). An isolated PS is known to be uniquely characterized by the energy parameter α [5]. For $\alpha > 0$, the PS has a single bound s state with an energy $-\alpha^2/2$. The PS may emit only s waves. Therefore, expression (1) for a system of PS contains only the terms with $l = 0$. If the solution is sought for a_{1g} symmetry (only this solution will be considered in what follows), when all the sources emit waves of equal amplitudes and phases, expression (1) can be represented as

$$\Psi^+ = b \exp(i\eta) \Sigma_j h_0^+(k|\mathbf{r} - \mathbf{r}_j|). \quad (3)$$

In the one-center case,

$$\Psi^+ = b \exp(i\eta) \{ j_0(kR) h_0^+(kr) + j_4(kR) h_4^+(kr) Y_{4m}(\Omega) C + \dots \}, \quad r > R, \quad (4)$$

where R is the radius of a sphere encompassing the PS and j_l are spherical Bessel functions. If $k \approx \pi/R$, $j_0(kR) \approx 0$ and the first term in (4) is small. In this case, the term with $l = 4$ is dominant and the system is surrounded by a centrifugal barrier of height $V_{\max} = l(l+1)/R^2 = 20/R^2$. For α's close to unity, the system may have a resonant state with an energy of about $(\pi/R)^2$, which is much lower than $V_{\max}$. This is a typical geometric resonance, which exists only in an energy range including the point $(\pi/R)^2$. It quickly breaks down when α increases and E_{res} goes beyond this range. As follows from this example, the resonance energy depends on the geometric factor R; hence, the name "geometric resonance."

It is essential that this resonance exists when the second term in (4) has a large l ($l = 4$). Owing to this, the resonance energy is much lower than the barrier height and the residence time of an electron in the system is large. For low-symmetry systems, to the second term in the solution Ψ^+, which transforms with respect to identical irreducible representation, there corresponds a small l (1 or 2) and the energy $(\pi/R)^2$ exceeds the height of the centrifugal barrier with this l. In such a system, geometric resonances are absent.

More emphatically, a potential dependence of a geometric resonance is described by the trajectory of the pole of the corresponding S matrix in the complex plane E or k. It is well known (see, e.g., [6]) that the real coordinate of the S-matrix pole in the fourth quadrant of the plane E equals the resonance energy E_{res}, while the imaginary one is equal to the half-width Γ of the corresponding maximum of the elastic-scattering cross section. In the case of a geometric resonance, the pole approaches the real axis as the potential grows and then moves away from this axis (the S-matrix pole trajectory for a system of eight PS is shown in Fig. 2). For a hybrid resonance, however, the pole usually monotonically tends to the real axis. At small k's, the value of Γ

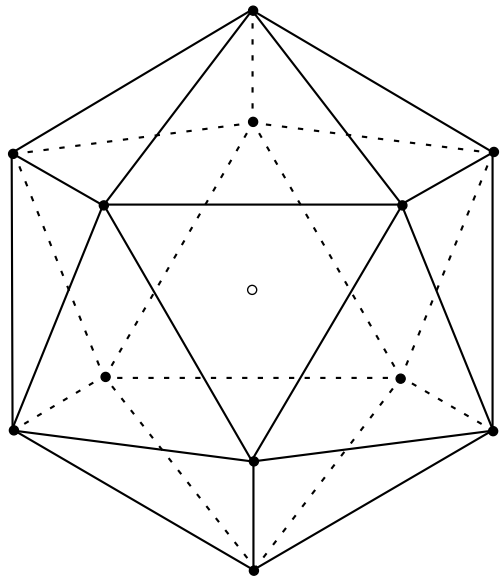


Fig. 1. System of point scatterers located at the vertexes of the icosahedron and at its center.

for hybrid resonances is proportional to $(\mathrm{Re}k)^{2l+1}$ with $l = l_{\min}$. For geometric resonances, such an estimate of Γ (with l corresponding to the resonance) is valid only in a narrow energy range ΔE.

It is worth stressing that geometric resonances arise not only in PS models. They have also been observed in real multiatomic systems, such as $LiBiS_2$ and $NaBiS_2$ compounds [7].

To conclude this section, we will briefly mention other features of geometric resonances. They have higher energies and larger quantum number l than hybrid resonances. Unlike the latter, which, in small systems, can be predicted in terms of the MO LCAO minimum-basis approach, geometric resonances cannot be predicted in such a way. This, in particular, means that, if each of the centers in a small system has its own localized (or quasi-localized) state and the total number of the states is N, then the system has N localized and quasi-localized states, exclusive of geometrical resonances (localized or quasi-localized states outside the system are not considered). Thus, geometrical resonances arise only in combined high-symmetry systems. Their presence means that the combined potential of even a small system is in some respects stronger than the sum of the potentials of isolated centers, since it is capable of holding a larger number of physically separated states.

GEOMETRIC RESONANCES IN AN ICOSAHEDRAL-SYMMETRY SYSTEM OF POINT SCATTERERS

Consider a PS system that has the highest (icosahedral) point symmetry. For the representation Y_h of the a_{lg} icosahedron group, the sequence of the l values is 0, 6, 8, … . Therefore, a multicenter system with this symmetry may have a distinct geometric resonance at $l = 6$.

Let 12 identical PS be located at the vertexes of an icosahedron. One more PS is placed at its center (Fig. 1) in order to improve the controllability of the scattering power of the system. Let the energy parameter of the central scatterer be α_0 and those at the periphery be α_1. We will seek the poles of the S matrix that are related to the a_{lg} solution. To do this, we will take advantage of the general equation for the S matrix poles in the muffin–tin approximation [4]:

$$\det\left\| [1 + i\cot\delta_l^{(j)}]\delta_{jj'}\delta_{LL'} + \Sigma_{L''} H_{LL'L''}^{j,j'} h_{l''}^{+}(k|\mathbf{r}_j - \mathbf{r}_{j'}|) \right\| = 0, \quad (5)$$

where $\delta_l^{(j)}$ is the phase shift on the jth atomic sphere (complex for complex k's) and $H_{LL'L''}^{j,j'} = 4\pi i^{l-l'-l''} Y_{L''}(\widehat{|\mathbf{r}_j - \mathbf{r}_{j'}|}) \int Y_L Y_{L'} Y_{L''} d\Omega$ are k-independent real structure constants.

For the case of point scatterers, only the terms with $l = 0$ and $l' = 0$ are left. All other nonzero constants are unities. The values of $\cot\delta_l^{(j)}$ are replaced by $-\alpha_j/k$ (j = 0, 1).As a result, Eq. (5) takes the form

$$D_0 D_1 - 12[h_0(kr_4)]^2 = 0, \quad (6)$$

where $D_0 = (1 - i\alpha_0/k)$; $D_1 = (1 - i\alpha_1/k + 5h_0^+(kr_1) + 5h_0^+(kr_2) + h_0^+(kr_3))$; $r_1 = d$, $r_2 = 1.618033989d$, $r_3 = 1.902113033d$, and $r_4 = r_3/2$ are distances between the PS in the system; and d is the icosahedron edge (in the calculations, d was taken to be equal to 1 a.u., which coincides with internuclear spacings in real multiatomic systems by order of magnitude).

First, we will find the poles for the system without the central PS. In this case, the function Ψ^+ has the same form as (4) for an eight-PS system, but the subscript 4 here is replaced by 6. Eq. (6) is reduced to $D_1 = 0$. Results are given in Fig. 2. Curve *2* describes the pole behavior for the icosahedral system when α_1 varies in wide limits (each point of the curve corresponds to a certain value of α_1). This curve approaches the real axis in the narrow range of Rek. At the point of closest approach, the parameters are the following: Rek = 3.300, Im$k = -1.204 \times 10^{-3}$, and $\alpha_1 = 2.597$. As follows from the theory (see the previous section), for a system of PS located on a sphere of radius R, the energy of geometric resonance must correlate with the first root of the equation $j_0(\mathrm{Re}kR) = 0$. This root equals π. In our case, $R = r_4 = 0.9511$ a.u.; hence, Rek = 3.303, which virtually coincides with the coordinate of the point where curve *2* is at the minimum distance from the real axis.

For comparison, Fig. 2 shows the trajectory of the S matrix pole for a system of eight PS arranged at the vertexes of a cube with an edge $d = 1$ a.u. (curve *1*). In this case, at the point of closest approach, Rek = 3.522, Imk = –0.05917, and α_1 = 1.653. The radius of the sphere is 0.8660 a.u. From the condition Re$kR = \pi$, we find Rek = 3.628. This value is also close to the coordinate of the point of closest approach.

Obviously, the pole trajectory for the icosahedron approaches the real axis much closer than in the case of the cube. From these data, it is easy to estimate the duration of the resonance state in the system: $\tau \sim \hbar/\Gamma$, where $\Gamma = 2(\mathrm{Re}k)(\mathrm{Im}k)$. It turns out that the maximum duration of the geometric resonance state in the icosahedron is 52 times longer than in the cube.

Now let us try to extend the residence time of an electron by varying the potential of the additional PS at the center of the icosahedron. The central PS causes an additional term with $l = 0$ to appear in Eq. (4). As a result, the energy of geometric resonance does not satisfy the simple relationship $j_0(\mathrm{Re}kR) = 0$ any longer and becomes a complex function of R, α_0, and α_1.

As follows from Eq. (6), when α_0 increases from a large negative value to a small positive one, the pole trajectory shifts from the right to the left, moving parallel to itself, and approaches the real axis. Curve *3* in Fig. 2 is plotted for $a_0 = 0$. At the point of closest approach to the real axis, Rek = 1.6517, Im$k = -1.484 \times 10^{-7}$, and α_1 = 3.6855. In this case, the residence time is 16200 times greater than that for the icosahedron without the central PS. Its absolute value is as high as 10^{-10} s, which is a giant time for any atomic process (for the cube, $\tau \sim 10^{-15}$ s, which is only one order of magnitude larger than the period of rotation of valence electrons in an atom). Thus, by varying the parameter of the central PS, one can alter the resonance energy and substantially increase the residence time of an electron.

In the foregoing, we considered geometric resonances in the system of point scatterers, which represent the simplest model of real multiatomic systems. However, the discussed effect of long electron residence in a system primarily reflects the fact of system symmetry; hence, the results obtained for the PS model remain valid also for more realistic models. (Investigations into systems of three-dimensional scatterers that were performed in the muffin–tin approximation support the above statement.)

A CONTROLLABLE TRAP FOR ELECTRONS

Since an icosahedral system is unique in terms of electron residence time, it is of interest to consider its possible applications. We have to assume that the construction of a stable cluster composed of real atoms with an internuclear spacing of about 1 a.u. (similarly to the considered system) is a possibility. In this case, the atoms should be arranged so that the resonance energy falls into the range where the value of D_L^+ with $l = 0$ in (2) is minimum. Recent successful attempts to construct various exotic systems suggest that such a system can also be created. Assume also that the temperature of the system is low, so that thermal vibrations of the atoms cannot significantly distort the icosahedral structure. Then, such a system could be used as a trap for free electrons with a kinetic energy equal to the geometric resonance energy.

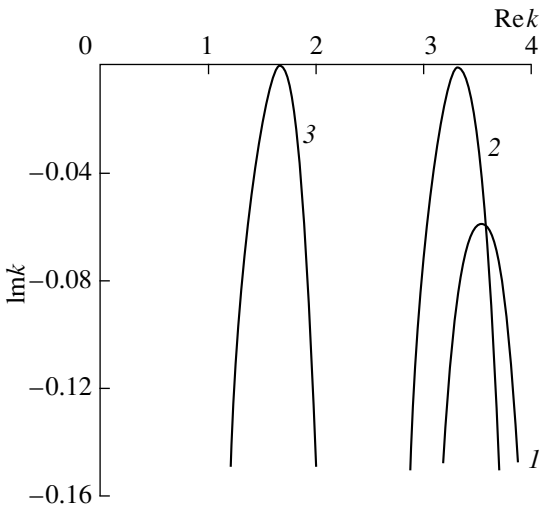


Fig. 2. S matrix pole trajectories corresponding to geometrical resonances in the PS systems: (*1*) eight PS at the vertexes of the cube with an edge length $d = 1$ a.u., (*2*) twelve PS at the vertexes of the icosahedron with an edge $d = 1$ a.u., and (*3*) thirteen PS at the vertexes and at the center of the icosahedron with an edge $d = 1$ a.u. ($\alpha_0 = 0$).

In this work, the trap concept alone is put forward without considering many issues that may be of importance in a specific implementation. The most essential of them are certainly those concerned with the effect of inelastic processes on electron scattering. It is not inconceivable that this effect will be significant in actual systems and the trap cannot be implemented. The orientation of spins in the system and in incident electrons, as well as some other problems, are also of concern. They all call for further detailed investigation and are beyond the scope of this article. Therefore, having assumed that a stable cluster of a desired configuration is feasible, we will make the next step forward, i.e., assume that all the above factors cannot deteriorate the holding capacity of a trap in a real system.

Now consider how one can vary the residence time of an electron by applying an external field. Here, two situations may arise: (1) uniform compression (or extension), when the radius of the scatterer changes or (2) a shift of the central atom from its equilibrium state, for example, under the action of a strong electrostatic field. Since the effective charge of the central atom and those of the peripheral atoms are unlike (it is hardly probable that they equal zero), the external field must displace the central one from the middle point.

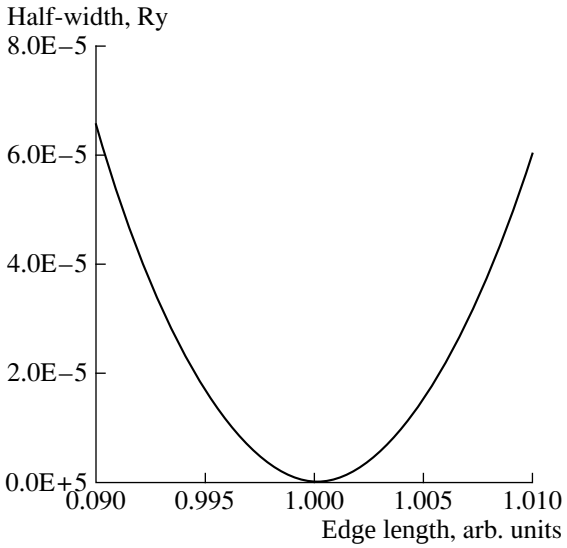


Fig. 3. Half-width Γ of the resonance peak vs. edge length *d* of the icosahedron for the PS system with $\alpha_0 = 0$ and $\alpha_1 = 3.6855$.

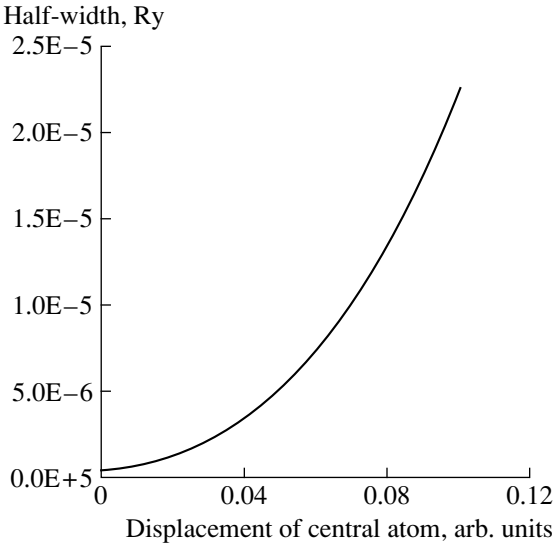


Fig. 4. Half-width Γ of the resonance peak vs. central atom offset from the middle point for the PS system with $d = 1$ a.u., $\alpha_0 = 0$, and $\alpha_1 = 3.6855$.

To study these situations, it is appropriate to see how Γ depends on the variable parameters of our model system, which consists of PS located at the center and vertexes of the icosahedron. For situation 1 (uniform compression), let us evaluate a Γ vs. icosahedron edge *d* dependence. If the potential parameters (α_0 and α_1) remain fixed, Γ varies as shown in Fig. 3. Here, α_0 and α_1 are chosen such that Γ takes the minimum value at $d = 1$ a.u. ("undisturbed" state; see previous section). Obviously, the greater the disturbance, the greater Γ and, hence, the smaller the residence time τ of an electron in the system. If, for example, *d* is increased by 0.5%, τ decreases 32-fold.

For situation 2 (the central atom is shifted from the middle point by *s*), Eq. (6) becomes invalid, since the 12 atoms at the vertexes are no longer identical and the icosahedral symmetry breaks. One has to solve a more intricate equation [derived from general Eq. (5)] whose complexity depends on the offset direction. Calculations for the offset *s* of the central atom toward one of the vertexes are shown in Fig. 4. It is seen that an increase in Γ and a decrease in τ are significant when *s* is small. When the two variables (offset *s* in situation 2 and a change in *d* in situation 1) are equal, τ changes to a lesser extent in situation 2. It is clear, however, that to shift the central atom by some distance is actually easier than to change *d* by the same value.

Note also that, in the presence of an external field (even if the system remains undisturbed), the symmetry of a state trapping an electron is reduced (inverse operation is eliminated), which also cuts the residence time of the electron.

CONCLUSION

A system of point scatterers that are placed at the center and vertexes of an icosahedron offer unique properties which show up at resonant scattering. The theoretically predicted residence time of an electron in this system was found to be 10^{-10} s even if the kinetic energy of an incident electron is relatively high (10–30 eV). This effect is due to the small dimensions of the system and a large value of the orbital quantum number *l*, which characterizes the resonance state outside the system. We believe that such systems can be used in nanotechnological projects.

REFERENCES

1. S. Trajmar, D. F. Register, and A. Chutjian, Phys. Rep. **97** (5), 219 (1983).
2. Yu. F. Migal, J. Phys. B **24** (19), 4181 (1991).
3. Yu. F. Migal, J. Phys. B **25** (18), 3849 (1992).
4. Yu. F. Migal, J. Phys. B **26** (17), 2755 (1993).
5. Yu. N. Demkov and V. N. Ostrovskii, *Zero-Range Potentials and Their Applications in Atomic Physics* (Leningr. Gos. Univ., Leningrad, 1975; Plenum, New York, 1988).
6. J. R. Taylor, *Scattering Theory* (Wiley, New York, 1972; Mir, Moscow, 1975).
7. A. A. Lavrent'ev, Yu. F. Migal', and I. Ya. Nikiforov, Zh. Strukt. Khim. **33** (2), 60 (1992).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 12, 2000, pp. 1529–1537. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 18–27.

GASES AND LIQUIDS

Impact of a High-Velocity Drop on an Obstacle

A. V. Chizhov and A. A. Schmidt

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: Anton.Chizhov@pop.ioffe.rssi.ru

Received March 1, 2000

Abstract—Processes arising when a high-velocity liquid drop strikes a rigid obstacle or a liquid layer were investigated using numerical simulation. The flow pattern being formed features a complicated interaction of compression shock and expansion waves between each other and with free surfaces, the initiation of a cumulative jet flow, and the formation of cavitation areas. Factors governing the interaction process are analyzed. Obtained results are compared with experimental data. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

The impact of a liquid drop on an obstacle at a velocity comparable to the sound speed in the liquid generates shock and expansion waves interacting with each other and with free surfaces, the evolution of a cumulative jet, and the formation and collapse of cavitation bubbles. Therefore, the phenomenon of high-velocity drop impact is closely related to many fundamental problems in mechanics of continua and physics of strength. This phenomenon is also of interest in applied problems, such as the erosion action of liquid–gas streams, coating application, mining, working of hard materials, cleaning of surfaces, production of new materials, astrophysics problems, etc. [1–6].

Comprehensive reviews on impacts of drops on obstacles can be found in [1, 5, 7–9]. However, the phenomenon of high velocity has not been adequately studied even qualitatively. The objective of this paper is to describe the detailed pattern of a drop impact on an absolutely rigid surface and on a thin liquid layer under conditions at which liquid compressibility shows up most vividly, that is, at impact velocities comparable to the sound speed in the liquid. Emphasis will be on the effect of liquid viscosity and surface tension, interaction between impact products and ambient gas, mechanism of initiation and disruption of the cumulative jet, jet velocity evaluation, formation mechanism of cavitation cavities, loads experienced by the obstacle, differences between plane and axisymmetric impacts, change in the flow structure upon striking the liquid layer, and reasons for experimental spread and discrepancy between analytical and experimental estimates.

DROP IMPACT PATTERN

Generally, drop–obstacle interaction produces a liquid flow with a well-developed wave structure and heavily deformable free surface (Fig. 1).

One interesting feature of a convex-drop impact is that, at the initial impact stage, the free surface, which does not contact the rigid one, does not deform. The compression region of the drop is bounded by a shock wave adjacent to the contact domain boundary (see Fig. 1a). This is because the velocity of motion of the contact boundary $V_E = V_0 \cot\beta(t)$ (where V_0 is the initial velocity of the drop and $\beta(t)$ is the angle between the free surface and the rigid wall), being infinitely large at the contact instant $t = 0$, decreases, remaining greater than the shock wave velocity up to a certain instant t^c. Therefore, disturbances propagating from the wall have no time to interact with the free surface. The compression of the liquid is maximal at the contact periphery and continues to grow with time.

At the critical instant t^c, the shock wave leaves the contact boundary and interacts with the free surface to form a shock wave in the ambient gas and an expansion wave propagating inside the drop. The free surface starts to deform and a near-wall high-velocity cumulative jet is formed (Fig. 1b). The time of jet formation depends on viscous and surface effects in the liquid near the wall, and the jet velocity far exceeds the impact one.

When the shock wave in the drop approaches its top, the expansion wave, following the shock wave, causes the formation of a toroidal cavitation area whose cross section is shown in Fig. 1c. At the final stage of interaction, the expansion wave collapses at the symmetry axis and a vast cavitation area with the greatest rarefaction near the axis is formed (Fig. 1d). As the expansion wave propagates toward the rigid surface, the cavitation area occupies nearly the whole drop excepting a thin layer near the free surface and the near-wall jet region. The instabilities develop, the liquid coating breaks down, and the drop takes a shape of a crown and disintegrates into small fragments [10].

FLOW FEATURES

The basic works on this subject [11, 12] contain experimental data corroborating the qualitative impact pattern described above, in particular, strongly nonuni-

1063-7842/00/4512-1529$20.00

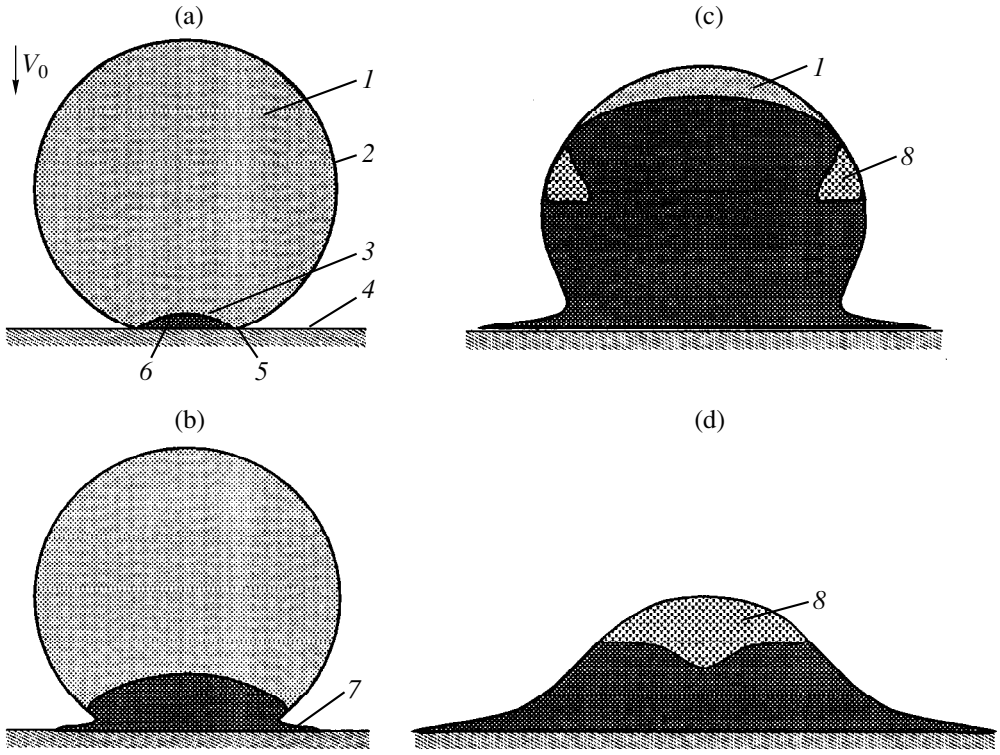


Fig. 1. Schematic of drop impact: (a) before spreading; (b) jet initiation; (c) shock wave approaches the drop top, toroidal expansion region is formed; and (d) initiation of vast expansion area with the cavitation region. (*1*) Undisturbed liquid, (*2*) free drop surface, (*3*) shock wave, (*4*) rigid surface, (*5*) contact boundary, (*6*) compressed liquid area, (*7*) jet, and (*8*) cavitation region.

form pressure distribution over the rigid surface and the shock-wave structure of the flow inside the drop. It was found that the lateral spreading starts only after the shock wave has been detached from the contact domain and has traveled along the free surface of the drop. The value of the angle $\beta(t)$ between the drop and obstacle surfaces, as well as the lateral spreading velocity, was measured at the instant of flow initiation. In the range of drop velocities from 30 to 100 m/s, this angle varied from 10 to 20° and the spreading velocity exceeded the impact velocity approximately tenfold.

There are two typical experimental conditions for studying the impact of a liquid on an obstacle: (*A*) impact on the side surface of a cylindrical layer of a waterlike gel of radius $R_0 = 5$ mm at a velocity $V_0 = 110$ m/s (Fig. 2) [12] and (*B*) impact on a spherical drop of radius $R_0 \approx 1.2$ mm at a velocity $V_0 = 660$ m/s (Fig. 3) [10].

In most practically important cases, it is assumed that (1) the ambient gas and processes occurring on the drop surface have an insignificant effect on the liquid flow inside the drop; (2) the flow in the drop outside the jet and near its outflow is inviscid; (3) viscosity influences the flow inside the jet and near its outflow, as well as the time instant the high-velocity flow appears and the jet size; and (4) surface tension forces affect the flow inside the drop and jet only slightly. The flow obeys the laws of conservation of mass and momentum for a viscous compressible liquid. The state of the liquid is described by the Tait barotropic equation [13]

$$p = B[(\rho/\rho_0)^\gamma - 1] + p_0, \qquad (1)$$

where, for water at temperatures from 20 to 60°C, it is usually accepted that $\rho_0 = 1000$ kg/m^3, $p_0 = 10^5$ Pa, $\gamma = 7$, and $B = 3.214 \times 10^8$ Pa. The sound speed in a liquid under normal conditions is defined as $a_0 = \sqrt{B\gamma/\rho_0}$.

The Start of Spreading

At the initial flow stage, when the shock wave is attached to the contact perimeter, the flow pattern in the vicinity of this boundary is similar to that of steady interaction between a supersonic flow and a rigid wedge with an angle β. The problem of water flow about a wedge is considered in [13]. It has two solutions, with strong and weak shocks. For a drop, the weak-shock solution is realized [14]. The solution exists, that is, the shock is attached, as long as the angle β does not exceed a certain critical value $\beta^c = \beta(t^c)$ that specifies the instant t^c of shock detachment from the rigid surface. The critical angle β^c can be found analytically using the associated relationship for a shock wave (the approximate solution is given in [15]). To the value of the critical angle found, there corresponds the

density ρ^c at the instant t^c (Fig. 4). At this instant, the highest pressure in a spherical drop is distributed over a ring (Fig. 5) and far exceeds the pressure of a one-dimensional impact $p^{1D} = \rho_0 D V_0$, where the shock wave velocity D at $M_0 = V_0/a_0 < 1.2$ is well approximated by a linear function $D = a_0 + kV_0$ (with $k = 2$ for water).

In [1], various hypotheses for the initial time of spreading are discussed. Numerical simulation shows that, without viscous effects and surface tension, the jet appears exactly at the instant t^c. In a real drop, the above factors retard jet initiation. For example, in experiments *B*, viscous effects play a decisive part: the shock wave is deflected from the contact only slightly (the bend of the free surface in the interaction zone is as small as several degrees) and does not form a radial jet (see frame *3* in Fig. 2).

What are the reasons for jet initiation retardation? The flow inside the drop shows appreciably differing flow scales in the jet vicinity and in the rest of the drop volume. Viscous forces are significant near the jet origin. At the initial stages of detachment, the shock wave moves away from the obstacle slowly, while the viscous boundary layer inside the jet origin may grow fast. Thus, up to a certain time instant, the initiation of the cumulative jet is retarded by viscous forces, which is confirmed by the simulation results (Fig. 6).

Comparing the characteristic Reynolds and Weber numbers ($\mathrm{Re} = \rho_0 V_j L/\mu$ and $\mathrm{We} = PL/\sigma$, where V_j and L are the scales of the velocity and the least cross size of the jet, respectively; μ and σ are the viscosity and surface tension coefficients, respectively; and P is the pressure scale, which can be roughly estimated as $p^{1D} \approx \rho_0 a_0 V_0$, one can estimate values of the interaction parameters at which the surface tension effect is dominant. This requires the ratio Re / We to be considerably larger. Assuming that $V_j \approx 10V_0$, we obtain $\mathrm{Re/We} = \sigma/a_0\mu$. It should be noted that this ratio depends only on the physical parameters of a medium (μ, σ, and a_0) and equals 0.05 for water. Notice also that the viscosity and surface tension effects do not influence the compression ρ^c of the liquid.

Jet Velocity

Under the assumption that viscous and surface tension forces are absent, the jet velocity V_j^c can be estimated by considering sudden spreading of the liquid into a vacuum with a velocity V_v and its acceleration in the jet throat, formed by the rigid wall and the detached shock wave. In this case, we obtain

$$V_j^c = \frac{V_v}{\sin\beta^c} + V_2 = \frac{1}{\sin\beta^c}\int_0^{p^c}\frac{dp}{a\rho} + V_2, \qquad (2)$$

where $p^c = p(\rho^c)$ is the pressure behind the shock wave and V_2 is the induced flow velocity at $\beta = \beta^c$.

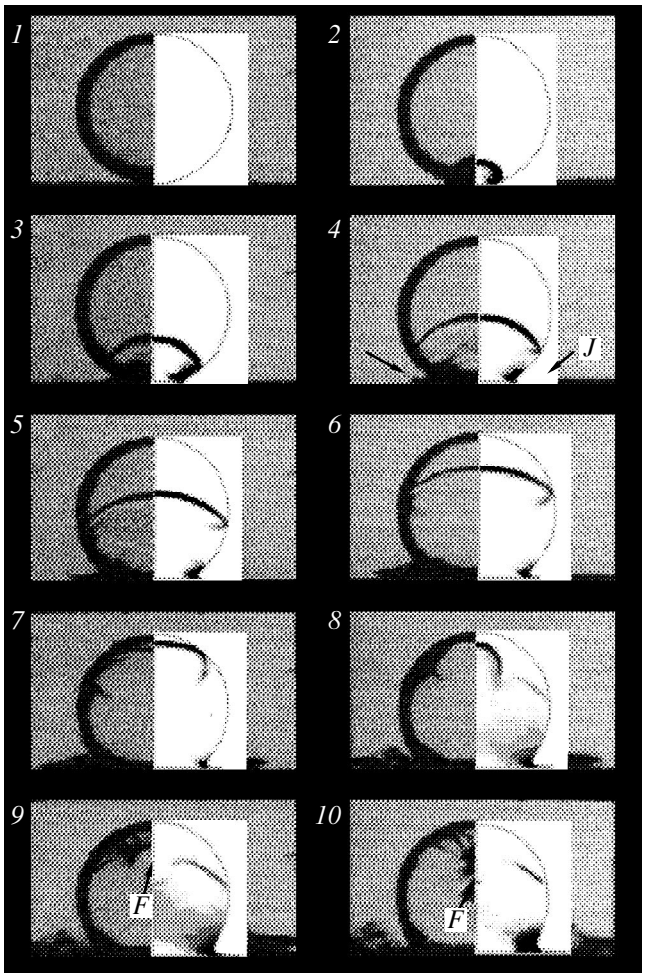


Fig. 2. Impact of a 10-mm cylindrical drop with a velocity of 110 m/s. The time step between the images is 1 μs. *J*, cumulative jet; *F*, region of expansion wave focusing.

Having found the critical values of p^c and V_2 from the relationships for a shock wave at the time instant t^c or using Fig. 4 (usually $V_2 \ll V_j^c$), we come to the impact velocity dependence on the jet velocity (Fig. 7, solid line *1*).

Another way of estimation uses the fact that the flow in the drop is akin to self-similar spreading. For the latter, an analogue of the Bernoulli integral is derived. Then, with its help, one obtains the same formula as in [1]. In deriving this formula, the flow was assumed to be steady in the spread point frame of reference and the Bernoulli equation

$$V_j^s = V_0\left(\frac{1}{\tan\beta} + \frac{1}{\sin\beta}\right) \qquad (3)$$

was applied.

Both ways of jet velocity estimation give close results (Fig. 7, dashed and solid lines *1*). This is supported by the results of numerical simulation for an

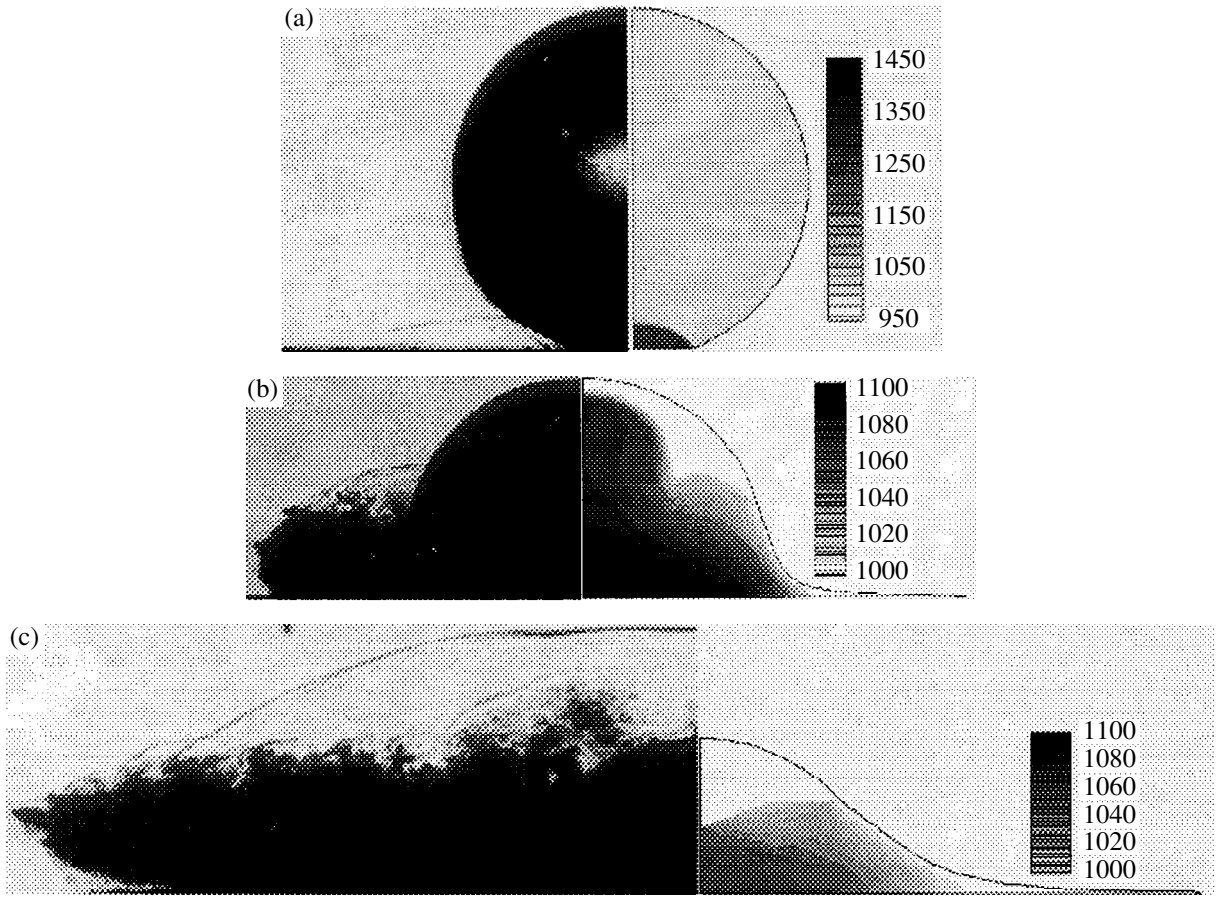


Fig. 3. Impact of a rigid body on a water drop of radius 1.2 mm with a velocity of 660 m/s in air. t = (a) 0.1, (b) 1.274, and (c) 2.0 μs.

ideal impact. For real drops, the jet velocity may be several times smaller than the estimates of V_j^c and V_j^s due to the action of viscous and surface tension forces.

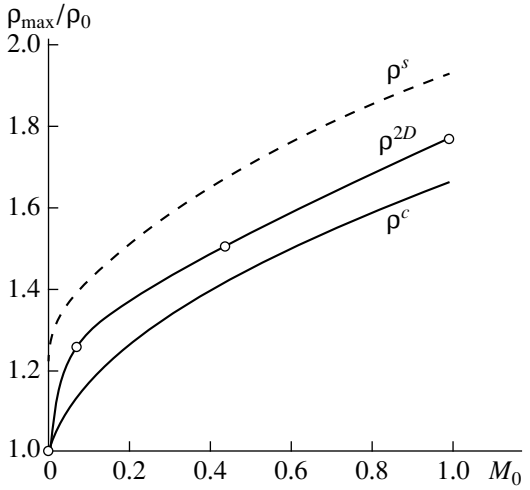


Fig. 4. Highest water density in the drop against the initial drop velocity. ρ^c is the density at the instant of shock wave detachment from the contact boundary, ρ^s is the density estimation under the assumption of a self-similar flow in the jet throat, and ρ^{2D} is the numerical simulation results.

The experimental results (curve *2* and data points *3–7* in Fig. 7) and the calculation (point *10*) represent data for collisions of real water drops. Experimental dependence *2* was obtained in [1]; data points *3* and *4*, in [10]; data points *5* and *6* were obtained when a rigid body flying in air struck a drop about 3 mm in diameter [10, 16–18]; and data point *7*, in [1]. Squares *8* and *9* are results of calculation in the frame of the model of an ideal barotropic liquid with parameters corresponding to experiments *3* and *6*, and squares *10* and *11* refer to similar calculations allowing for viscosity. Note that in [6] the relative jet velocity, $V_j/V_0 > 1500$, was experimentally measured at low impact velocities ($V_0 \sim 0.01$ m/s).

From these data, we can conclude that the assumption on the nonviscous liquid in the drop is valid and that the cumulative effect and viscous forces are significant in the near-wall region of the flow.

Maximum Pressure

To find the maximum pressure in the liquid during the impact, we will estimate the pressure, assuming that drop spreading is accompanied by jet formation. The

estimate is derived like formula (3):

$$\int_0^{p^s} \frac{dp}{\rho} = \frac{V_0^2}{2\sin^2\beta}. \quad (4)$$

If a flow similar to liquid wedge spreading were set in instantaneously at $\beta = \beta^c$, then the pressure p^s or the corresponding density ρ^s would be observed in the drop (see Fig. 4). Moreover, the density would decrease with increasing angle β according to formula (4). Actually, the flow regime akin to self-similar is retarded, and the pressure turns out to be smaller. This is corroborated by the calculated values of the largest density ρ^{2D} shown in Fig. 4 by circles. Thus, ρ^s turns out to be the upper estimate, and the density ρ^c at the instant of shock wave detachment, the lower one.

Now let us find the shape of the rigid surface that provides the highest pressure at a given impact velocity of a spherical drop. Since the maximum pressure in a smooth convex drop striking a flat target is close to the pressure at the instant the shock wave is detached from the contact boundary, a higher pressure may be achieved when the shock wave is focused on the symmetry axis, provided that immediately before focusing the angle between the free surface and the wall is equal to the critical value β^c. In this case, the contact arranges itself into a ring at the initial impact stage. A similar pattern was observed upon modeling drop fall into a conical crater with a half-angle of $\pi/2 - \beta^c$.

"Irregular" Shock Wave–Free Surface Interaction

In experiments [11, 12, 19], it was found that lateral spreading of the drop at the impact stage, when the bow shock wave has not yet reached the drop top, is localized in the near-wall region, while the edges of the compression wave on the free surface are at a noticeable distance from the wall (Fig. 1b). This effect is explained by the fact that the expansion wave propagating from the free surface interacts with the shock wave and suppresses the latter near the wall. For such "irregular" interaction [14], the fields of the hydrodynamic parameters obtained by simulating the impact at an initial drop velocity $V_0 = 660$ m/s are given in Fig. 8.

Jet Disruption Mechanism

A high-velocity jet moving in an ambient gas experiences a strong counteraction from this gas. Estimates based on known scales and jet velocities [20] allow one to check the conditions of jet instability in air, compare the times of jet formation and instability evolution, and elucidate a possible disintegration mechanism of the jet edge. Based on the jet-thickness-dependent characteristic Weber, Laplace, and Bond numbers, one can separate out various disruption mechanisms (by analogy with drop disintegration in gas flows [20]). For example, in experiments *B*, surface tension forces and deceleration forces of the ambient gas were found to cause jet disruption even at the early stage of propagation. The disruption mechanism is continuous stripping of the jet surface layer for which $60 < We < 1000$. In this case, up to a certain stage of deformation, disruption is due only to the separation of the boundary layer in the form of fine droplets from the jet surface. The droplet size may be tenths of the initial cross scale of the jet. These fine droplets experience a strong influence of the ambient gas and also interact with each other and with the jet.

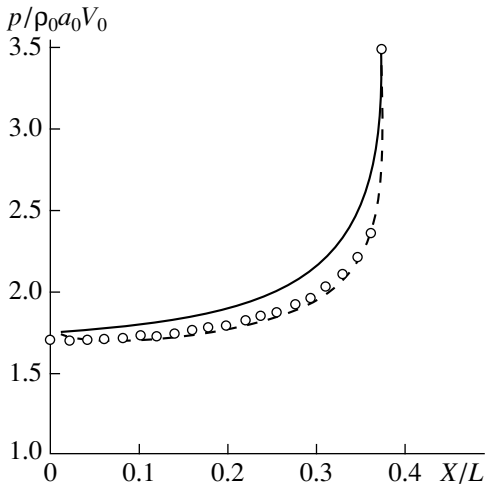


Fig. 5. Pressure distribution over the obstacle (the origin is at the symmetry axis). Solid line, acoustic model [16]; circles, nonlinear model [17]; and dashed line, two-dimensional calculation. The scale factor is $L = RV_0/a_0$.

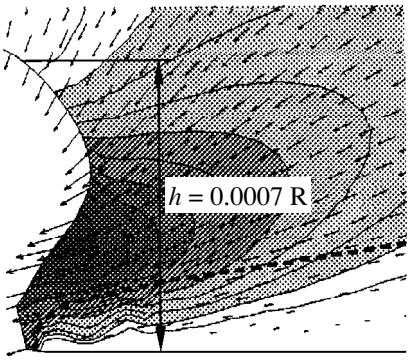


Fig. 6. Horizontal velocity field. The dashed line shows the edge of the boundary layer retarding the formation of the nonviscous core of the high-velocity jet.

EXPERIMENTS VS. NUMERICAL SIMULATION

Experiment A

In [12, 19], cylindrical "drops" of various cross sections were prepared from a gel and placed between glass plates. An impact on the side surface was performed by a flat metal striker to provide a practically

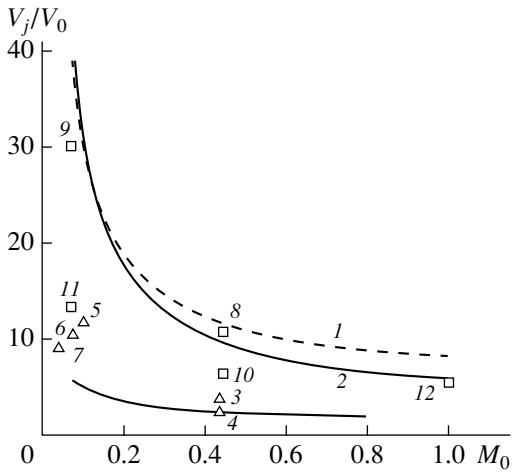


Fig. 7. Jet velocity against the initial water drop velocity. (*1*) Estimate for a nonviscous drop; triangles *3–7* and curve *2*, experimental data; squares, numerical simulation with (*8*, *9*, and *12*) and without (*10* and *11*) allowance for liquid viscosity.

plane flow. In photos, one can see the distinct shock-wave flow pattern inside the drop, the jet, and regions where cavitation cavities form. In particular, an impact by a disk with an initial velocity of 110 m/s (experiment *A*) was analyzed. Important features of the interaction were demonstrated. The first one, essential in erosion destruction processes, is a high-velocity near-wall jet, which appears within a time after the collision. The jet velocity measured from the photos was 1170 m/s, that is, approximately ten times greater than the impact one. Another feature is the formation of cavitation areas near the rigid surface and in that part of the drop most distant from the disk wall. Measurements with the help of piezoelectric transducers strengthened the theoretical conclusion that the pressure at the contact periphery is higher than at its center. The experiments also demonstrated that the jet originates much later than the shock wave is detached from the contact boundary.

Figure 2 shows photos taken in [12] for ten time instants at 1-μs intervals. On the right of the photos, the calculated pressure gradient fields obtained in [21] are displayed. Results of physical modeling and numerical simulation imply that the configurations of the high-gradient regions—shock and expansion waves at time instants *1–9*—qualitatively coincide. In both cases, the cumulative jet arises at time instant *4*, while the analytical solution neglecting viscosity and surface tension gives the spreading instant preceding instant *2*. The experimental and calculated velocities are equal to 1170 and 1470 m/s, respectively, which appears to be a good agreement. Note that the model that ignores viscous effects gives the velocity 3300 m/s. This favors the conclusion that the influence of viscous forces is appreciable.

The dark diffuse regions near the free surface behind the shock wave in Fig. 2 at time instants *6–8* are areas where toroidal cavitation originates and expands. At instant *9*, the cavitation region is seen behind the expansion wave in the form of diffuse dark fringes on the left and as a white region on the right. At time instant *10*, the expansion wave is focused in the region that corresponds to the center of the calculated cavitation region. The shape of the dark area and its size, which is less than that of the cavitation region, are evidence for the complicated formation of cavities in the gel. They also suggest that the ultimate strength of the gel is higher than that of water.

Thus, it follows that the experimental and analytical results are in good agreement and that the mathematical model stated above is adequate to the phenomenon being considered. The analysis of the calculations validated and extended the treatment of the experimental data proposed by the authors of [12].

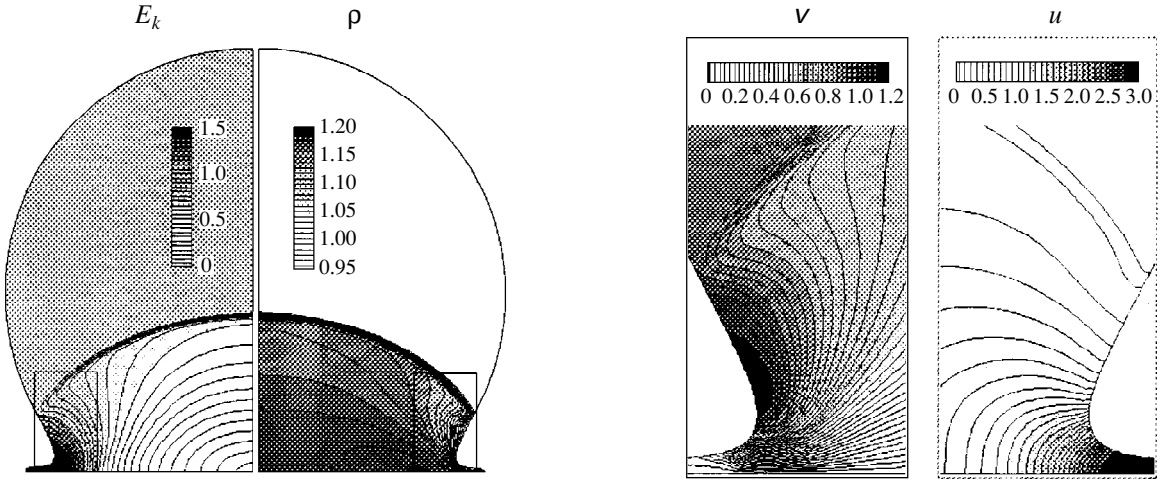


Fig. 8. Impact of a spherical drop with a velocity of 660 m/s. "Irregular" interaction between the shock wave and the free drop surface is observed when the shock wave almost does not change the drop shape and drop spread is localized near the wall. E_k, kinetic energy; ρ, liquid density; and *u* and *v*, longitudinal and transverse (relative to the wall) velocity components, respectively.

Influence of the Axial Symmetry of the Impact

Experimental studies on impacts of spherical drops and those of other essentially three-dimensional shapes are difficult because of the nontransparency of small drops due to a large curvature of their free surfaces. To overcome this challenge, drops of plane or cylindrical geometry are used in experiments. The differences in flow patterns produced by drops of spherical and cylindrical geometry are elucidated by numerical simulation.

At the stage of shock wave initialization near the rigid surface and at the initial stage of its detachment from the contact, the flows are practically identical. As the shock wave propagates toward the drop top, it gradually decreases, since its front expands and the wave interacts with the expansion wave. In a spherical drop, the shock wave decreases and decelerates faster than in a cylindrical drop; therefore, in the latter case, the shock wave reaches the drop top more quickly. Qualitatively, the spreading patterns for cylindrical and spherical drops are identical.

Experiment B

In the experiments carried out at the Ioffe Physicotechnical Institute [10, 18], photos of a drop striking the end face of a cylinder flying in air with a velocity of 660 m/s were taken. Figures 3a–3c show combined spreading patterns for three time instants. The left sides of each of the panels depict the drop and the external gas flow; and the right sides, the density distribution fields. The experimental data allow us to trace the drop shape evolution, the ambient gas flow, and jet spraying. However, the internal drop structure remains unclear. Numerical simulation gives better insight into the distribution of the gas-dynamics parameters inside the liquid. The experimental and numerical results in combination make it possible to visualize the phenomenon under study as a whole.

At the left of Fig. 3a, the shock wave in the gas surrounding the drop is seen. This wave formed when the drop entered the shock layer of the gas about the flat-end impacting body. The wave reflected from the drop is also shown. At the right of this figure, the dark area corresponds to the compressed liquid of the drop with a still indistinguishable jet.

Most of the left part of Fig. 3b is occupied by a cloud of sprays formed when the front, initially very thin, edge of the jet disintegrates. The shock wave in air resulting from the radial motion of the liquid near the wall is also seen. The spreading liquid forces out the air with a supersonic velocity, thus forming the shock wave. The origination and propagation of this shock wave in the gas are confirmed by the results of numerical simulation [22]. At the right of Fig. 3b, the light area represents the cavitation region formed after the reflection of the primary shock wave from the upper part of the free surface of the drop and the propagation of the expansion waves inside the drop. Since the mathematical model does not involve a jet disruption mechanism, the cloud of sprays does not figure as a final result in the calculation.

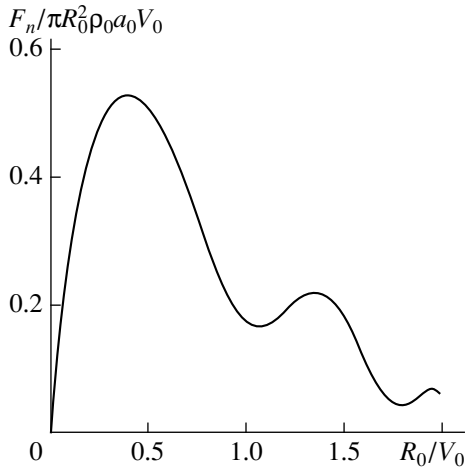


Fig. 9. Time variation of the total force exerted on the obstacle by the drop.

Force Acting on the Obstacle

The time dependence of the liquid pressure force acting on the obstacle is shown in Fig. 9. It was obtained by considering a water drop striking a flat surface with a velocity $V_0 = 660$ m/s.

Alternating compression and expansion waves propagating near the obstacle make this dependence nonmonotonic. The maximum force is observed at a time instant $t \approx 0.4R_0/V_0$, coincident with the formation instant of the toroidal cavitation area behind the shock wave near the free surface (Fig. 1c). Subsequently, the expansion wave propagating from the lower part of the free surface increasingly shrinks the compressed liquid region near the wall. Then, this wave is focused on the symmetry axis, forming one more cavitation area. With time, this new area joins the cavitation area that formed when the bow shock wave reflected from the "upper" part of the drop (Fig. 1d). Since the shock wave propagating over the free drop surface gradually decreases, the liquid away from the symmetry axis retains the velocity vector directed toward the target. This liquid is then decelerated by the surface, causing one more compression cycle, and the cavitation area collapses. The force acting on the target reaches a local minimum at an instant $t \approx 1.1R_0/V_0$ and then starts to increase again. Since the drop–wall distance has decreased nearly twice, the new compression–expansion cycle is shorter than the first one and the next minimum of the force on the wall is observed at an instant $t \approx 1.8R_0/V_0$. The absolute maximum of the force amounts to approximately half the pressure force of a one-dimensional impact. In the acoustic approximation, it equals $\rho_0 a_0 V_0$ under the assumption that the force acts on an area equal to the maximum midsection of the drop.

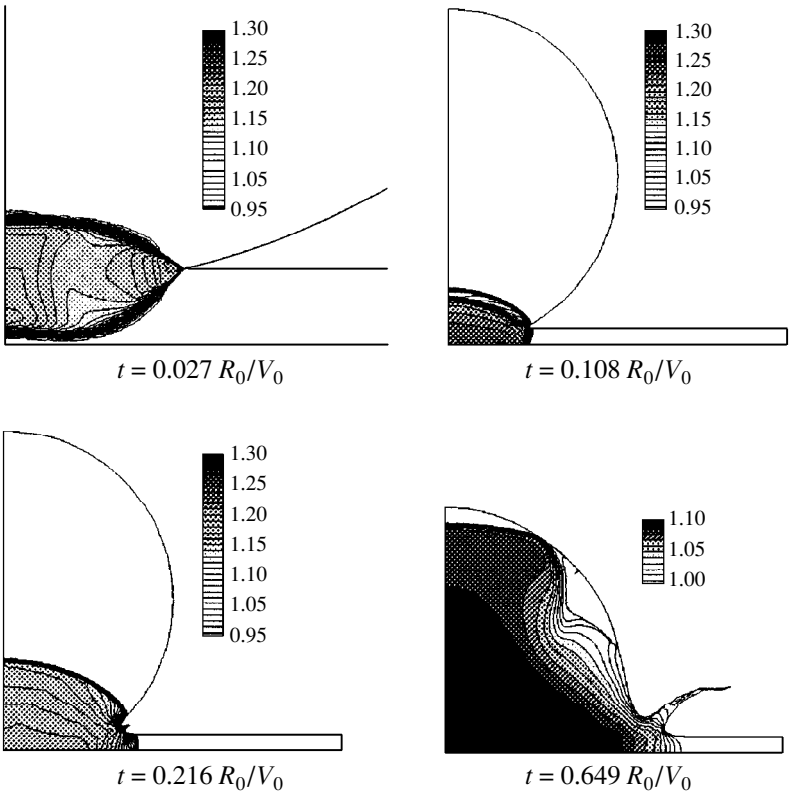


Fig. 10. Drop–liquid layer collision with a velocity of 660 m/s.

Drop–Liquid Layer Collision

Repeat collisions between drops and rigid obstacles frequently lead to the formation of a liquid layer on the collision surface. In this case, the qualitative interaction pattern markedly changes. As an illustration, Fig. 10 represents the density distributions at various instants of a drop impact on a water layer with a velocity of 650 m/s for an H/R_0 ratio of 0.1, where H is the layer thickness and R_0 is the drop radius.

The qualitative discrepancy between liquid–solid and liquid–liquid impacts shows up even at the initial interaction stage. Near the drop–layer contact surface, two shock waves arise. One propagates into the drop, and the other, toward the rigid surface. The strength of either wave is comparable to half the strength of the initial shock wave striking the rigid surface. The second wave reflects from the surface and then follows the first one. At a certain time instant, the first wave is detached from the boundary between the two free surfaces. However, in the presence of weak viscosity, drop spreading is retarded up to the instant the second shock wave, which causes the cumulative jet, reaches the contact boundary. Thus, the instant of jet origination does not depend on the Reynolds number if its value is sufficient to delay the jet up to the time instant the shock wave reflected from the surface arrives but cannot delay jet formation after arrival of the second wave.

Once the shock wave reflected from the surface has caught up with the first one, the flow pattern inside the drop becomes qualitatively similar to that observed for the liquid–solid impact. The jet flows at an increasing angle to the layer surface, producing a liquid train around the drop.

CONCLUSIONS

In this work, we analyzed the collision of a liquid drop with an obstacle, the collision velocity being comparable to the sound speed in a liquid. In such an interaction, the medium compressibility is essential. The general qualitative pattern of the flow and its features are described in the first two sections.

At the initial stage of interaction, drop spreading is not observed. At a certain time instant, the radial cumulative jet appears. Its velocity can be evaluated using formulas (2) and (3). The jet origination may be appreciably retarded by viscous forces. The liquid undergoes the largest compression at the initial stage (where the free drop surface is in contact with the obstacle and in the throat of the radial jet). The pressure distribution and its maximum can be estimated from Figs. 4 and 5 and formula (4). The time dependence of the total force exerted on the obstacle by the drop is nonmonotonic (Fig. 9).

An important feature of the process is "irregular" interaction between the shock wave and the free drop surface (Fig. 8). It should also be noted that the jet moving in the gas surrounding the drop interacts with it and may be disrupted.

The numerical simulation provided better insight into the flow behavior (Figs. 2 and 3). Together with the experimental data, it allowed us to make up the total pattern of the liquid flow inside the drop and the gas flow outside (Fig. 3). Also, the practically important case of drop–liquid layer collision was considered (Fig. 10), and the role of the shock wave reflected from the layer bottom was emphasized.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (grants Nos. 95-01-00521 and 98-01-01117).

REFERENCES

1. *Erosion*, Ed. by C. Preece (Academic, New York, 1979; Mir, Moscow, 1982).
2. G. S. Springer, *Erosion by Liquid Impact* (Scripta, Washington, 1976; Moscow, 1981).
3. M. Rein, in *Proceedings of IUTAM Symposium on Waves in Liquid/Gas and Liquid/Vapor Two-Phase Systems, 1995*, pp. 171–190.
4. D. N. Contractor, J. Fluids Eng. **94** (2), 207 (1972).
5. M. Rein, Fluid Dyn. Res. **12**, 61 (1993).
6. J. G. Black and G. H. Miller, in *Abstracts of the 21st International Symposium on Shock Waves, 1997*, p. 25.
7. A. A. Korobkin and V. V. Pukhnachov, Annu. Rev. Fluid Mech. **20**, 159 (1988).
8. A. L. Gonor and V. Ya. Rivkind, Itogi Nauki Tekh., Ser.: Mekh. Zhidk. Gaza **17** (1982).
9. M. Lesser and J. Field, Annu. Rev. Fluid Mech. **15**, 97 (1983).
10. I. A. Dukhovskiĭ and P. I. Kovalev, Prib. Tekh. Éksp., No. 6, 102 (1996).
11. J. Brunton and J. J. Camus, in *Proceedings of the III International Conference on Rain Erosion, 1970*, p. 327.
12. M. Lesser and J. Field, in *Proceedings of the XVIII International Symposium on Shock Waves, 1991*, p. 61.
13. S. Ridah, J. Appl. Phys. **64** (1), 152 (1988).
14. A. V. Chizhov and A. A. Shmidt, Pis'ma Zh. Tekh. Fiz. **22** (3), 57 (1997) [Tech. Phys. Lett. **22**, 117 (1997)].
15. F. J. Heymann, J. Appl. Phys. **40**, 5113 (1969).
16. M. Lesser, Proc. R. Soc. London, Ser. A **377** (1770), 289 (1981).
17. A. V. Chizhov, Candidate's Dissertation in Mathematical Physics (St. Petersburg State Technical University, St. Petersburg, 1998).
18. I. A. Dukhovskiĭ, P. I. Kovalev, and A. A. Shmidt, Pis'ma Zh. Tekh. Fiz. **10** (11), 649 (1984) [Sov. Tech. Phys. Lett. **10**, 274 (1984)].
19. J. P. Dear and J. E. Field, J. Appl. Phys. **63** (4), 1015 (1988).
20. A. I. Ivandaev, A. G. Kutushev, and R. I. Nigmatulin, Itogi Nauki Tekh., Ser.: Mekh. Zhidk. Gaza **16**, 209 (1981).
21. A. V. Chizhov and A. A. Schmidt, in *Proceedings of the IX International Conference on Finite Elements in Fluids, Italy, 1995*, Vol. 2, p. 1019.
22. M. Hirano, K. Tokayama, and J. Falcovitz, in *Proceedings of the 21st International Symposium on Shock Waves, 1997*, p. 192.

Translated by N. Mende

Technical Physics, Vol. 45, No. 12, 2000, pp. 1538–1548. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 28–38.

GASES AND LIQUIDS

Evolution of Cluster Size Distribution During Nucleation with Rapidly Changing Dynamic Gas Processes

Yu. E. Gorbachev and I. S. Nikitin

Institute for High-Performance Computing & Data Bases, St. Petersburg, 198005 Russia

e-mail: gorbachev@hm.csa.ru

ivan@fn. csa.ru

Received January 10, 2000; in final form, April 27, 2000

Abstract—A non-steady-state solution of the Zel'dovich equation for the cluster size distribution function has been obtained. The solution is applicable in the entire range of cluster sizes. The relaxation time of the quasi-steady-state distribution and the partial time lags of cluster formation were derived and analyzed. The dependence of these times on cluster size was investigated. Expressions for the quasi-steady-state flux and the induction times have been derived. The accuracy of these calculations is confirmed by comparison with numerical calculations by other authors. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Problems relating to the non-steady-state effects in the process of condensation, in particular the gradual relaxation of the nucleation flux to its steady-state value, had been studied for a long time [1]. Two approaches were used in the analysis of deviations of the nucleation process from the steady-state regime. On the one hand, attempts were made to simplify the non-steady-state problem and, on the other hand, corrections were applied to the results of the steady-state theory.

In [2, 3], a theory of homogeneous condensation has been developed and for the first time, the concept of process stages for isothermal conditions was introduced and the time of relaxation to the steady-state distribution (kinetic relaxation time) was estimated. In this theory, the problem of the approach to the steady-state regime has not been ultimately solved. In most of the studies, the so-called time lag was determined [1, 4–6] and is defined as the time taken by a flux in dimensional space to reach its steady-state value. A number of studies have dealt with the numerical computations of non-steady-state nucleation [7].

The most important results in the analysis of the evolution of the cluster size distribution function during condensation were obtained in [8, 9]. In addition to determining the characteristic times of relaxation to the quasi-steady-state, the authors have derived a quasi-steady-state cluster distribution and a corresponding expression for the critical nucleation flux, and have suggested a self-contained system of condensation equations. Using different approaches, an attempt was made in [10–12] to consider the nucleation process in the framework of a quasi-chemical model and to reformulate of the classic nucleation theory by Volmer–Weber–Becker–Döring–Frenkel–Zel'dovich with the use of kinetic theory techniques based on the separation of fast and slowly varying parameters of the system [13–15].

The classic nucleation theory, in addition to a number of other assumptions (for example, see [16]), is based on the assumption of a great difference in the time scale for variations of the monomer and g-mer concentrations with $g \geq 2$. In [17], it was noted that the probability of nucleation as a result of fluctuations, which is considered an equilibrium characteristic, consequently causes the distribution function to not correspond to complete statistical equilibrium and it is applicable only for times that are short compared with the formation time of critical nuclei.

In studies of the nucleation process, two characteristic cluster size distributions are usually considered: an equilibrium distribution, n_g^e, which turns to zero for the flux of clusters in a space of size ($j(n_g^e) = 0$), and a steady-state distribution, n_g^s, which is defined by the relation $j(n_g^s) = \text{const}$. In the latter case, for $g \geq 2$, $\dot{n}_g = 0$ and $\dot{n}_1 \neq 0$. Both functions n_g^e and n_g^s are defined by a set of gasdynamic variables that describe the macroscopic state of the system. An increase in the rate of variation of one variable, for example, the monomer concentration, will result in a deviation in the distribution from its equilibrium value, because the solution is dependent on the derivative of the function obtained in a zero approximation by a small parameter (in this case on $\dot{n}_g^s$). Data providing evidence that the nucleation process depends not only on the gasdynamic parameters but also on their rate of variation, were discussed in [10, 11].

1063-7842/00/4512-1538$20.00

In the present work, a method is proposed for finding the distribution function that is sensitive to the variation rate of the system's dynamic gas parameters. This method is based on a simplified approach to the system description by distinguishing fast and slow processes in the system, as suggested in [13–15]. An expression for the distribution function has been obtained and is referred to below as a quasi-steady-state distribution. The temporal behavior of the distribution function as it approaches the quasi-steady-state was analyzed. Individual time lags for formation of the quasi-steady-state distribution of the clusters were introduced and their relation to the time lag of the quasi-steady-state nucleation flux have been established.

FORMULATION OF THE PROBLEM

We assume a differential form of the equation for the cluster size distribution function $n_g(g, t)$

$$\frac{\partial n_g}{\partial t} = -\frac{\partial j_g}{\partial g}, \tag{1}$$

where the flux j_g is given by

$$j_g = -D_g n_g^e \frac{\partial}{\partial g}\left(\frac{n_g}{n_g^e}\right) = -D_g\left(\frac{\partial n_g}{\partial g} + \frac{n_g}{kT}\frac{\partial \Phi}{\partial g}\right), \tag{2}$$

D_g is the diffusion coefficient; $n_g^e = n_1\exp(-\Phi/kT)$ is an equilibrium cluster size distribution function; $\Phi(g)$ is the minimum formation energy of a drop of g-monomers; and T is the vapor temperature. The small parameter of the theory is considered to be a quantity ε

$$\varepsilon = \frac{\Delta}{g_*} \ll 1, \quad \Delta = \left(-\frac{1}{2kT}\frac{\partial^2 \Phi}{\partial g^2}\bigg|_*\right)^{-1/2}. \tag{3}$$

Here, Δ is the halfwidth of the function Φ and the asterisk indicates that the quantity is calculated for $g = g_*$, where g_* is the critical cluster size determined for the condition at which the function Φ reaches a maximum: $\partial\Phi/\partial g|_* = 0$. In the steady-state case, where j = const, the distribution function has a known form

$$n_g^s = j^s n_g^e\left(\int_g^\infty (D_g n_g^e)^{-1} dg\right)$$
$$= n_g^e\left(1 - j^s\int_1^g (D_g n_g^e)^{-1} dg\right) \simeq \frac{1}{2} n_g^e erfc\left(\frac{g - g_*}{\Delta}\right). \tag{4}$$

In deriving the latter relationship, a quadratic approximation is used for the formation energy $\Phi(g)/kT = \Phi(g_*)/kT - (g - g_*)^2/\Delta^2$. The steady-state flux j^s is given by

$$j^s = \left(\int_1^\infty (D_g n_g^e)^{-1} dg\right)^{-1} \simeq \frac{D_* n_*^e}{\Delta\sqrt{\pi}}. \tag{5}$$

To determine how the distribution function approaches its quasi-steady-state value, the non-steady-state distribution function $n_g(g, t)$ is of great interest. In order to obtain it from Eq. (1), a new quantity $\tilde{n}_g$ is introduced [10]

$$\tilde{n}_g = n_g/n_g^s. \tag{6}$$

This function has a number of advantages over the more commonly used quantity, n_g/n_g^e, which is discussed below. One advantage is that $\tilde{n}_g$ is a smoother function of g than n_g/n_g^e, which makes the diffusion approximation applicable over a wider range in the supercritical region. With the new variables, Eq. (1) takes the form

$$\frac{\partial \tilde{n}_g}{\partial t} = \dot{g}\frac{\partial \tilde{n}_g}{\partial g} + \frac{\partial}{\partial g} D_g \frac{\partial \tilde{n}_g}{\partial g} - \tilde{n}_g \frac{\partial \ln n_g^s}{\partial t}, \tag{7}$$

where

$$\dot{g} = \frac{j^s}{n_g^s} + D_g \frac{\partial \ln n_g^s}{\partial g} = -\left(\frac{2j^s}{n_g^s} + \frac{D_g \partial \Phi}{kT\,\partial g}\right). \tag{8}$$

The coefficient of the first derivative, usually denoted $\dot{g}$, has the meaning of drift in the dimension space, and differs from Einstein's expression $\dot{g} = -D_g\Phi'_g/kT$. As shown in [12] with both expressions coinciding in the subcritical region, in the supercritical region, expression (8) diminishes, whereas Einstein's expression increases, and changes sign at $g \simeq g_*$. An important property of expression (8) is the fact that it does not turn to zero at any point in the range of g values, which significantly simplifies the analysis of Eq. (7). The boundary conditions for our problem are

$$\tilde{n}_g(g, t)\big|_{g=1} = 1; \quad \tilde{n}_g(g, t) \longrightarrow 0, \quad g \longrightarrow \infty, \tag{9}$$

where the first equality following from the observation that $n_g \longrightarrow n_1^s = n_1$ for $g \longrightarrow 1$ (for vapor, the distribution is equilibrium). The second equality follows from its existence in the system of maximum size clusters, which is confirmed in the analysis below.

The purpose of the following analysis is to consider Eq. (7) under boundary conditions (9) to determine the characteristic time periods of the system evolution and clarify the features of its solution for long-time scale.

ANALYSIS OF THE BASIC EQUATION

Let us write Eq. (7) in dimensionless variables taking into account the physical meaning of all its terms. The first two terms on the right-hand side govern the kinetic processes characterized with a kinetic relaxation time τ^k, which is assumed to correspond to the relaxation time of the critical nucleus [2, 3],

$$\tau^k = \frac{\pi\Delta^2}{4D_*}. \tag{10}$$

As shown in the analysis below, this is the quantity that belongs to the basic characteristics determining the temporal dependence of the g-mer concentrations (see (23)).

The last term on the right-hand side of Eq. (7) can be defined as a gasdynamic source, because it describes the deviation of the distribution function from its steady-state value at varying gasdynamic quantities (to be defined below). This term is characterized by a time scale that is dependent on the cluster size g and is called gasdynamic (because n_g^s depends only on gasdynamic variables)

$$\tau_g^{gd} = \left|\frac{d\ln n_g^s}{dt}\right|^{-1}. \tag{11}$$

For small clusters $\tau_g^{gd} \sim 1/g$, and taking into account that for supercritical sizes, $n_g^s \simeq n_*^e/2 + Ag^{-\alpha}$, $\alpha > 0$, $\tau_g^{gd} \sim \tau_*^{gd} = |d\ln n_*^e/dt|^{-1}$ is obtained. Thus, the characteristic gasdynamic time, which decreases in the subcritical region, becomes constant in the supercritical region.

Let us convert t into a dimensionless quantity T depending on the time scale within which the behavior of the system is considered. We introduce the dimensionless parameters Sh, μ_g, and $\tilde{\mu}_g$

$$\mathrm{Sh} = \tau^k/T, \quad \mu_g = \tau^k/\tau_g^{gd}, \qquad \tilde{\mu}_g = \mu_g \mathrm{sgn}\{d\ln n_g^s/dt\}. \tag{12}$$

The final form of our equation in dimensionless variables (t' is the dimensionless time, $u = g/g_*$) will be

$(\partial/\partial t \longrightarrow (1/T\partial/\partial t' - ug_*^{-1}\dot{g}_*\partial/\partial u)$, $\partial/\partial g \longrightarrow g_*^{-1}\partial/\partial u$, $\dot{g}_* = \partial g_*/\partial t)$

$$\mathrm{Sh}\frac{\partial\tilde{n}_u}{\partial t'} = \varepsilon\sqrt{\pi}\dot{\tilde{g}}\frac{\partial\tilde{n}_u}{\partial u} + \varepsilon^2\frac{\pi}{4}\frac{\partial}{\partial u}\frac{D_u}{D_*}\frac{\partial\tilde{n}_u}{\partial u} - \tilde{\mu}_u\tilde{n}_u, \qquad \dot{\tilde{g}} = (\dot{g} + u\dot{g}_*)/|g_{\max}|. \tag{13}$$

Here, the drift velocity is normalized to its maximum absolute value (see (18)). Since the coefficients in Eq. (13) depend on the gasdynamic variables, their further analysis requires knowledge of the rate of variation of these parameters. The gasdynamic variables include supersaturation $s = p/p_e$ (p is the vapor pressure and p_e is the equilibrium vapor pressure under the same conditions), the vapor temperature T, and the monomer concentration. To these, the velocity v will be added if convection effects are taken into account; on the left-hand side of Eq. (7), the term $\nabla(\tilde{n}_g v)$ appears. Variation of these parameters is described by corresponding equations of which only the equation for the monomer concentration is shown here, because of its special role

$$\partial n_1/\partial t = j_1 + \sum_{g=1}^{\infty} j_g \approx \int_1^{\infty} j_g dg. \tag{14}$$

An analysis of this equation will be given below.

If the time scales for variations of gasdynamic and kinetic quantities (which include concentrations of g-mers with $g \geq 2$) are different, the system description can be considerably simplified. The time scale to be used in the analysis depends on the fastest gasdynamic process. For simplicity, we describe the above gasdynamic parameters using a single time scale; namely, the quantity τ_1^{gd} (see (11)). Without loss of generality, we can also assume that $T = \tau_1^{gd}$, which leads to $\mathrm{Sh} = \mu_1$.

In this way, variable rates of kinetic and gasdynamic quantities can be estimated

$$\mu_1\frac{\partial\tilde{n}_u}{\partial t'} = O(1) + O(\mu_g), \quad \frac{\partial\ln n_1}{\partial t'} = O(1).$$

At $\mu_1 \ll \max\{1, \mu_g\}$, these equations have the usual form of singularly disturbed equations [18], and a separation of kinetic and gasdynamic variables can be carried out; that is, Eq. (13) can be analyzed for weakly varying gasdynamic parameters. In this case, the time evolution of its solution can be derived. In solving Eq. (13) below, we use the Laplace transformation method with the Laplace transform of the distribution function denoted as $N(g, p)$.

Let us make use of another small parameter ε. If the quantity $\dot{\tilde{g}}$ does not become small on the order of ε, the diffusion term in (13) can be neglected at all g. In order to obtain a uniform approximation applicable for small $\dot{\tilde{g}}$ as well, we consider Eq. (13) in three characteristic regions: a) the subcritical region, $1 \ll g < g_*(1-\varepsilon) = g_-$; b) the near-critical region, $|g - g_*| \leq \varepsilon g_* = \Delta$; and c) the supercritical region, $g_*(1+\varepsilon) = g_+ < g$. In these regions, Eq. (13) takes a simpler form.

In the region of subcritical sizes (a), the diffusion term in (13) can be neglected, which gives

$$-\mu_1\tilde{n}(0,g)+(\mu_1 p+\tilde{\mu}_g)N-(\dot{g}+g\dot{g}_*/g_*)\tau^k\frac{dN}{dg}=0, \qquad (15)$$
$$\dot{g}\simeq-\frac{D_g}{kT}\frac{\partial\Phi}{\partial g}.$$

Taking into account the boundary condition $N(g,p)|_{g\to 1}=1/p$, which follows from the first of the expressions in (9), the following solution to this equation is (the expression for $\Lambda^{(a)}$ is given in Appendix 1):

$$N^{(a)}(g,p)=\frac{1}{p}\exp(-(\mu_1 p+\tilde{\mu}_g)\tau_a(g)/\tau^k)+\Lambda^{(a)}, \qquad (16)$$
$$\tau_a(g)=-\int_1^g(\dot{g}+g\dot{g}_*/g_*)^{-1}dg',$$

where $\tau_a(g)$ is the decay time of a nucleus of size g and, as shown below, is simultaneously the time lag for the establishment of the steady-state distribution function in the region considered.

In the region of near-critical sizes (b), the magnitudes of the first and second terms on the right-hand side of Eq. (13) can be of the same order. Then, if a new variable $z=(u-1)/\varepsilon$ is used, Eq. (13) becomes

$$\mu_1\frac{\partial\tilde{n}_z}{\partial t}-\tau^k(\dot{g}(g_*)+\dot{g}_*)\frac{\partial\tilde{n}_z}{\partial z}-\frac{\pi}{4}\frac{\partial^2\tilde{n}_z}{\partial z^2}+\tilde{\mu}_*\tilde{n}_z=0, \qquad (17)$$
$$z=\frac{g-g_*}{\varepsilon g_*}.$$

Here D_g, μ_g, and $\dot{g}$ have been replaced by their values at point g_*, which corresponds to the zero order approximation in parameter ε. The quantity $\dot{g}$ reaches its maximum value in region (b) near point g_*. Taking into account that $n_*^e=2n_*^s$ and making use of (5), we obtain the following expression, which is accurate to within ε

$$\dot{g}(g_*)\simeq\dot{g}_{\max}\simeq-\frac{4D_*}{\Delta\sqrt{\pi}}\simeq-\frac{4j^s}{n_*^e}. \qquad (18)$$

Applying the Laplace transformation to Eq. (17), we find a solution which meets the requirement of decreasing at $z\gg 1$ (an expression for $\Lambda^{(b)'}$ is given in Appendix 1)

$$N^{(b)}(z,p)$$
$$=C_1\exp(2\xi(1-\sqrt{1+\xi^{-2}(\mu_1 p+\tilde{\mu}_*)})z/\sqrt{\pi})-\Lambda^{(b)}, \qquad (19)$$
$$\xi=\left(1+\frac{\sqrt{\pi}\dot{g}_*}{4}\tau^k\right).$$

The constant C_1 is found under the conditions that the solutions of (19) and (16) coincide at a point $z=-1$ at the boundary between regions a and b or at $g_-=g_*(1-\varepsilon)$.

In the region of supercritical sizes c, the diffusion term in Eq. (13) becomes negligible and, just as in the subcritical region, can be omitted. Therefore, we get (the quantity $\Lambda^{c'}$ is defined by expression (42))

$$N^{(c)}(g,p)=C_2\exp(-(\mu_1 p+\tilde{\mu}_*)\tau_c'(g)/\tau^k)+\Lambda^{(c)'}, \qquad (20)$$
$$\tau_c'(g)=-\int_{g_+}^g\frac{dg}{\dot{g}+g\dot{g}_*/g_*},\quad g_+=g_*(1+\varepsilon).$$

The constant C_2 is found from the coincidence of solutions $N^{(b)}$ and $N^{(c)}$ at point g_+ at the boundary of regions b and c.

For convenience of further analysis, we introduce the following quantities

$$\tau_b=\tau_a(g_-),\quad \tau_c(g)=\tau_b+\tau_c'(g), \qquad (21)$$

which also have the meaning of the disintegration time of a nucleus of size g and, simultaneously, of a time lag for establishing the steady-state distribution function. The final form of the three solutions will be

$$N^{(i)}(g,p)=\frac{1}{p}\exp\left(-(\mu_1 p+\tilde{\mu}_i)\frac{\tau_i}{\tau^k}+\left(\frac{2\xi}{\sqrt{\pi}}\right.\right.$$
$$\left.\left.\times\left(1-\sqrt{1+\frac{(\mu_1 p+\tilde{\mu}_*)}{\xi^2}}\right)(z_i+1)\right)\right)+\Lambda^{(i)}; \qquad (22)$$
$$i=a,b,c,$$

where $\tilde{\mu}_a=\tilde{\mu}_g$; $\tilde{\mu}_{b,c}=\tilde{\mu}_*$; $z_a=-1$ and $z_b=z$, $z_c=+1$. $\Lambda^{(i)}(g,p)$ is defined by formulas (40) and (44) in Appendix 1.

Let us return to the variable t by performing the inverse Laplace transform. $n^{(a)}$ can be expressed in an analytical form, and $n^{(b)}$ and $n^{(c)}$ can be represented as quadratures, for which asymptotic expressions can be derived as $t\to\infty$ (see Appendix 2). Inverse Laplace transforms for $\Lambda^{(i)}$ at $i=a,b,c$ are given in Appendix 1, where it is also shown that, for times greater than τ_a, τ_b and τ_c, respectively, the contribution of initial conditions is small and it will be neglected here after.

Finally, for the normalized distribution function, the following representation is obtained (in which time is a

dimensional quantity):

$$\tilde{n}^{(i)}(g, t) = \Theta(t - \tau_i)\tilde{n}^{qs(i)}(g)(1 + G_i(g_i, t))(i = a, b, c),$$
$$G_i(g_i, t) = \exp(2\xi\sqrt{1 + \xi^{-2}\tilde{\mu}_*}(z_i + 1)/\sqrt{\pi}) \times \frac{\xi(z_i + 1)}{\pi(1 + \xi^{-2}\tilde{\mu}_*)}\left(\frac{\tau^k}{t - \tau_i}\right)^{3/2}\exp\left(-(1 + \xi^{-2}\tilde{\mu}_*)\frac{t - \tau_i}{\tau^k}\right), \quad (23)$$

where $g_a = g_-$, $g_b = g$, $g_c = g_+$, and z_i corresponds to g_i (therefore, $G_a(g_a, t) = 0$).

The quasi-steady-state distribution function is given by

$$\tilde{n}^{qs(i)}(g) = \exp(-\tilde{\mu}_i\tau_i/\tau^k + 2\xi(z_i + 1) \times (1 - \sqrt{1 + \xi^{-2}\tilde{\mu}_i})/\sqrt{\pi})(i = a, b, c), \quad (24)$$

where $\tilde{\mu}_i$ is defined below expression (22).

The asymptotic expressions (23) (at $i = b, c$) have been derived under the assumption that $t - \tau_b \gg \tau^k$ and $t - \tau_c(g) \gg \tau^k$, respectively. In particular, it is seen from (23) at $i = b$ that the kinetic time in the quasi-steady-state case is different from that introduced in (10), and is defined by the relationship $\tau'^k = \tau^k/(1 + \xi^{-2}\tilde{\mu}_*)$, in agreement with the result obtained in [8]. From these relationships, it is also seen that the time lag for establishment of the quasi-steady-state distribution τ_d can represent clusters of any size by a compound formula

$$\tau_d = \tau_a(g) \quad \text{at} \quad 1 \le g \le g_-,$$
$$\tau_d = \tau_b = \tau(g_-) \quad \text{at} \quad g_- < g < g_+, \quad (25)$$
$$\tau_d = \tau_c(g) \quad \text{at} \quad g_+ \le g,$$

and the relationship $\tau_c(g_+) = \tau_a(g_-)$ ensures its continuity as a function of g. During the time τ_d, drops in the system do not nucleate and flux in the dimensional space is absent. At this point it should also be noted that because τ_d is not dependent on $\tilde{\mu}_g$ when gasdynamic parameters have no effect on the nucleation process ($\tilde{\mu}_g \longrightarrow 0$), the same quantity, τ_d has the meaning of a time lag for the establishment of the steady-state size distribution of clusters.

THE TIME LAG OF THE QUASI-STEADY-STATE DISTRIBUTION

We shall calculate the times $\tau_a(g)$, τ_0, and $\tau_c(g)$ using the following approximate expression for the cluster formation energy Φ

$$\Phi(g)/kT = a(g - 1)^{2/3} - b(g - 1), \quad (26)$$

where $a = 4\pi\sigma(3\nu/4\pi)^{2/3}/kT$ is dependent on surface effects; σ is the surface tension; ν_l is the volume per monomer in a liquid phase; and $b = \ln s$, s being a supersaturation.

Then it follows from expression (3) that $\Delta = 3g_*^{2/3}/\sqrt{a}$ ($g_* = (2a/3b)^3$).

The diffusion coefficient D_g is related to the nucleus size via the association constants K_g^+, $D_g = K_g^+ n_1 \sim K_1^+ n_1 g^{2/3}$, where K_g^+ is the number of molecules attached to a cluster in a unit time and $K_1^+ = \pi\alpha_k(3\nu_l/4\pi)^{2/3}\sqrt{8kT/\pi m}$, where α_k is a condensation coefficient.

The quantity $\tau_a(g)/\tau^k$, which has a meaning at $2 \le g \le g_-$, is defined by expression (16). The second addends in the denominator of both this expression and a similar expression for $\tau_c(g)/\tau^k$, (21) are on the order of ε^2. Therefore, in calculating the time lag, they can be neglected. Then, we obtain for $\tau_a(g)$ ($u = g/g_*$)

$$\tau_a(g)/\tau^k = \frac{2}{\pi}\left(\frac{1 - g^{1/3}}{g_*^{1/3}} + \ln\frac{1 - (1/g_*)^{1/3}}{1 - (g/g_*)^{1/3}}\right). \quad (27)$$

Over the specified period of g variation, this quantity rises from $2(2^{5/3} - 1)/(5\pi g_*^{5/3})$ to its maximum value of $\tau_b/\tau^k \simeq 2/\pi(\ln 3/\varepsilon - 11/6)$, which is achieved at the boundary of the subcritical and near-critical regions.

In deriving $\tau_c(g)$, we make use of expression (21). At $g \ge g_+$ the diffusion component of the flux (2) becomes negligible (n_g^s is very weakly dependent on g in the considered range of cluster sizes) and we can write $j_{st}/n_g^s = -D_g\Phi_g'/kT$. Then, $\dot{g} = D_g\Phi_g'/kT$, which leads to the following expression for $\tau_c(g)$:

$$\tau_c(g)/\tau^k = \frac{2}{\pi}\left[\frac{1 - g_-^{1/3} + g^{1/3} - g_+^{1/3}}{g_*^{1/3}} + \ln\frac{(1 - (1/g_*)^{1/3})((g/g_*)^{1/3} - 1)}{(1 - (g_-/g_*)^{1/3})((g_+/g_*)^{1/3} - 1)}\right] \simeq \frac{2}{\pi}\left[\ln\left(\frac{9}{\varepsilon^2}((g/g_*)^{1/3} - 1)\right) + (g/g_*)^{1/3}\right]. \quad (28)$$

From the above calculations, it is seen that with increasing number g, the time lag τ_d monotonically rises. Therefore, at any particular moment of time t the equation $\tau_d(g)$ has a unique solution $g = g_m(t)$. Taking into account the ladder-like form of the distribution function given by expression (23), it will be natural to call the cluster size g_m maximum at the moment t.

At small g, $\tau_d \ll \tau^k$; therefore, there exists some cluster size g_k for which $\tau_d(g_k) = \tau^k$. If $g_k < g_-$, then $g_k/g_* \simeq (1 - e^{-2})^3$; in the opposite case, $g_k/g_* \simeq (1 + e^{-2})^3$. At $\varepsilon < 3e^{-2}$, we have the second case. Because of the smallness of ε, it can be stated that in the subcritical region, the time lag is less than the kinetic time and therefore, the cluster concentration in this size range attains its quasi-steady-state value in a time determined exclusively by the kinetic time τ^k. In the supercritical region, τ_d continues to increase and at some cluster size, g_M becomes equal to the gasdynamic time $\tau_d(g_M) = \tau_*^{gd}$, which depends weakly on g in this size range. The last relation can be written as $\tau_d(g_M)/\tau^k = \mu_*^{-1}$. Thus, formation of clusters with sizes greater than g_M occurs in gasdynamic times, when the size distribution function has already been formed. In this aspect, the clusters of size g_M can be considered as the largest ones formed in the nucleation process. The existence of a maximum cluster size confirms the correctness of both the boundary conditions at $g \longrightarrow \infty$ and the equation for monomers (14) used for completion of the quasi-chemical model of condensation [1, 12]. This point will be considered in more detail in the next section.

FLUX j

Consider now the flux of g-mers j_g which is an important characteristic of the nucleation process. It can be written in the form [12]

$$j(g, t) = j^s \tilde{j}(g, t),\ \tilde{j}(g, t) = \tilde{n}_g\left(1 - D_g \frac{n_g^s}{j^s}\frac{\partial \ln \tilde{n}_g}{\partial g}\right). \quad (29)$$

In what follows, we consider a normalized flux $\tilde{j}(g, t)$.

As noted above, in the subcritical and near-critical regions, the time it takes the system to reach its quasi-steady-state value is less than τ^k. Then, for times greater than the kinetic time, we will obtain for the quasi-steady-state flux $\tilde{j}^{qs(a)}(g, t)$ in the subcritical region ($g \le g_-$) the expression

$$\tilde{j}^{qs(a)}(g) = \tilde{n}^{qs(a)}(g)\left(1 + \sqrt{\pi}\varepsilon g_* \frac{D_g}{D_*}\frac{n_g^s}{n_*^e}\frac{\partial \tilde{\mu}_g \tau_a(g)/\tau^k}{\partial g}\right). \quad (30)$$

From these relationships, it follows that, because of the large value of the ratio n_g^s/n_*^e, even at small $\tilde{\mu}_g$ values the flux of g-mers in the subcritical region can substantially deviate from the steady-state value. In the near-critical region we have

$$\tilde{j}^{qs(b)}(g) = \tilde{n}^{qs(b)}(g)\left[1 - \xi\frac{n_g^s}{n_*^s}(1 - \sqrt{1 + \xi^{-2}\tilde{\mu}_*})\right]. \quad (31)$$

In particular, for clusters of critical size $g = g_*$, the quasi-steady-state flux will have the form

$$\tilde{j}_*^{qs} = (1 - \xi(1 - \sqrt{1 - \xi^{-2}\tilde{\mu}_*})) \times \exp(-\tilde{\mu}_*\tau_b/\tau^k + 2\xi(1 - \sqrt{1 + \xi^{-2}\tilde{\mu}_*})/\sqrt{\pi}). \quad (32)$$

For the supercritical region, as long as $\tau^k/\tau_c > \mu_1$, we get

$$\tilde{j}^{qs(c)}(g) = \tilde{n}^{qs(c)}\left[1 - \varepsilon\tilde{\mu}_* \frac{2}{3\sqrt{\pi}}\frac{D_g}{D_*}\frac{n_g^s}{n_*^e}\frac{(g/g_*)^{-1/3}}{(1 - (g/g_*)^{1/3})}\right] \simeq \tilde{n}^{qs(c)}(g). \quad (33)$$

In the supercritical region, $\tau_c(g)/\tau^k > 1$; therefore, the dependence on parameter $\tilde{\mu}_g$, contained practically only in $n^{qs(c)}(g)$, is enhanced.

Thus, it turns out that, at different g, variations of gasdynamic parameters produce different deviations of the flux of g-mers from its steady-state value. In the subcritical region, especially for small sizes, the dependence on $\tilde{\mu}_g$ is greater and the flux deviation can be large even for weak gasdynamic sources in Eq. (7). In the near-critical region, the dependence on $\tilde{\mu}_g$ is significant only if this quantity is not small. In the supercritical region, dependence on $\tilde{\mu}_g$ where $g < g_m(t)$ is enhanced, whereas the flux is totally absent in the opposite case.

It should also be noted that non-steady-state effects can both increase and decrease the flux depending on the sign of $\tilde{\mu}_g$.

Consider again Eq. (14). As the steady-state is approached, $j_g \longrightarrow$ const and the integral in the right-hand side of this equation should have been diverging. The existence of the maximum size makes it possible to set the upper limit of the integral at g_M and thus eliminate the problem of divergence. In a discrete description (quasi-chemical model), the corresponding replacement should be done in summing up the fluxes. Note that this problem has not been discussed in the literature, as well as the problem of the consecutive derivation of the condensation equations (for the condensate mass fraction) from the equations of the quasi-chemical model. With increasing g_m, the condition of infinitesimal μ_1 will be violated, resulting in drastic changes in both n_1 and the distribution function n_g, which should be interpreted as a jump of condensation.

INDUCTION TIME t_{ind} FOR THE FLUX

Of particular interest is the time lag for the formation of the quasi-steady-state flux in dimensional space. In the literature [4, 5, 8], only the time in which the flux attains its steady-state value was considered. The value calculated below characterizes the attainment of the quasi-steady-state flux value and represents a more general characteristic because the steady-state induction time can be observed only if the inequality $\mu_g \ll 1$ holds during the whole process of nucleation. The quantity defined here acquires its steady-state value when $\mu_g \longrightarrow 0$.

We will define the induction time as the coordinate of the point at which the asymptote of the total flux $N(g, t) = \int_0^t d\,tj(g, t)$ intersects the time axis (at large time scale, the partial flux $j(g, t)$ becomes constant and, consequently, N(g, t) becomes linear in time):

$$\lim_{t \to \infty} N(g, t)/[t - t_{\text{ind}}] = j_g^{gs}. \tag{34}$$

In calculating t_{ind} for the near-critical region, we use an integral representation (54) of the distribution function (see Appendix 2)

$$\tilde{n}^{(b)}(z, t) = n^{qs(b)}(g) - f(g, \mu)I(t),$$

$$f(g, \mu) = \frac{\xi(z+1)}{\pi}\exp\left(-\frac{\tilde{\mu}_* \tau_b}{\tau^k} + \frac{2\xi(1+z)}{\sqrt{\pi}}\right)\Theta(t - \tau_b),$$

$$I(t) = \int_0^{\tau^k/(t-\tau_b)} dx x^{-1/2}\exp\left(-\frac{a_1}{x} - a_2 x\right), \tag{35}$$

$$a_1 = 1 + \xi^{-2}\tilde{\mu}_*, \quad a_2 = (\xi(z+1))^2/\pi,$$

where parameter ξ is defined by formula (19).

Using definition (2) of the flux and identity (29), we write the total flux as

$$N(g, t) = (t - \tau_b)\left(j^s n^{qs(b)}(g) - D_g n_g^s \frac{\partial}{\partial g} n^{qs(b)}(g)\right)$$

$$- j^s f(g, \mu)\int_{\tau_b}^{t} I(t)dt + D_g n_g^s \frac{\partial}{\partial g}\left(f(g, \mu)\int_{\tau_b}^{t} I(t)dt\right).$$

The coefficient of $(t - \tau_b)$ is nothing less than the quasi-steady-state flux $j^{sq(b)}$. At $t \longrightarrow \infty$ the total flux is approximated by a linear function of time (34), and the following expression for the time lag in the near-critical region is obtained

$$t_{\text{ind}}^{(b)}(g) = \tau_b + \frac{\tau^k \xi}{\sqrt{\pi(1 + \xi^{-2}\tilde{\mu}_*)}}\left[(z+1)\right.$$
$$\left. - \sqrt{\pi}\frac{D_g}{D_* n_*^e}\left(1 - \xi\frac{n_g^s}{n_*^s}(1 - \sqrt{1 + \xi^{-2}\tilde{\mu}_*})\right)^{-1}\right]. \tag{36}$$

In particular, for the steady-state case we get

$$t_{\text{ind}}^{(b)}(g)|_{\tilde{\mu}_* = 0} = \tau_b + \frac{\tau^k \xi}{\sqrt{\pi}}\left[(z+1) - \sqrt{\pi}\frac{D_g}{D_*}\frac{n_g^s}{n_*^e}\right]. \tag{37}$$

In a similar manner, for the supercritical region we obtain

$$t_{\text{ind}}^{(c)}(g) = \tau_c + \frac{2\xi\tau^k}{\sqrt{\pi(1 + \xi^{-2}\tilde{\mu}_*)}}. \tag{38}$$

For the subcritical region, directly from (23) (at $i = a$), we obtain

$$t_{\text{ind}}^{(a)}(g) = \tau_a(g). \tag{39}$$

In the subcritical region, the induction time coincides with the time lag for a cluster of corresponding size (and is therefore small) and is independent of μ_g; i.e., it is independent of the rates of the gasdynamic processes. In the near-critical and especially in the supercritical regions where $\tau_d \gg \tau^k$, t_{ind} also differs little from the corresponding individual time lags and the dependence on μ becomes weaker with increasing g. Non-steady-state effects can both decrease and increase the induction time, depending on the sign of $\tilde{\mu}_g$.

CALCULATIONS AND DISCUSSION

Results of calculations for a number of characteristics of the nucleation process are given. In all calculations, the same values of the basic parameters are taken as in [4]: the critical size $g_* = 52$, and the parameter responsible for the influence on the formation energy of surface effects (see (26)) is $a = 11.81$. In this case, we have $\Delta = 12.2$, $\varepsilon = 0.23$, $g_- = 39$, $g_+ = 64$, the kinetic time $\tau^k = 2.6 \times 10^{-6}$ s, and the time lag in the near-critical region $\tau_b = 2.4 \times 10^{-6}$ s.

The calculated dependences of the partial time lag of the formation of the quasi-steady-state distribution function on the cluster size are given in Fig. 1. From (28), it follows that in the supercritical region at $g/g_* > 3\ln(9/\varepsilon^2)$, the time lag assumes an asymptotic dependence on g: $\tau_b(g)/\tau^k \sim (2\pi)(g/g_*)^{1/3}$.

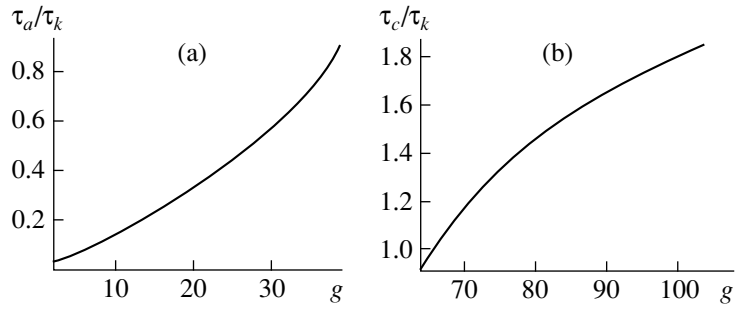


Fig. 1. Dependence of the time lag τ_d for relaxation to the quasi-steady-state value on cluster size (a) for the subcritical region ($\tau^{(a)}$) and (b) for the supercritical region ($\tau^{(c)}$).

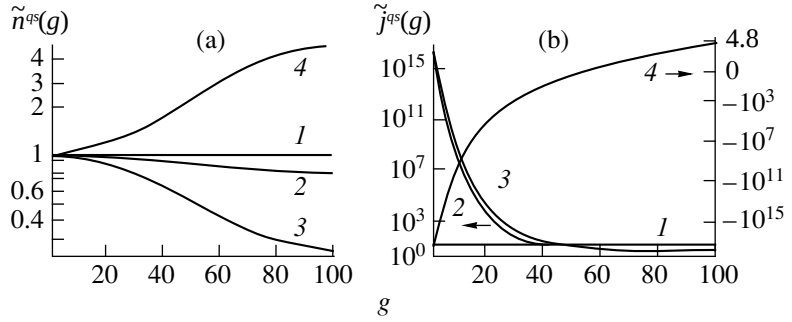


Fig. 2. Variation with cluster size g of (a) the normalized quasi-steady-state distribution $\tilde{n}^{qs}$ and (b) the normalized quasi-steady-state flux $\tilde{n}^{qs}$ at different values of parameter $\tilde{\mu}$: (*1*) 0; (*2*) 0.1; (*3*) 0.5; and (*4*) –0.5. Curves *1* combine two graphs with different scales for $g \le 30$ and $g < 30$.

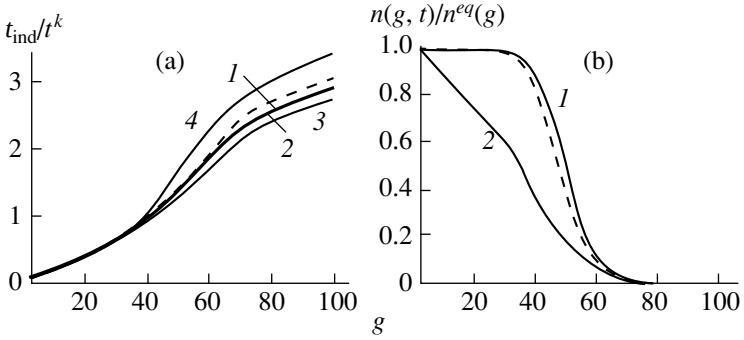


Fig. 3. Dependence on cluster size of (a) t_{ind}/t^k and (b) the distribution function $n(g, t)$. Dashed line indicates data from [4]; $\tilde{\mu}$ = (*1*) 0, (*2*) 0.5.

Figure 2 shows the variation with cluster size of the normalized quasi-steady-state distribution $\tilde{n}^{qs}(g)$ and the normalized quasi-steady-state flux $\tilde{j}^{qs}(g) = j^{qs}(g)/j^s$ at different values of parameter μ.

With increasing cluster size, the quasi-steady-state distribution exponentially but relatively slowly deviates from its steady-state value and in the supercritical region, the rate of deviation decreases. The quasi-steady-state flux in dimension space behaves in a different way. At positive μ, a rapid growth of small clusters takes place that considerably exceeds the steady-state flux. In the subcritical region, the flux sharply drops (by several orders of magnitude) to its steady-state value and then, in the supercritical region, weakly diminishes and remains below the steady-state flux. For this reason, in the supercritical region, the quasi-steady-state concentration of clusters turns out to be lower than the steady-state concentration. Thus, a considerably higher flux is needed in the subcritical region than in the supercritical region in order to compensate for the variation in gasdynamic parameters and to ensure a steady-state character of the process. In some respects the situation is opposite if μ is negative; small clusters rapidly evaporate (the flux turns negative and

large in absolute magnitude), whereas the supercritical clusters grow slowly at a rate only several times higher than the steady-state rate. This makes the distribution function in the supercritical region higher than its steady-state value. This cannot continue for long because of the lack of clusters of near-critical size, which provide flux into the supercritical region. This is the physical reason for the instability of the process at negative μ at too high absolute values.

Shown in Fig. 3 are the induction times (time lags for the quasi-steady-state flux formation) at different values of μ ($\left|j^{qs}\right|_{\tilde{\mu}=0} = j^s$) calculated with the use of analytical expressions (36), (38), and (39), and the normalized distribution functions at the time moment t = 0.009 s. In both cases, good agreement has been obtained with the calculation of similar parameters for which expressions in the form of quadratures have been proposed in [4] at $\mu = 0$. It is also seen from the calculations that, with increasing influence of the gasdynamic effects in the supercritical region at positive μ, the induction time slowly decreasing and, at negative μ, is increasing. The graph on the right indicates that by changing the rate of variation of gasdynamic parameters, it is possible to control the cluster size distribution function.

CONCLUSIONS

In the present work, by separating fast and slow variables, a non-steady-state solution (in the form of quadratures) of the Zel'dovich equation for the cluster size distribution function with times exceeding the kinetic time τ^k given by relationship (10) has been obtained. The solution is applicable in the entire range of cluster sizes. For times $(t - \tau_d) \gg \tau^k$, corresponding asymptotic solutions have been obtained. Times of relaxation to the quasi-steady-state distribution and partial time lags of cluster formation have been derived and analyzed. The dependence of these times on the cluster size has been analyzed. Expressions for the quasi-steady-state flux and the induction time have been obtained. The accuracy of these calculations is confirmed by comparison with numerical calculations by other authors.

Analysis of the behavior of the time lag yields the following results. In the subcritical region, the induction time is less than the kinetic time, and the evolution of both the distribution function and all other characteristics of these clusters is determined by the value of τ^k. This greatly simplifies the description of the evolution of subcritical clusters. On the other hand, for most applications, $\mu_1 < 1$ and quasi-steady-state values of the parameters can be used. All of these quantities have been calculated in this work and their form is substantially more simple than in [4, 6, 8].

The situation differs in principle in the supercritical region where the induction time is a determining factor and the quasi-steady-state behavior is disturbed starting with some cluster size. The delay in evolution of the distribution function causes the appearance of a maximum cluster size, which makes a correct interpretation possible for the monomer equation.

The performed analysis provided a clearer picture of the evolution of the distribution function towards its quasi-steady-state value and identifies the main parameter controlling the shape of the quasi-steady-state distribution.

APPENDIX 1

$$\Lambda^{(a)}(g, p) = -\frac{\mu_1}{\tau^k}\int_1^g dg' \frac{\tilde{n}(0, g')}{\dot{g}(g') + g\dot{g}_*/g_*} \times \exp(-(\mu_1 p + \tilde{\mu}_g)(\tau_a(g) - \tau_a(g'))/\tau^k), \quad (40)$$

$$\Lambda^{(b)'}(g, p) = \frac{-\mu_1}{\sqrt{\pi}\sqrt{1 + \mu_1 p + \tilde{\mu}_*}}\int_{-1}^z dz' \times \exp(2\xi(1 - \sqrt{1 + \xi^{-2}(\mu_1 p + \tilde{\mu}_*)})(z - z')/\sqrt{\pi})\tilde{n}(0, z), \quad (41)$$

$$\Lambda^{(c)'}(g, p) = -\frac{\mu_1}{\tau^k}\int_{g_+}^g dg' \frac{\tilde{n}(0, g')}{\dot{g}(g') + g'\dot{g}_*/g_*} \times \exp(-(\mu_1 p + \tilde{\mu}_g)(\tau_c'(g) - \tau_c'(g'))/\tau^k), \quad (42)$$

$$\Lambda^{(b)}(g, p) = \frac{\mu_1}{\tau^k}\exp(2\xi(1 - \sqrt{1 + \xi^{-2}(\mu_1 p + \tilde{\mu}_*)})) \times (1 + z)/\sqrt{\pi})\int_1^{g_-} dg' \frac{\tilde{n}(0, g')}{\dot{g}(g') + g'\dot{g}_*/g_*} \times \exp(-(\mu_1 p + \tilde{\mu}_g)(\tau_a(g) - \tau_a(g'))/\tau^k) - \Lambda^{(b)'}, \quad (43)$$

$$\Lambda^{(c)}(g, p) = \exp(-(\mu_1 p + \tilde{\mu}_*)\tau_c'(g)/\tau^k) \times \Lambda^{(b)}(g_+, p) + \Lambda^{(c)}(g, p)'. \quad (44)$$

Without derivation, we give the results of the inverse Laplace transforms of the functions $\Lambda^{(i)}(g, p) \longrightarrow \lambda^{(i)}(g, t)$, $t = a, b, c$.

In calculating the integral with respect to g', the property of the δ-function will be used where $\delta(t - \tau(g')) = d\tau/dg'^{-1}\delta(t(g') - g')$, $\dot{g} < 0$, and expression (16) will be used for $\tau_a(g)$. The result will be (time dimensional)

$$\lambda^{(a)} = -\tilde{n}(0, g_0(t))\exp(-\tilde{\mu}_* t/\tau^k). \quad (45)$$

This equality is true if g_0 belongs to the interval $[1, g]$; otherwise, $\lambda^a = 0$. Here, g_0 is the value of g' for which $t - \tau_a(g) - \tau_a(g') = 0$ ($g_0 = g$ at $t = 0$). Thus, the maximum time for which $\lambda^{(a)}$ is nonzero is $\tau_a(g)$, i.e.,

the relaxation of the initial distribution is determined not by the exponent but by the first cofactor in (45).

Expressions for $\lambda^{(b)}$ and $\lambda^{(c)}$ have been obtained assuming that $t - \tau_d \gg t^k$.

$$\lambda^{(b)} = -\frac{z+1}{\pi\tau^k}\int_1^{g_-} dg' \frac{\tilde{n}(0,g')}{\dot{g}(g') + g'\dot{g}_*/g_*}\left(\frac{t-\hat{\tau}(g')}{\tau^k}\right)^{-3/2}$$

$$\times \exp\left(\frac{\hat{\tau}(g') - (\tilde{\mu}_* + 1)t}{\tau^k} + \frac{(z+1)^2\tau^k}{\pi(\hat{\tau}(g') - t)} + \frac{2}{\sqrt{\pi}}(z+1)\right)$$

$$+ \frac{1}{\pi}\sqrt{\frac{\tau^k}{t}}\exp\left(\frac{\tilde{\mu}_* + 1}{\tau^k}t\right)\int_{-1}^{z} dz'\tilde{n}(0,z') \qquad (46)$$

$$\times \exp\left(\frac{2}{\sqrt{\pi}}(z-z') - \frac{\tau^k}{\pi t}(z-z')^2\right),$$

$$\hat{\tau}(g') = \tau_a(g_-) - \tau_a(g');$$

i.e., at $t > \tau = \max\{\tau^k, \tau_b\}$, $\lambda^{(b)}$ is negligibly small,

$$\lambda^{(c)} = l^{(c)} - \frac{2}{\pi\tau^k}\int_1^{g_-} dg' \frac{\tilde{n}(0,g')}{\dot{g}(g') + g'\dot{g}_*/g_*}$$

$$\times\left(\frac{t - (\tau_c(g) - \tau_a(g'))}{\tau^k}\right)^{-3/2}$$

$$\times \exp\left(\frac{(\tau_c(g) - \tau_a(g')) - (\tilde{\mu}_* + 1)t}{\tau^k}\right. \qquad (47)$$

$$\left. + \frac{4}{\sqrt{\pi}}\left(1 + \frac{\tau^k}{(\tau_c(g) - \tau_a(g')) - t}\right)\right)$$

$$-\frac{1}{\pi}\int_{-1}^{1} dz'\tilde{n}(0,z')\left(\frac{t-\tau_c'(g)}{\tau^k}\right)^{-1/2}\exp\left(\frac{\tau_c'(g) - (\tilde{\mu}_* + 1)t}{\tau^k}\right.$$

$$\left. + \frac{(1-z')^2\tau^k}{\pi(\tau_c'(g) - t)} + \frac{2}{\sqrt{\pi}}(1-z')\right),$$

where the quantity $l^{(c)}$ is $l^{(c)} = \tilde{n}(0, \tilde{g}_0(t))\exp(-\tilde{\mu}_* t/\tau^k)$ if $\tilde{g}_0$ belongs to the interval $[g_+, g]$ and is otherwise zero; $\tilde{g}_0$ is the value of g' for which $t - (\tau_c(g) - \tau_c(g')) = 0$, i.e., the maximum time for which $l^{(a)}$ is nonzero is $\tau_c(g)$; thus, at $t > \tau_c(g)$, $\lambda(c)$ can be neglected.

Asymptotics for $\tilde{n}^{(b)}(g, t)$ and $\tilde{n}^{(c)}(g, t)$ have been derived under more rigorous conditions; therefore, we will not take into account further the effect of the initial distribution function.

APPENDIX 2

INVERSE LAPLACE TRANSFORMS OF THE FUNCTIONS $N^{(b)}(z, p)$ AND $N^{(c)}(z, p)$

Consider the function $N^{(b)}(z, p)$

$$N^{(b)}(z,p) = p^{-1}\exp\left(2\pi^{-1/2}(1 - \sqrt{1 + \mu_1 p + \tilde{\mu}_*})\right.$$

$$\left. \times (1+z) - (\mu_1 p + \tilde{\mu}_*)\frac{\tau_b}{\tau^k}\right). \qquad (48)$$

Let us introduce the following notation: $\alpha = 2\sqrt{\mu_1/\pi}(z+1)$, $b = (1 + \tilde{\mu}_*)/\mu_1$, $\gamma = \mu_1\tau_b/\tau^k$, and $\beta = \tilde{\mu}_*\tau_b/\tau^k - 2\pi^{-1/2}(z+1)$; then $N^{(b)}(p)$ shall have the form $N^{(b)}(p) = F(p)G(p)$, where $F(p) = p^{-1}\exp(-\beta - \gamma p)$ and $G(p) = \exp(-\alpha\sqrt{p+b})$. Using the convolution transformation rule [19] we get for $\tilde{n}^{(b)}(t')$

$$\tilde{n}(t') = (fg)(t') = \int_0^{t'} d\tau f(\tau)g(t'-\tau), \qquad (49)$$

where $f(t')$ and $g(t')$ are inverse Laplace transforms for functions $F(p)$ and $G(p)$, respectively.

For $f(t')$ and $g(t')$ we find

$$f(t') = \exp(-\beta)\Theta(t'-\gamma),$$

$$g(t') = \exp(-t'b)\frac{\alpha}{2\sqrt{\pi}}\frac{1}{t'^{3/2}}\exp\left(-\frac{\alpha^2}{4t'}\right). \qquad (50)$$

Finally, $\tilde{n}^{(b)}(t')$ will assume the following integral representation

$$\tilde{n}^{(b)}(t') = \frac{\alpha}{2\sqrt{\pi}}\exp(-\beta)\int_0^{t'} d\tau\Theta(\tau-\gamma)(t'-\tau)^{-3/2}$$

$$\times \exp\left(-(t'-\tau)b - \frac{\alpha^2}{4(t'-\tau)}\right). \qquad (51)$$

Substituting $x = \xi^{-2}\mu_1(t' - \tau)^{-1}$ into Eq. (51) and transforming to dimensional time $t' - \gamma = \mu_1(t - \tau_b)/\tau_1^{qd}$, we get

$$\tilde{n}^{(b)}(z,t) = \frac{\xi(z+1)}{\pi^{3/2}}\exp(-\beta)\Theta(t-\tau_b)$$

$$\times \int_{q/\lambda}^{\infty} dx x^{-1/2}\exp\left(-\frac{a_1}{x} - a_2 x\right), \qquad (52)$$

where $\lambda = (t - \tau_b)/\tau^k$; $a_1 = (1 + \xi^{-2}\tilde{\mu}_*)$; and $a_2 = (\xi(z + 1))^2/\pi$.

As we are interested in the limit at large t when $\lambda > 1$, it is convenient to rewrite the integral in (52) as

a difference $\int_{1/\lambda}^{\infty} = \int_0^{\infty} - \int_0^{1/\lambda}$. The first integral can be expressed through a modified second-order Bessel function $K_{1/2}(x)$

$$\int_0^{\infty} dx x^{-1/2} \exp\left(-\frac{a_1}{x} - a_2 x\right) = 2\left(\frac{a_1}{a_2}\right)^{1/4} K_{1/2}(2\sqrt{a_1 a_2}) = \sqrt{\frac{\pi}{a^2}} \exp(-2\sqrt{a_1 a_2}). \quad (53)$$

Then, the distribution function can be represented in the following form

$$\tilde{n}^{(b)}(z, t) = \Theta(t - \tau_b)\tilde{n}^{qs(b)}(g) \times \left[1 - \frac{\xi(z+1)}{\pi}\exp\left(2\xi\sqrt{\frac{1+\xi^{-2}\tilde{\mu}_*}{\pi}}(1+z)\right) \times \int_0^{\tau^k/(t-\tau_b)} dx x^{-1/2}\exp\left(-\frac{a_1}{x} - a_2 x\right)\right], \quad (54)$$

where the quasi-steady-state distribution $\tilde{n}^{qs(b)}(g)$ is given by (24).

Let us consider the remaining integral at $\lambda \longrightarrow \infty$, i.e., at $t - \tau_b \gg \tau^k$, and rewrite it in the form ($y = \lambda x$)

$$\int_0^{1/\lambda} dx \ldots = \frac{1}{\sqrt{\lambda}}\int_0^1 dy f(y)\exp(\lambda F(y, \lambda)), \quad F(y, \lambda) = -\frac{a_1}{y} - \frac{a_2 y}{\lambda^2}, \quad f(y) = y^{-1/2}. \quad (55)$$

The function $F(u, \lambda)$ has a maximum at point $y_m(\lambda) = \lambda\sqrt{a_1/a_2}$; therefore, at $\lambda \geq \sqrt{a_2/a_1}$ and integration limits [0; 1], $F(y, \lambda)$ reaches its maximum value at $y_0 = 1$. Then, the main term of the integral (55) asymptote at $\lambda \longrightarrow \infty$ will have the form

$$\int_0^1 dy f(y) \approx -\frac{f(y_0)\exp(\lambda F(y, \lambda))}{\lambda F'_y(y_0, \lambda)}, \quad (56)$$

where $F(y_0, \lambda) = -a_2/\lambda^2 - a_1$, $f(y_0) = 1$, and $F'_y(y_0, \lambda) = -a_2/\lambda^2 + a_1$.

Ultimately, for $\tilde{n}^{(b)}(z, t)$ at $t \longrightarrow \infty$ we get the asymptotic expression (23) (at $i = b$). In a similar manner, an integral expression for $\tilde{n}^{(c)}(z, t)$ will have the form

$$\tilde{n}^{(c)}(g, t) = \tilde{n}^{qs(c)}(g)\Theta(t - \tau_c(g)) \times \left[\left(1 - \frac{2\xi}{\pi}\right)\exp\left(4\xi\sqrt{\frac{1+\xi^{-2}\tilde{\mu}_*}{\pi}}\right) \times \int_0^{(\tau^k/t-\tau_c(g))} \frac{dx}{\sqrt{x}}\exp\left(-\frac{(1+\xi^{-2}\tilde{\mu}_*)}{x} - \frac{4\xi^2 x}{\pi}\right)\right] \quad (57)$$

with a quasi-steady-state distribution $\tilde{n}^{qs(c)}(g)$ (24) given by (23) ($i = c$).

REFERENCES

1. F. F. Abraham, *Homogeneous Nucleation Theory* (Academic, New York, 1974).
2. F. M. Kuni, in *Problems of Theoretical Physics III* (Leningr. Gos. Univ., Leningrad, 1988).
3. F. M. Kuni, Preprint No. ITP-84-178 E (Kiev, 1984).
4. B. Shizgal and J. C. Barrett, J. Chem. Phys. **91** (10), 6505 (1989).
5. D. T. Wu, J. Chem. Phys. **97**, 2644 (1992).
6. V. A. Shneidman and M. C. Weinberg, J. Chem. Phys. **97**, 3621 (1992).
7. D. Lippman and W. C. Schive, J. Chem. Phys. **92** (7), 4426 (1990).
8. V. A. Shneidman, Zh. Tekh. Fiz. **57** (1), 131 (1987) [Sov. Phys. Tech. Phys. **32**, 76 (1987)].
9. V. A. Shneidman, Phys. Rev. A **44**, 2609 (1991).
10. Yu. E. Gorbachev and V. Yu. Kruglov, Pis'ma Zh. Tekh. Fiz. **16** (8), 1 (1990) [Sov. Tech. Phys. Lett. **16**, 283 (1990)].
11. Yu. E. Gorbachev and V. Yu. Kruglov, Prikl. Mekh. Tekh. Fiz., No. 1, 146 (1991).
12. Yu. E. Gorbachev, Report on 20th International Symposium on Rarefied Gas Dynamics, Beijing, 1996 (Peking Univ. Press, Beijing, 1997), p. 853.
13. E. G. Kolesnichenko, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 3, 96 (1981).
14. N. G. van Kampen, Phys. Rep. **124**, 69 (1985).
15. N. G. van Kampen, J. Stat. Phys. **46**, 709 (1987).
16. Ya. I. Frenkel', *Collection of Selected Works* (Akad. Nauk SSSR, Moscow, 1959), Vol. 3, pp. 358–407.
17. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
18. S. A. Lomov, *Introduction to the General Theory of Singular Perturbance* (Nauka, Moscow, 1981).
19. M. V. Fedoryuk, *Saddle Point Approximation* (Nauka, Moscow, 1977).

Translated by B. Kalinin

Technical Physics, Vol. 45, No. 12, 2000, pp. 1549–1552. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 39–42.

SOLIDS

Twinning in Boron-Ion-Irradiated Bismuth Single Crystals Subjected to Mutually Orthogonal Electric and Magnetic Fields

O. M. Ostrikov

Mozyr State Pedagogical Institute, Mozyr, Gomel oblast, 247760 Belarus

Received November 1, 1999

Abstract—The effect of electric pulses (duration 10^{-5} s and current density up to 60 A/mm^2) and an external magnetic field (0.2 T) on the behavior of wedgelike twin ensembles in boron-irradiated bismuth single crystals was studied. The irradiation energy and dose were, respectively, 25 keV and 10^{17} ion/cm^2. When the crystals were exposed to the magnetic field, both irradiation and pulse application enhanced the mobility of twinning dislocations. An equilibrium condition for twinning dislocations in the irradiated material simultaneously subjected to the electric and magnetic fields was derived. © *2000 MAIK "Nauka/Interperiodica".*

Twinning in bismuth single crystals, both nonirradiated [3–5] and irradiated by various ions [1, 2], is today the subject of much investigation. Of interest is the influence of mutually orthogonal pulse electric [2, 4] and permanent magnetic [5] fields applied to these crystals on the process of plastic deformation by twinning. The urgency of this investigation stems from the fact that both ion irradiation and field application allow researchers to effectively control the physical properties of solids. In this work, we studied twinning in bismuth single crystals subjected to a variety of external effects. It is hoped that our results will serve to extend a set of methods for controlling the ductility of structure materials. Physically, the interaction and kinetics of twinning dislocations in irradiated crystals under electric and magnetic fields is of fundamental importance.

EXPERIMENTAL

Single crystals were Bridgman-grown from 99.999%-pure raw material. The respective initial densities of basal and pyramidal dislocations were 10^5 and 10^3 cm^{-2}. Samples measuring $4 \times 5 \times 10$ mm were cut along the (111) cleavage plane.

The as-cleaved (111) surface of single-crystal bismuth were irradiated by boron ions with an energy of 24 keV and a dose of 10^{17} ion/cm^2. The samples were deformed by concentrated load, i.e., by the diamond pyramid of a conventional PMT-3 microhardness meter. The indentation load was varied from 5×10^{-2} to 30×10^{-2} N, and the indentation time was 5 s. In parallel with indentation, an isolated near-triangular current pulse of duration 10^{-5} s was applied to the crystal. The current density in the pulse was varied from zero to 60 A/mm^2. The pulse amplitude was controlled with a GI-3M pulse generator; and its shape, by an S1-17 oscilloscope. The crystals were deformed in an external magnetic field of 0.2 T, orthogonal to the current density direction (Fig. 1).

During microscopic examination, we measured the mean number of twins near the indent, as well as the length and width of the mouth of individual twins.

RESULTS AND DISCUSSION

Six to eight wedge twins are usually observed near a Vickers diamond pyramid indent on the (111) surface of bismuth single crystals. Statistical characteristics (the mean length and width of twins and their mean number) of this twin ensemble remain more or less constant [5]. They depend only on the conditions under which the material is deformed: material purity, temperature, presence of magnetic and electric fields, and surface defects. Therefore, from the evolution of twin ensembles, one can judge the ductile properties of twin materials subjected to external actions.

Dependences of the mean twin length L on indentation load P (Fig. 2) show that the mobility of twinning dislocations (hence, their path) in bismuth single crystals increases under both boron irradiation and the simultaneous action of electric and magnetic fields.

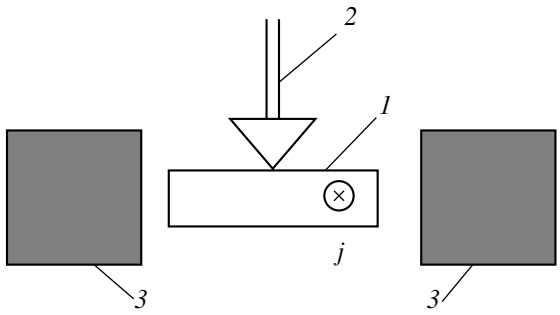


Fig. 1. Indentation in external electric and magnetic fields: (*1*) sample, (*2*) indenter, and (*3*) magnet.

1063-7842/00/4512-1549$20.00

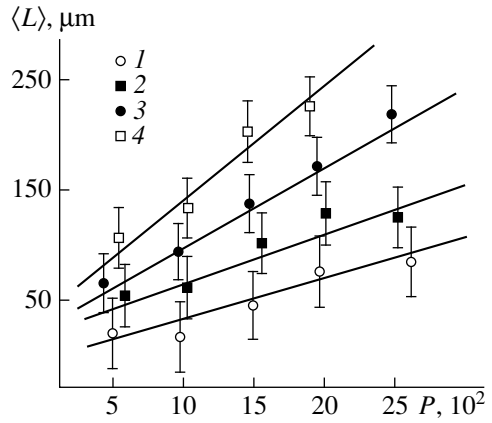


Fig. 2. Mean twin midline length vs. indentation load. (*1*) Nonirradiated sample, (*2*) boron-irradiated sample, (*3*) nonirradiated sample subjected to current pulses of density 60 A/mm^2 and an external magnetic field of 0.2 T, and (*4*) irradiated sample subjected to current pulses of density 60 A/mm^2 and an external magnetic field of 0.2 T.

This fact can be explained by stresses due to the implanted impurity and current-induced pinch effect (electron plasma confinement).

Let τ_{xy}^i be implant-induced cleavage stresses and τ_{xy}^j and τ_{xy}^B be those induced by electric and magnetic fields, respectively. Then, at the instant of electrical pulse application, the crystal is subjected to the sum of the stresses

$$\tau_{xy}^e = \tau_{xy}^i + \tau_{xy}^j + \tau_{xy}^p + \tau_{xy}^B, \quad (1)$$

where τ_{xy}^p is a cleavage stress induced by a concentrated load.

These stresses exert the force per unit length of a twinning dislocation

$$f_x = b\tau_{xy}^e, \quad (2)$$

where b is the Burgers vector.

Accordingly, the normal components of the strain tensor,

$$\sigma_{xx}^e = \sigma_{xx}^i + \sigma_{xx}^j + \sigma_{xx}^p + \sigma_{xx}^B, \quad (3)$$

$$\sigma_{yy}^e = \sigma_{yy}^i + \sigma_{yy}^j + \sigma_{yy}^p + \sigma_{yy}^B, \quad (4)$$

cause climb forces to arise:

$$f_x = -b\sigma_{yy}^e, \quad f_y = -b\sigma_{xx}^e, \quad (5), (6)$$

In their presence, crystal lattice defects are easier to overcome by dislocations. Forces (2), (5), and (6) extend the twin length, which was observed in experiments. Expressions for stresses acting on twinning dislocations are given below, where equilibrium conditions for the dislocations in the presence of the external effects are derived.

Equations for dislocation equilibrium at each of the twin boundaries (coordinate axes with respect to twin boundaries are shown in Fig. 3) are as follows [6]:

$$\begin{aligned} -a\tau_0 \int_{-L}^{L} \frac{\rho(t)dt}{t-x} + \tau_1\{h(x)\eta_2''(x) - 2a\rho(x)\eta'(x)\} \\ = \sigma^e(x,0) + \eta_1(x)\left(\frac{d\sigma^e}{dy}\right)_0 + S_1(x), \\ -a\tau_0 \int_{-L}^{L} \frac{\rho(t)dt}{t-x} - \tau_1\{h(x)\eta_1''(x) - 2a\rho(x)\eta'(x)\} \\ = \sigma^e(x,0) + \eta_2(x)\left(\frac{d\sigma^e}{dy}\right)_0 + S_2(x). \end{aligned} \quad (7)$$

Here, a is the interatomic distance in the direction perpendicular to the twin plane, constants τ_0 and τ_1 are given by

$$\tau_0 = \frac{2\alpha\lambda_{1111}}{\pi(m_1+m_2)}, \quad \tau_1 = \alpha\lambda_{1111}. \quad (8)$$

Here, α is the surface tension at the twin–host crystal interface, λ_{1111} is the Lamé coefficient, m_1 and m_2 depend on the elastic modula of a medium, $\rho(x) = -h'(x)$, $h(x)$ is the twin thickness at each point of the $0X$ axis ($h(x) = \eta_1(x) - \eta_2(x)$), $\eta = \eta_1(x)$ is the equation for the upper twin boundary, $\eta = \eta_2(x)$ is the equation for the lower twin boundary [in this case, $\eta_1(x) \geq \eta_2(x)$], and the function $\eta(x) = 1/2(\eta_1(x) + \eta_2(x))$ locates the twin midline.

Relationships (7) were obtained under the assumption that the quantities of $h(x)$ and $\eta(x)$ are small. It is valid if a twin is thin. Let $S_1(x)$ and $S_2(x)$ be inelastic forces at the upper and lower twin boundaries, respectively. Their difference may stem from the presence of barriers for dislocations on one of them, various conditions for dislocation origination, etc. [6] If these forces are equal, ($S_1(x) = S_2(x) = S(x)$), and the twin asymmetry is due to external stress gradients only, one obtains the basic approximation with respect to the parameter h/L [6]:

$$-a\tau_0 \int_{-L}^{L} \frac{\rho(t)dt}{t-x} = \sigma^e(x,0) + S(x). \quad (9)$$

In Eqs. (7) and (9), the function $\sigma^e(x, y)$ specifies the distribution of nonuniform external stresses at the twin plane; and $\partial\sigma^e(x)/\partial y$ in Eq. (7), their gradient.

As was noted, if irradiation is combined with a pulse electric current and an external magnetic field applied to the crystal, its surface is deformed by the action of the implant-induced stresses, electron plasma confined due to the pinch effect, magnetic-field-induced stresses, and concentrated load.

Implant-induced stresses in a near-surface layer of thickness $h = 2R_p + \Delta R_p$ can be determined from the relationship [1, 2]

$$\sigma^i = k(x/h), \tag{10}$$

where k is the material constant.

In the rest of the crystal, which is much thicker than the separated near-surface layer, implant-induced stresses can be put equal to zero. This, in particular, follows from (10) if h is infinitely large. Then, (10) can be recast as

$$\begin{aligned} k(x/h), \quad 0 \le x \le 2R_p + \Delta R_p, \\ \sigma^i = 0, \quad x > 2R_p + \Delta R_p. \end{aligned} \tag{11}$$

The expression for voltage distribution due to the electric current is given by [7]

$$\sigma^j = \sigma_0^j \exp(-x/l), \tag{12}$$

where l is the free path of electrons, x is the coordinate reckoned from the normal to the surface, $\sigma_0^j = Pp_F j/e$, P is the Fuchsian diffuseness parameter ($0 < P \le 1$), p_F is the Fermi momentum, j is the current density, and e is the electron charge.

The effect of the magnetic field in (7) can be included with the relationship

$$\sigma^B = \frac{\mathrm{grad}\,U(x)}{b^2}, \tag{13}$$

where $U(x)$ is the surface barrier potential [8].

The effect of the magnetic field can also be taken into account using formula (12). It is assumed that, in the presence of a magnetic field, the electron concentration in the near-surface layer changes in comparison with the case when the field is absent. This is due to the Hall effect. Whether the electron concentration will increase or decrease depends on the mutual arrangement of the vectors of the magnetic field and electric current density. A change in the concentration in (12) causes a change in j and, hence, inner stresses induced by the electric current in the crystal. This shows up as the extension or shortening of wedge twins, which was observed in [9].

Stress fields due to the concentrated load can be determined from the $H = f(\sigma)$ curve. According to [10],

$$\sigma^p = H/a, \tag{14}$$

where H is the material microhardness and $a = 3$ for bismuth.

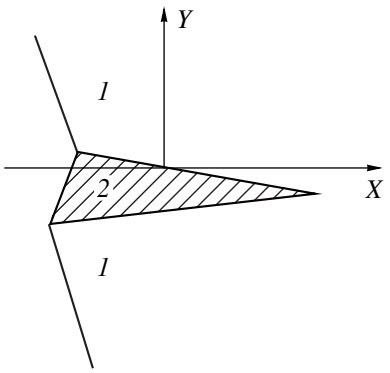


Fig. 3. Wedge twin: (*1*) host crystal and (*2*) twin lamella.

The resulting stresses in the crystals are equal to the superposition of stresses (11)–(14):

$$\begin{aligned} \sigma^e(x, 0) \\ = k(x/h) + \sigma_0^j \exp(-x/l) + H/a + \frac{\mathrm{grad}\,U(x)}{b^2} \end{aligned} \tag{15}$$

for $0 \le x \le 2R_p + \Delta R_p$ and

$$\sigma^e(x, 0) = \sigma_0^j \exp(-x/l) + H/a + \frac{\mathrm{grad}\,U(x)}{b^2} \tag{16}$$

for $x > 2R_p + \Delta R_p$.

Substituting (15) and (16) into (7) yields the equilibrium equations for twinning dislocations when the electric current passes through the irradiated crystal. In these equations, the functions $\eta_1(x)$ and $\eta_2(x)$ must be continuous at the boundary between the separated implanted layer and the rest of the crystals.

CONCLUSION

Thus, the evolution of twin ensembles in boron-irradiated (energy 25 keV, dose 10^{17} ion/cm^2) bismuth single crystals was studied. The crystals were subjected to current pulses (duration 10^{-5} s, current density to 60 A/mm^2) and an external magnetic field of 0.2 T. It was found that (1) boron ion implantation enhances the mobility of twinning dislocations and (2) current pulses in the presence of an external magnetic field improve the ductility of the bismuth single crystals, both irradiated and nonirradiated.

From the dislocation model of twinning, we derived the equilibrium equations for twinning dislocations at twin boundaries in the irradiated crystal exposed to external electric and magnetic fields.

REFERENCES

1. V. S. Savenko, V. V. Uglov, O. M. Ostrikov, and A. P. Khodoskin, Pis'ma Zh. Tekh. Fiz. **24** (8), 1 (1998) [Tech. Phys. Lett. **24**, 287 (1998)].

2. V. S. Savenko, V. V. Uglov, O. M. Ostrikov, *et al.*, Fiz. Met. Metalloved. **85** (5), 97 (1998).
3. V. S. Savenko and O. M. Ostrikov, in *Proceedings of the IV International Conference "Effect of Electromagnetic Waves on Plasticity and Strength od Materials," Voronezh, 1996*, p. 20.
4. V. S. Savenko and O. M. Ostrikov, Izv. Vyssh. Uchebn. Zaved., Chern. Metall., No. 6, 12 (1998).
5. V. S. Savenko and O. M. Ostrikov, Izv. Nats. Akad. Nauk Belarusi, Ser. Fiz.-Mat. Nauk, No. 2, 96 (1998).
6. A. M. Kosevich and V. S. Boĭko, Usp. Fiz. Nauk **104** (2), 201 (1971) [Sov. Phys. Usp. **14**, 286 (1971)].
7. A. M. Roshchupkin and I. L. Bataronov, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 3, 57 (1996).
8. A. I. Pinchuk, Author's Abstract of Candidate's Dissertation (Minsk, 1998).
9. A. I. Pinchuk, V. S. Savenko, S. D. Shavreĭ, *et al.*, in *Proceedings of the IV International Conference "Effect of Electromagnetic Waves on Plasticity and Strength of Materials," Voronezh, 1996*, p. 19.
10. S. I. Bulychev, V. P. Alekhin, and M. U. Shorokhov, Fiz. Khim. Obrab. Mater., No. 5, 69 (1979).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 12, 2000, pp. 1553–1556. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 43–46.

SOLIDS

The Switchable Component of Ferroelectric Ceramics Polarization and Pulsed Emission of Electrons

V. N. Ivanov and L. M. Rabkin

Rostov State University, pr. Stachki 194, Rostov-on-Don, 344090 Russia

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

Received January 31, 2000

Abstract—The switchable component of the polarization of ferroelectric ceramics subjected to a pulse electric field was found with the use of the Preisach model. The model parameters were found for PLZT-9/65/35 ceramics from an experimentally found hysteresis loop. The threshold field strength for the switchable component was established to be close to the threshold strength of electron emission. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Electron emission from a ferroelectric surface has been the subject of investigation over the last two decades. Current densities of practical value have been obtained in PLZT ceramics with compositions close to phase transitions [1]. Intense emission is observed under the action of electrical pulses with an amplitude much higher than the coercive field and a duration of several tens of nanoseconds. The pulse repetition rate is usually no more than 1 kHz.

It was established [2] that the emission appears when the vector of spontaneous polarization changes its orientation. However, to gain a better understanding of the emission mechanism, one has to study the near-surface distribution of a potential that appears due to repolarization (polarization switching). Potential evaluations made in [1, 2] ignore volume charges due to repolarization. These charges arise because of field and polarization nonuniformities caused by the complex geometry of one of the electrodes, which is made in the form of a strip array [1–4]. The situation is further aggravated by the fact that, with time, free charges, which are always present in ferroelectrics, screen both volume and surface charges arising during spontaneous polarization reversal. Obviously, self-consistent potential computations for a ferroelectric in an inhomogeneous external field must start with setting a correlation between the field and polarization. This is the aim of this work.

Repolarization in a polycrystalline ceramics is a complex process, because the material is inhomogeneous. Moreover, PLZT compositions we are interested in lie near phase transitions, which are diffuse. Therefore, repolarization in these compositions can adequately be described only in terms of a phenomenological model with parameters numerically evaluated from experiment.

Thermodynamically, complex phase transitions can be described only qualitatively. Such an approach, however, fails to quantitatively characterize ferroelectric polarization. In this work, we apply the Preisach statistical method [5], which was developed for ferromagnetics and applied to ferroelectric ceramics in [6, 7].

THE PREISACH MODEL OF FERROELECTRIC CERAMICS POLARIZATION

This model [6] assumes that all domains have the same spontaneous polarization P_s and one-third (on average) of the domains are aligned with the polarizing field. Also, each of the domains has its own coercive field $E_c > 0$ and exhibits a rectangular hysteresis loop. The effect of the surroundings is characterized by an effective inner field E_i, which can both enhance and weaken the external polarizing field.

In the depolarized state under thermal equilibrium, the spontaneous polarizations of the domains are aligned with the inner field E_i. When the external field E is applied, domains for which $E \geq E_i + E_c$ repolarize. In this case, the basic branch of the hysteresis loop for the orientational polarization P_R vs. field E dependence is given by

$$P_R(E) = (2P_s/3)\int_0^{E}\int_0^{E-E_c} f(E_c, E_i)dE_c dE_i, \qquad (1)$$

where $f(E_c, E_i)$ is the function of domain distribution over coercive and inner fields that is normalized

1063-7842/00/4512-1553$20.00

according to the condition

$$\int_{0}^{\infty}\int_{-\infty}^{+\infty} f(E_c, E_i)dE_c dE_i = 1. \quad (2)$$

For a varying electric field with an amplitude E_m applied to the ceramics, the polarization maximum will be

$$P_m = P_R(E_m). \quad (3)$$

When the field returns to zero, the remanent polarization is

$$P_r = P_m - P_b, \quad (4)$$

where

$$P_b(E) = (2P_s/3)\int_{0}^{E_m/2E_m - E_c}\int_{E_c} f(E_c, E_i)dE_c dE_i. \quad (5)$$

If an alternating-sign field with an amplitude E_m is applied, the descending branch of the hysteresis loop is given by

$$P_d(E) = P_m - (2P_s/3)\int_{0}^{(E_m - E)/2E_m - E_c}\int_{E + E_c} f(E_c, E_i)dE_c dE_i, \quad (6)$$

and the ascending one, by

$$P_a(E) = -P_m + (2P_s/3)\int_{0}^{(E_m + E)/2\ E - E_c}\int_{E_c - E_m} f(E_c, E_i)dE_c dE_i, \quad (7)$$

where E is the instantaneous value of the alternating-sign electric field.

UNSCREENABLE PART OF POLARIZATION CHARGES

Let a train of short unipolar voltage pulses with a large repetition period be applied to a ceramic sample (emitter of electrons). During a pulse, the polarization equals P_m, being P_r for the long remaining time. One can suppose that polarization charges due to the remanent polarization P_r will be neutralized by free charges with time.

Polarization charges due to the switchable part of the polarization $P_b = P_m - P_r$ will remain unneutralized after passing the pulse front at least for a short time. It is these charges that specify the potential distribution in the sample during an electron emission current pulse.

Thus, the evaluation of the potential in the sample requires the knowledge of P_b. To calculate P_b, it is necessary to find the distribution function $f(E_c, E_i)$.

THE SELECTION OF THE DISTRIBUTION FUNCTION

The polarization vector of a ceramic is the mean of a large number of small contributions from individual domains. The domain coercive field depends on its structure and boundary conditions, while the inner field is defined by the interaction of a domain with its environment. These fields can generally be thought of as correlated random quantities. Therefore, the distribution function is selected in the form of a two-dimensional normal distribution

$$f(E_c, E_i) = \frac{1}{2\pi\sigma_c\sigma_i(1-r^2)^{1/2}} \times \exp\left\{\frac{(E_c - E_m)^2\sigma_c^{-2} - 2r(E_c - E_{cm})E_i\sigma_c^{-1}\sigma_i^{-1} + E_i^2\sigma_i^{-2}}{2(1-r^2)}\right\}, \quad (8)$$

where σ_c^2 and σ_i^2 are the variances of the coercive and inner fields, respectively; r is the correlation factor for the fields; and E_{cm} is the mean coercive field.

Expression (8) is an approximate representation of the distribution function. Negative values of the coercive field make no sense, as reflected by normalizing condition (2). However, with regard for the properties of a Gaussian (normal) distribution, one can use (8) if $\sigma_c < E_{cm}/3$. The parameters P_s, E_{cm}, σ_c, σ_i, and r of the distribution function are found experimentally, specifically, from a hysteresis loop.

THE DETERMINATION OF THE DISTRIBUTION FUNCTION PARAMETERS FROM HYSTERESIS LOOP

When using a hysteresis loop to find the parameters, one should realize that, along with the orientational component, it also includes the linear part $k\varepsilon_0 E$ (k is the dielectric susceptibility coefficient), which is associated with quasi-elastic displacements of electron and ion shells.

Six equations necessary to find the parameters k, P_s, E_{cm}, σ_c, σ_i, and r are derived by considering the fields and polarizations at several characteristic points of the

loop and also the slopes of the loop at these points. To be definite, we will take the descending branch [expression (6)]. At the top of the loop, where the field is E_m, the maximum experimentally found polarization P_{me} is related to P_m from (3) as

$$P_{me} = P_m + k\varepsilon_0 E_m. \quad (9)$$

For $E = 0$, the experimental remanent polarization P_{re} satisfies, according to (6), the equality

$$P_{re} = P_m - (2P_s/3)\int_0^{E_m/2}\int_{E_c}^{E_m - E_c} f(E_c, E_i)dE_c E_i. \quad (10)$$

Similarly, the absolute value of the experimental coercive field E_{ce} for the zero experimental polarization meets the equality

$$P_m - (2P_s/3) \times \int_0^{(E_m + E_{ce})/2}\int_{(E_c - E_{ce})}^{E_m - E_c} f(E_c, E_i)dE_c dE_i - k\varepsilon_0 E = 0. \quad (11)$$

The slope K_m of the experimental hysteresis loop at its top and the derivative

$$dP_m/dE_m = (2P_s/3)\int_0^{E_m} f(E_c, E_m - E_c)dE_c, \quad (12)$$

found from (3) and (1), are related by

$$K_m = dP_m/dE_m + k\varepsilon_0. \quad (13)$$

Similarly, the slopes K_0 and K_c of the experimental loop at the points $E = 0$ and $E = -E_{ce}$, respectively, are related to the derivatives of (6) in the vicinity of these points:

$$dP_d/dE|_{E=0} = (2P_s/3)\int_0^{E_m/2} f(E_c, E_c)dE_c, \quad (14)$$

$$dP_d/dE|_{E=-E_{ce}} = (2P_s/3)\int_0^{(E_m + E_{ce})/2} f(E_c, E_c - E_{ce})dE_c \quad (15)$$

as

$$K_0 = dP_d/dE|_{E=0} + k\varepsilon_0, \quad (16)$$

and

$$K_c = dP_d/dE|_{E=-E_{ce}} + k\varepsilon_0. \quad (17)$$

Thus, we arrive at the system of six Eqs. (9)–(11), (13), (15), and (17). Using the experimentally found parameters E_m, P_{me}, P_{re}, E_{ce}, K_m, K_0, and K_c and taking into account (1), (3), (12), (14), and (16), we numerically minimize the residues to find the susceptibility k and five parameters (P_s, E_{cm}, σ_c, σ_i, and r) of the distribution function that give the best fit between the calculated and experimental hysteresis loops.

Table

t, °C	P_s, μC/cm^2	E_{cm}, kV/cm	σ_i, kV/cm
–20	98	12.6	6.1
+20	74	3.73	11.4
+50	69	2.60	23.0

AN EXAMPLE OF SWITCHABLE COMPONENT CALCULATION

The switchable component P_b of the polarization [expression (5)] was calculated using experimental hysteresis loops [8] for PLZT-9/65/35 at temperatures between –20°C and +70°C.

It was shown that, at temperatures away from the phase transition temperature, the best fit is provided if σ_c values are much lower than the mean coercive field E_{cm} and inner field variance σ_i. Then, one can employ the simpler approximation for the distribution function:

$$f(E_c, E_i) = (2\pi)^{-1/2}\sigma_i^{-1}\exp\{-E_i^2/(2\sigma_i^2)\}\delta(E_c - E_{cm}), \quad (18)$$

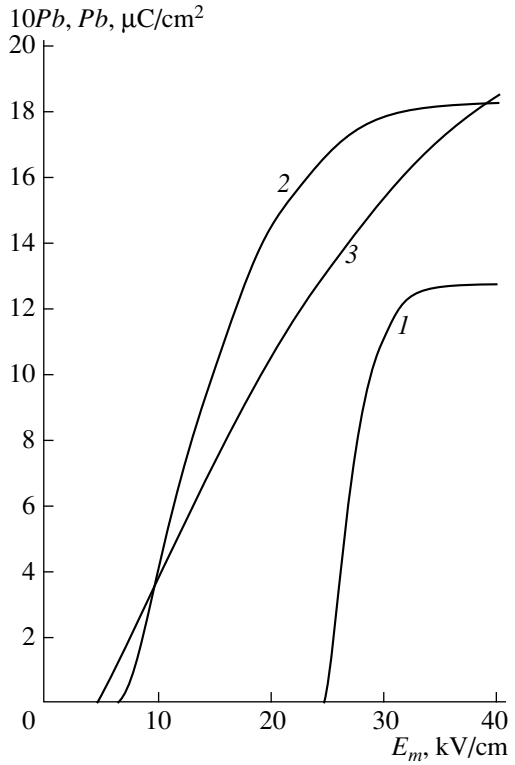
where δ is the Dirac function. As a result, we come to the following expression that relates the switchable polarization P_b and pulse amplitude E_m:

$$P_b = \begin{cases} 0, & E_m < 2E_{cm}, \\ (P_s/3)\{\mathrm{erf}[(E_m - E_{cm})/(2^{1/2}\sigma_i)] \\ \quad - \mathrm{erf}[E_{cm}/(2^{1/2}\sigma_i)]\}, & E_m > 2E_{cm}. \end{cases} \quad (19)$$

The table lists the calculated parameters of the distribution function that define the switchable part for three temperatures. As the temperature rises, the spontaneous polarization P_s and mean coercive field E_{cm} decrease, while the inner field variance σ_i grows. At a temperature of +70°C, which is close to that of P_s disappearance [9], the model becomes invalid, since the calculated $\sigma_c > E_{cm}$.

The figure plots P_b against E_m. At –20°C (the hysteresis loop is nearly rectangular), the switchable component is much less than at +20°C or +50°C, when the loops are largely inclined. As follows from (19), at –20°C, the value of $2E_{cm}$ is comparable to the highest E_m obtainable in the experiment and σ_i is small; on the contrary, at +20°C and +50°C, E_{cm} is small and σ_i is large.

The switchable part saturates at $E_m - E_{cm} \gg \sigma_i$ and is absent at $E_m < 2E_{cm}$. This implies the existence of a threshold electric field. Such a threshold for electron emission due to a pulsed electric field was found in [9].



Switchable polarization P_b vs. electric field strength E_m: (*1*) $10P_b$, t = –20°C; (*2*) P_b, t = +20°C; and (*3*) P_b, t = +50°C.

For the ceramics under study, the threshold roughly equals 6 kV/cm, which agrees well with the threshold field of switchable polarization at near-room temperatures. It is, therefore, natural to suggest that the switchable polarization, related polarization charges, and fields produced by these charges contribute to an electron emission mechanism.

CONCLUSIONS

(1) A mechanism of electron emission from ferroelectric ceramics subjected to a pulse electric field can be elucidated if the sample potential is calculated and polarization charges are taken into account.

(2) It is assumed that only the polarization component that switches at the short-pulse front generates polarization charges, which have no time to be neutralized during emission.

(3) The Preisach model is suggested for switchable polarization calculations.

(4) A method for the determination of the Preisach model parameters from hysteresis loops was developed.

(5) The Preisach model parameters and the switchable polarization vs. electric field strength dependence were evaluated for PLZT-9/65/35 ceramics at several temperatures.

(6) The switchable polarization has a threshold field, which is close to that for electron emission. This fact agrees well with the assumption that the electric fields of polarization charges (associated with the switchable polarization) play an essential part in electron emission from ferroelectric ceramics.

The Preisach model used in this work looks like a formal mathematical way of approximating an experimental hysteresis loop by the sum of elementary rectangular loops. However, it has a fundamental physical basis, which is supported, in particular, by the fact that the emission current is the superposition of a large number of short spikes due to the reorientation of small domains in ferroelectric ceramics [1, 10].

The developed method for computing the switchable part of the polarization in ferroelectric ceramics will be used in investigating polarization charges and their influence on the potential distribution in an emitting ceramic sample.

ACKNOWLEDGMENTS

This work was supported by the “Integration” program (project no. 5-82).

REFERENCES

1. H. Gundel, Integr. Ferroelectr. **2**, 207 (1992).
2. W. Zhang and W. Huebner, Ferroelectrics **221**, 199 (1999).
3. G. Rosenman, D. Shur, Kh. Garb, *et al.*, J. Appl. Phys. **82**, 772 (1997).
4. L. M. Rabkin and V. N. Ivanov, Pis’ma Zh. Tekh. Fiz. **24** (14), 54 (1998) [Tech. Phys. Lett. **24**, 565 (1998)].
5. F. Preisach, Z. Phys. **94**, 277 (1935).
6. A. V. Turik, Fiz. Tverd. Tela (Leningrad) **5** (4), 1213 (1963) [Sov. Phys. Solid State **5**, 885 (1963)].
7. A. V. Turik, Fiz. Tverd. Tela (Leningrad) **5** (10), 2922 (1963) [Sov. Phys. Solid State **5**, 2141 (1963)].
8. K. Carl and K. Geisan, Proc. IEEE **61** (7), 967 (1973).
9. H. Gundel, J. Handerek, H. Riege, *et al.*, Ferroelectrics **110**, 183 (1990).
10. H. Gundel, J. Handerek, and H. Riege, J. Appl. Phys. **69**, 975 (1991).

Translated by V. Isaakyan

Technical Physics, Vol. 45, No. 12, 2000, pp. 1557–1565. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 47–57.

SOLIDS

Autowave Dynamics of the Magnetic Flux in Nonideal Type-II Superconductors with Different Current–Voltage Characteristics

V. R. Romanovskiĭ

Russian Research Center Kurchatov Institute, Moscow, 123182 Russia

Received May 11, 1999; in final form, February 22, 2000

Abstract—An analysis was performed of the processes of penetration of a macroscopic electromagnetic field into superconducting media with different current-voltage characteristics induced by a variable external magnetic field or by carrier transport. It has been shown that even if a finite electrical voltage arises in the superconductor before the critical current is reached, the magnetic flux, as in the critical state model, penetrates into the superconductor at a finite rate as a characteristic electromagnetic wave. At its front, a special condition is fulfilled; namely, a smooth transition to the nonperturbed value of the strength of the electric field induced by an external perturbation. Formulas are given for the calculation of the penetration rate of the electromagnetic field into a superconductor. For corroboration of the formulated relationships, the study was compared with corresponding numerical calculations. © *2000 MAIK "Nauka/Interperiodica".*

The basis for the study of the physical characteristics of non-steady-state processes occurring in nonideal type-II superconductors in response to an external perturbation is the analysis of the magnetic flux dynamics. If the characteristic dimensions of the perturbation are far in excess of the London penetration depth, then the macroscopic Maxwell equations can be used to describe the superconductor electrodynamics. They should be supplemented with corresponding relationships that establish the correlations between characteristic superconductor parameters. One of the most important of these relationships is the current-voltage (*I–V*) characteristic of a superconductor.

Despite its importance, a general microscopic theory, which could explain the experimentally observed *I–V* characteristics of nonideal type-II superconductors over the entire range of the electric field strength variations, is still lacking. Therefore, phenomenological models are widely used for their description. These models provide averaged descriptions of the experimentally observed behavior of the electrical voltage arising inside a superconductor determined by the peculiarities of the movement of its vortex structure, the nonuniformity of its critical properties, etc.

It is known [1] that a nonideal type-II superconductor passes to the resistive state when the density of the current flowing in it exceeds a certain threshold value. Physically, this corresponds to conditions in which the Lorentz force exceeds the pinning force. In this case, a nonlinear region appears in the *I–V* characteristic, which is due to the creep of the flux and the gradual development of vortex lattice motion [2]. The nonlinear region in the *I–V* characteristics can be caused by many factors, namely pinning nonuniformity, defects of the vortex lattice structure, thermal activation of the magnetic flux, etc. With increasing current, the Lorentz force can mobilize the entire vortex lattice and the *I–V* characteristic of the superconductor can become linear. This is the region of so-called viscous flow of the magnetic flux [3].

For the description of these modifications of the *I–V* characteristics, the following models are used.

If the nonlinear region in the *I–V* characteristics is negligible, the approximation of the *I–V* characteristics can be obtained using a model of the viscous flow of the magnetic flux. This approximation is expressed as

$$E = \begin{cases} 0, & J < J_C, \\ (J - J_C)\rho_f, & J \geq J_C, \end{cases} \quad (1)$$

where J_C is the critical current density depending on the temperature and magnetic field induction, and ρ_f is the resistance to current flow.

From this model, Bean's model of the critical state [4] follows, in which the slope of the *I–V* characteristics is ignored. According to this approach, any perturbation induce a current density J_C in a superconductor. By means of this simplified model, the electrodynamics of the nonideal type-II superconductors can be described in a simple and illustrative form. In particular, the penetration rate of the electromagnetic field into a superconductor induced by an increasing external magnetic field can readily be obtained. For a planar superconductor, its value is equal to $v = \dot{B}/\mu_0 J_C$, where $\dot{B} = dB/dt$ is the rate of increase of the external magnetic field.

1063-7842/00/4512-1557$20.00

In the cases where the *I–V* characteristics are nonlinear in a wide range of the electric field strengths, various other models are used for the description of the *I–V* characteristics of the nonideal type-II superconductors. In particular, they can be described by power and exponential laws, which in the simplest cases have the form [5, 6]

$$E = E_C\left(\frac{J}{J_C}\right)^n, \tag{2}$$

$$E = E_C\exp\left(\frac{J - J_C}{J_\delta}\right). \tag{3}$$

Here, J_C is the current density determined at the electric field strength E_C; J_δ and n are the current density and the exponent, respectively, that determine the slope of the *I–V* characteristics. *I–V* characteristics of the form (2) are typical of superconductors in which the current dependence of the potential barrier height is logarithmic. Equation (3) is based on the idea of thermal activation of the superconductor vortex structure [2, 3]. Some macroscopic effects also lead to exponential *I–V* characteristics in the nonideal type-II superconductors, particularly the inhomogeneity of the superconductor's physical properties. In such materials, the superconducting transition in different grains occurs at different temperatures, magnetic inductions, and current densities. Apart from the inhomogeneity of the critical parameters of the bulk of the superconductor, it can have a longitudinal inhomogeneity of the superconducting properties (the so-called "sausage links" effect). The linear term in the exponent of the *I–V* characteristic (3) becomes quadratic. However, in such cases, the shape of the *I–V* characteristics is satisfactorily given by the simple and illustrative Eq. (2).

By virtue of the mentioned peculiarities, the values of J_δ and n vary widely. For nonideal type-II superconductors at liquid helium temperatures, the following values are typical: $E_C = 10^{-5}$–10^{-2} V/m; $J_\delta = (0.01$–$0.02)J_C$; and $n \sim 50$–100. The current-voltage characteristics of the high temperature superconductors, as a rule, can be described in terms of model (2), in which the exponent usually varies in the range $n \sim 5$–30.

At $n \longrightarrow \infty$ and $J_\delta \longrightarrow 0$, models (2) and (3) transform into the critical state model. However, it should be noted that in Eq. (3) the electric field strength at $J = 0$ is equal to $E_0 = E_C\exp(-J_C/J_\delta)$. In order to avoid such an uncertainty, Eq. (3) should be written as

$$E = E_C\exp\left(\frac{J - J_C}{J_\delta}\right) - E_0. \tag{4}$$

Therefore, for the correctness of the obtained results, *I–V* characteristics of the form (4) will be used below.

Overall, the equations for the current-voltage characteristics give a fairly accurate account of a wide range of experimental data obtained from measurements of the current-carrying capacity of real superconductor wires. So, the *I–V* characteristics of physically homogeneous wires with a small number of superconducting fibers are quite steep. Therefore, the viscous flow model for such wires is suitable for the analysis of stability conditions of the superconducting state with respect to thermal perturbations and to the loss calculations in variable magnetic fields [5, 6]. At the same time, the fabrication technology of superconductor wires is becoming more complicated and practical applications currently deal with superconducting materials such as composites with very thin superconducting fibers and wires with a complex cross sectional structure with nonlinear current-voltage characteristics. In these cases, the choice between models (2) and (3) is mainly guided by details of the interpretation of the experimental data. Since none of these models is universal, both power and exponential laws are widely used to describe the *I–V* characteristics of real wires [2–11]. However, it should be noted that despite the constant attention given to the study of the magnetic flux dynamics in type-II superconductors (see, for example, [12–16]), an exhaustive analysis of the main physical laws of their electrodynamic behavior that takes into account all the phenomenological models of the *I–V* characteristics in real semiconductors is still lacking. In this study, a solution to the system of the equations

$$\begin{gathered}\mathrm{rot}\mathbf{E} = -\partial\mathbf{B}/\partial t, \quad \mathrm{rot}\mathbf{B} = \mu_0\mathbf{J}, \\ \mathrm{div}\mathbf{B} = 0, \quad \mathrm{div}\mathbf{J} = 0\end{gathered} \tag{5}$$

with *I–V* characteristics (1), (2), and (4) was employed in order to describe the main regularities of electromagnetic field penetration into nonideal type-II superconductors initiated by a variable external magnetic field and by inducing carrier transport. Main attention is paid to the diffusion processes occurring in superconducting materials that retain superconductivity at liquid helium temperatures.

Let us consider a well-cooled semi-infinite superconducting medium placed in a homogeneous external magnetic field parallel to its boundary surface (Fig. 1). Let the magnetic field B_a at the superconductor surface be equal to zero at the initial moment and then increase at a constant rate ($B_a = \dot{B}t$) in such a way that the superconductor temperature remains essentially constant and equal to the coolant temperature. For simplification, we will also suppose that the current density in the superconductor is independent of the magnetic field induction. After obvious transformations, the system of Eqs. (5) can be reduced to the following form:

$$\mu_0\partial\mathbf{J}/\partial t = \Delta\mathbf{E}. \tag{6}$$

Taking into account the character of the electromagnetic field variation and the regularity of the corresponding values at infinity, the necessary boundary and initial conditions with respect to the electric field

strength $E = E(x, t)$ for the considered one-dimensional geometry will be

$$\partial E/\partial x(0, t) = -dB/dt, \quad E(\infty, t) = 0, \quad E(x, 0) = 0. \tag{7}$$

Along with (7), an additional constraint will be required for further analysis,

$$\mu_0 \int_0^\infty J(x, t)dx = B_a, \quad t > 0, \tag{8}$$

which expresses the continuity in the variation of the magnetic induction. Excluding the current density in (7), the following one-dimensional equations can be obtained for the above forms of the *I–V* characteristics:

$$\frac{\partial^2 E}{\partial x^2} = \mu_0 \begin{cases} \dfrac{1}{\rho_f}\dfrac{\partial E}{\partial t} & \text{for } I\text{–}V \text{ characteristics of form (1),} \quad (9) \\ \dfrac{J_C}{nE}\left(\dfrac{E}{E_C}\right)^{\frac{1}{n}}\dfrac{\partial E}{\partial t} & \text{for } I\text{–}V \text{ characteristics of form (2),} \quad (10) \\ \dfrac{J_\delta}{E + E_0}\dfrac{\partial E}{\partial t} & \text{for } I\text{–}V \text{ characteristics of form (4).} \quad (11) \end{cases}$$

To solve the problem sets, it is convenient to use the dimensionless variables $e = E/E_x$, $X = x/L_x$, $\tau = t/t_x$, where

$$E_x = J_C\rho_f, \quad L_x = J_C\rho_f/\dot{B},$$

$$t_x = \mu_0 J_C^2 \rho_f/\dot{B}^2 \text{ for } I\text{–}V \text{ characteriscics of form (1),}$$

$$E_x = E_C, \quad L_x = E_C/\dot{B},$$

$$t_x = \mu_0 J_C E_C/\dot{B}^2 \text{ for } I\text{–}V \text{ characteristics of form (2),}$$

$$E_x = E_C, \quad L_x = E_C/\dot{B},$$

$$t_x = \mu_0 J_\delta E_C/\dot{B}^2 \text{ for } I\text{–}V \text{ characteristics of form (4).}$$

Then, in the dimensionless form, the electric field variation in space and time can be described by the initial boundary-value problems of the form

$$\frac{\partial^2 e}{\partial X^2} = \begin{cases} \dfrac{\partial e}{\partial \tau}, & (12) \\ \dfrac{1}{n}e^{(1-n)/n}\dfrac{\partial e}{\partial \tau}, & (13) \\ \dfrac{1}{e + e_0}\dfrac{\partial e}{\partial \tau}, & (14) \end{cases}$$

$$\frac{\partial e}{\partial X}(0, \tau) = -1, \quad e(\infty, \tau) = 0, \quad e(X, 0) = 0 \tag{15}$$

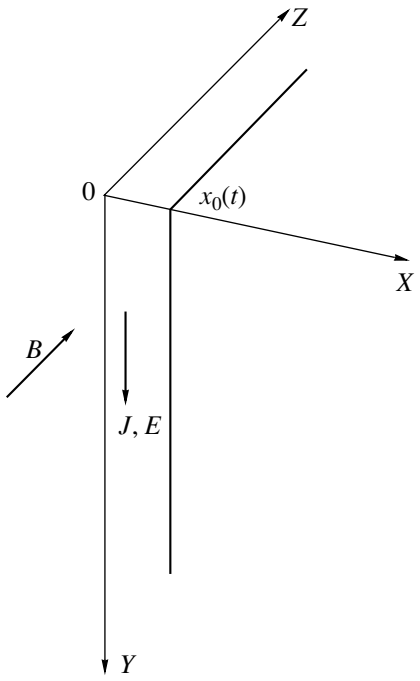


Fig. 1. The electromagnetic field distribution in a superconducting half-space.

subject to additional conditions

$$\int_0^\infty (1 + e)dX = \tau \tag{16}$$

for the *I–V* characteristics of form (1),

$$\int_0^\infty e^{1/n}dX = \tau \tag{17}$$

for the *I–V* characteristics of form (2)

$$\int_0^\infty [1 + \delta \ln(e + e_0)]dX = \delta\tau, \quad \delta = J_\delta/J_C, \quad e_0 = E_0/E_C \tag{18}$$

for the *I–V* characteristics of form (4).

The required distribution of the electric field strength will be looked for in the class of automodel solutions using the group properties of the differential Eqs. (12)–(14). Let us write them for each considered form of the *I–V* characteristics.

To solve problems (12) and (15), we will use invariants of the form

$$V(Z) = \partial e/\partial X, \quad Z = X/\sqrt{\tau}. \tag{19}$$

Then, going from differentiation with respect to the independent variables X, τ to differentiation with

respect to Z, the partial differential Eq. (12) is reduced to the ordinary differential equation

$$\frac{d^2V}{dZ^2} + \frac{Z}{2}\frac{dV}{dZ} = 0.$$

Using the appropriately transformed boundary conditions $V(0) = -1$, $V(\infty) = 0$, we obtain

$$V(Z) = C\,\mathrm{erf}(Z/2) - 1, \tag{20}$$

where

$$\mathrm{erf}(y) = \frac{2}{\sqrt{\pi}}\int_0^y \exp(-x^2)dx$$

is the error function and C is a positive constant.

The expression for calculating it is not given here, as it is too cumbersome.

Since the electric field is finite inside a superconductor for all $Z = X/\sqrt{\tau} > 0$, the inequality $\partial e/\partial X < 0$ is valid. Therefore, the region of existence of nonnegative values $e(X, \tau)$ is finite. This conclusion also follows from the continuity condition (16). Therefore, according to (15) and (19), the electric field distribution inside the superconductor is described by the expression

$$e(X, \tau) = X_0(\tau) - X + C\int_{X_0(\tau)}^{X} \mathrm{erf}\left(\frac{X}{2\sqrt{\tau}}\right)dX,$$

where $X_0(\tau)$ is the moving boundary of the magnetization region of the superconductor.

Using for its determination a modified condition (16)

$$\int_0^{X_0(\tau)} (1 + e)dX = \tau, \tag{21}$$

we obtain the motion equation for the magnetization front

$$\tau = X_0 + \frac{X_0^2}{2} + C\int_0^{X_0} [\Psi(X, \tau) - \Psi(X_0, \tau)]dX. \tag{22}$$

Here, $\Psi(X, \tau) = X\,\mathrm{erf}(X/2\sqrt{\tau}) + 2\sqrt{\tau/\pi}\exp(-X^2/4\tau)$. Differentiating Eq. (21) with respect to τ, it is easily found that $dX_0/\tau = \partial e/\partial X|_{X = X_0}$ is in the viscous flow model at the boundary of the magnetization region and the electric field strength profile, and, as in the critical state model, has a finite derivative with respect to the space coordinate. From (22) it follows that, contrary to the critical state model, the propagation rate of the magnetization front inside the nonideal type-II superconductor is generally not constant. However, if we take into account that the cross-sectional dimensions of the superconducting strands are small in real composite superconductors, it is easy to show that for them, $X_0 \ll 1$. Therefore, to a good approximation, $X_0 \sim \tau$; that is, as in the critical state model, the penetration rate of the magnetic flux into the superconductors with I–V characteristics (1) is essentially constant and in the dimensional form it is equal to $v = \dot{B}/\mu_0 J_C$.

For a superconducting media with a power-law I–V characteristic, the automodel solution can be obtained by introducing invariants of the form

$$V(Z) = e\tau^p, \quad Z = X/\tau^p, \quad p = n/(n + 1).$$

Then the problems (13) and (15) are reduced to the integration of an ordinary nonlinear equation

$$(n + 1)\frac{d^2V}{dZ^2} + ZV^{(1-n)/n}\frac{dV}{dZ} - V^{1/n} = 0 \tag{23}$$

with the boundary conditions

$$\frac{dV}{dZ}(0) = -1, \quad V(\infty) = 0. \tag{24}$$

In this case, the additional condition (17) transforms to

$$\int_0^{\infty} V^{1/n}dZ = 1. \tag{25}$$

In Fig. 2, a set of solutions to Eq. (23) is presented for I–V characteristics with $n = 80$, which is typical for the Nb–Ti alloy. Curves *1–3* are obtained by solving (23) numerically. Instead of the boundary conditions (24), the initial value of the derivative and variable values of invariant $V(0)$ were set at $Z = 0$; these values are shown in the figure caption. In addition, fulfillment of condition (25) was tested; this is shown in the inset in Fig. 2 for the given $V(0)$ values. Corresponding values are plotted for the integral

$$I(Z) = \int_0^Z V^{1/n}dZ.$$

As it follows from Fig. 2, the characteristic features of the plotted curves is the occurrence of the point Z_0, at which the function $V(Z)$ becomes zero. Its existence determines the allowed range of variation of the invariant Z, since at $Z > Z_0$ there is no real-value solution to (23). Accordingly, the second boundary condition (24) and the condition (25) should be modified. The latter will thereby determine the desired Z_0 value.

At some finite Z_0 value, $V(Z_0) = 0$ determines the special features of the process of penetration of the magnetic flux into the superconducting medium with power-law I–V characteristics. First, the screening currents produced in the superconductor by the external magnetic field will flow only through a finite region of the superconductor adjacent to its surface $0 \le X \le X_0(\tau) = Z_0\tau^p$ with $e(X_0, \tau) = 0$. Therefore, the screening

currents will spread inside the semiconductor at a finite rate. Second, at the front of the magnetization region, the conditions for variation of $V(Z)$ and, consequently, of the electromagnetic field, are special. Let us define them. For this purpose, we will differentiate Eq. (17) with respect to time, and the upper limit will be replaced in accordance with the above by $X_0(\tau)$. After simple transformations, we obtain

$$\left(e^{1/n}\frac{dX_0}{d\tau} + \frac{\partial e}{\partial X}\right)\Bigg|_{X = X_0} = 0.$$

Since $e(X_0\tau) = 0$, then $\partial e/\partial X(X_0\tau) = 0$. Therefore, $dV/dZ = 0$ at $Z = Z_0$. From (23), it follows that the second and all higher derivatives d^kV/dZ^k at $Z = Z_0$ are also equal to zero. Consequently, at $Z \geq Z_0$, the invariant $V(Z)$, and $e(X, \tau)$ at $X \geq X_0$ continue with null equation.

Using the noted properties of the invariant $V(Z)$, let us transform Eq. (23) into an equivalent second-order integral equation by grouping the corresponding terms in (23) and integrating it with respect to Z between the limits from 0 to Z and then from Z_0 to Z. As a result, we obtain

$$V(Z) = Z_0 - Z + \int_{Z_0}^{Z} dx \int_{0}^{x} V^{1/n}(y)dy - \frac{n}{n+1}\int_{Z_0}^{Z} V^{1/n}(y)dy.$$

The solution of this integral equation will be sought according to the method of successive approximations. As a zeroth approximation, we take

$$V_0(Z) = Z_0 - Z.$$

In this case, Z_0 in accordance with condition (25) is equal to

$$Z_0 = \left(\frac{n+1}{n}\right)^{n/(\tau - 1)}.$$

Substituting the zeroth approximation into the right side of the starting integral equation, the following approximation can be obtained:

$$V_1(Z) = V_0(Z) - pZ_0^{1/p}V_0(Z) + p^2Z_0V_0^{1/p}(Z).$$

In Fig. 2, curves *4* and *5* represent the zeroth and the first approximations of $V(Z)$, respectively. It is clearly seen that for the Nb–Ti alloy, even the zeroth approximation is an accurate enough solution of the integral equation. A more detailed analysis of possible solutions of the problems (23)–(25) shows that, at $n \gg 10$, the linear approximation adequately describes the sought-for invariants $V(Z)$. We will use it to determine the strength of the electric field induced inside the superconductor. After transition to the dimensional variables, we obtain

$$E(x, t) = \frac{dB}{dt}[x_0(t) - x], \quad 0 \leq x \leq x_0(t), \quad t > 0. \quad (26)$$

This relationship completely coincides with that of the critical state model [5, 6]. But for the *I–V* characteristics (2), the motion of the magnetization front is described by the expression

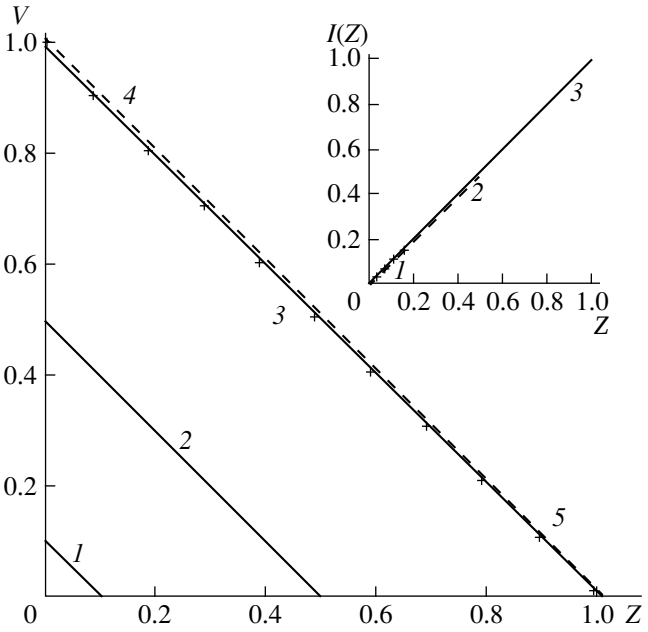


Fig. 2. A set of solutions to Eq. (23). $V(0)$: (*1*) 0.1, (*2*) 0.5, (*3*) 1.000017.

$$x_0(t) = \left(\frac{n+1}{n}\right)^{n/(n+1)}\left(\frac{dB}{dt}\right)^{\frac{n-1}{n+1}} t^{n/(n+1)}\left(\frac{E_C}{\mu_0^n J_C^n}\right)^{1/(n+1)},$$

according to which the magnetic flux propagates along the superconductor at a rate

$$\frac{dx_0}{dt}(t) = \left(\frac{dB}{dt}\right)^{\frac{n-1}{n+1}}\left(\frac{n}{n+1}\frac{E_C}{\mu_0^n J_C^n t}\right)^{1/(n+1)}.$$

It is not difficult to find that in the limit $n \longrightarrow \infty$, the above expressions become identical to those of the critical state model.

In Fig. 3, the electric field distributions inside the Nb–Ti superconductor are shown at different moments of time with the external magnetic field varying at a rate $dB/dt = 1$ T/s. Dashed lines are plots of $E(x, t)$ obtained with the use of the zeroth approximation. Solid lines are the results of the numerical solutions to the problems (2), (7) and (10) with a time increment $\Delta t = 10^{-8}$ s and space increment $\Delta x = 10^{-9}$ m. Its accuracy can be inferred from the results of calculation of the boundary of the magnetization region $x_0(t)$, which are shown in the inset in Fig. 3. The initial parameters in the calculations are $n = 80$, $E_C = 10^{-4}$ V/m, and $J_C = 4 \times 10^9$ A/m^2.

Let us consider the peculiarities of the magnetic flux dynamics in the superconductors with exponential *I–V* characteristics. Since $e_0 \ll 1$, the solution of (14)and (15) will be sought in the form of the expansion in

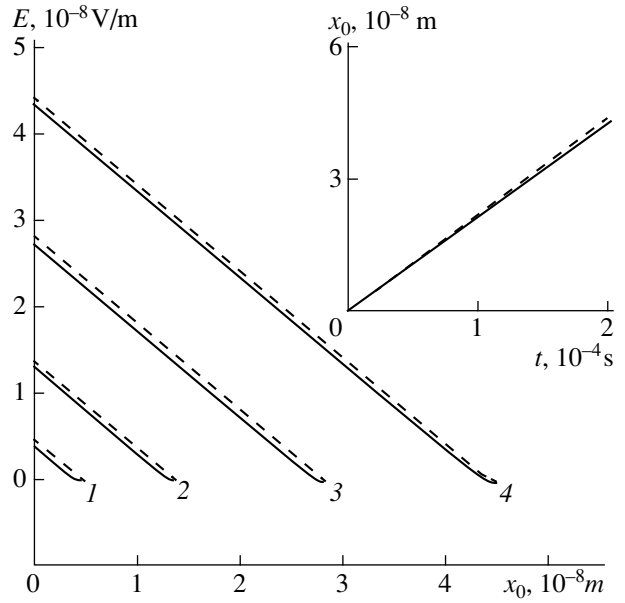


Fig. 3. The distribution of the electric field strength inside a superconductor with the power-law *I–V* characteristics. *t*, 10^{-4} s: (*1*) 0.2, (*2*) 0.6, (*3*) 1.257, (*4*) 2.

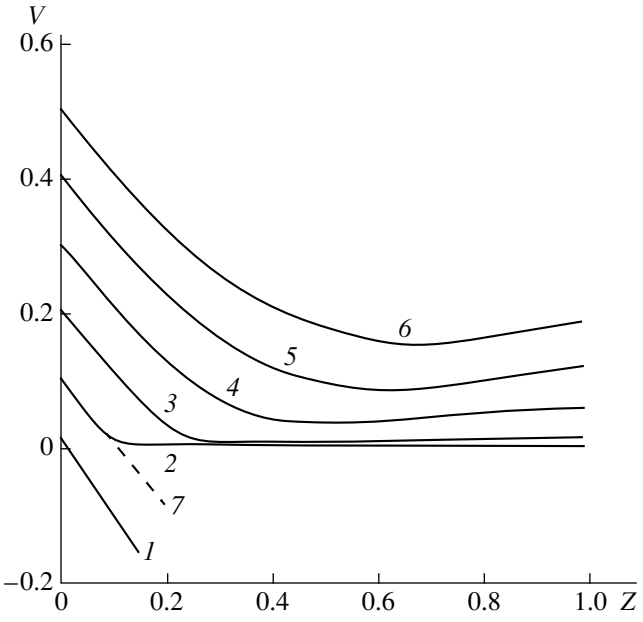


Fig. 4. A set of solutions to the problem (27). *V*(0): (*1*) 0.01, (*2*) 0.1, (*3*) 0.2, (*4*) 0.3, (*5*) 0.4, (*6*) 0.5, (*7*) approximate solution.

terms of a small parameter e_0:

$$e = \varepsilon_0 + \varepsilon_1 e_0 + \varepsilon_2 e_0^2 + \ldots$$

Then, in the zeroth approximation with respect to the expansion in terms of e_0, the initial problem can be reduced to the form

$$\varepsilon_0 \frac{\partial^2 \varepsilon_0}{\partial X^2} = \frac{\partial \varepsilon_0}{\partial \tau},$$

$$\frac{\partial \varepsilon_0}{\partial X}(0, \tau) = -1, \quad \varepsilon_0(\infty, \tau) = 0, \quad \varepsilon_0(X, 0) = 0.$$

It is easy to see that due to the smallness of e_0, the subsequent approximations will have practically no influence on $e(X, \tau)$ values. Therefore, to determine $e(X, \tau)$, it is sufficient to confine oneself to the zeroth approximation with respect to expansion in terms of e_0. Assuming $V(Z) = \varepsilon_0/\tau$, and $Z = X/\tau$, the boundary problem for the determination of the invariant $V(Z)$ will be

$$V\frac{d^2V}{dZ^2} + Z\frac{dV}{dZ} - V = 0,$$
$$\frac{dV}{dZ}(0) = -1, \quad V(\infty) = 0. \tag{27}$$

In Fig. 4, possible $V(Z)$ dependencies are presented, which were determined by the numerical solution of (27) for different initial values of $V(0)$ and in conformity with the initial boundary condition $Z = 0$. Since $V(Z)$ should be not negative according to its physical meaning and at $Z \longrightarrow \infty$ should be finite, the curves presented in Fig. 4 allow the following conclusions to be made. The form of the invariant should differ from the dependencies presented by curves *2–6*, which have a positive slope starting with some Z value. Therefore, the solution of (27), as in the cases considered above, should be sought in a limited range $0 \le Z \le Z_0$ with $V(Z_0) = 0$. Then, (18) and (27) give that at $Z = Z_0$, all derivatives with respect to Z are equal to zero; that is, $V(Z) = 0$ for all $Z \ge Z_0$. Consequently, the nontrivial solution of problems (4), (7) and (11), as in the cases of the critical state models and power-law *I–V* characteristics, is also in the form of an electromagnetic wave propagating through the superconductor at a finite rate. Its front is the boundary between the magnetization region and the region of nonperturbed initial values, with a smooth transition between these values and the corresponding values induced by the external magnetic field in the magnetization region.

To determine $V(Z)$, the differential Eq. (27) must be transformed into a corresponding integral equation. Dividing Eq. (27) by V, and integrating it first from 0 to Z and then from Z_0 to Z, the following second-order integral equation will be obtained

$$V(Z) = Z_0 - Z + \frac{Z^2 - Z_0^2}{2} - \int_{Z_0}^{Z} dx \int_0^x \frac{y}{V(y)} \frac{dV}{dy} dy.$$

From Fig. 4, it follows that the admissible values of the invariant $V(Z)$ exist in a range $Z \ll 1$. Therefore, even the linear approximation of the solutions of Eq. (27) presented in Fig. 4 for the case $V(0) = 0.1$ by the dashed line *7* will satisfactorily describe the $V(Z)$

values. Consequently, in the linear approximation, the electric field strength inside the superconductors can also be described by Eq. (26), in which the moving coordinate of the magnetization front for superconducting media with *I–V* characteristics of form (4) is the solution of a transcendental equation

$$B_a = \mu_0 J_C x_0(t) + \mu_0 J_\delta\left(\frac{E_0}{dB/dt} + x_0(t)\right) \times\left(\ln\frac{E_0 + x_0(t)dB/dt}{E_C} - 1\right) - \mu_0 J_\delta \frac{E_0}{dB/dt}\left(\ln\frac{E_0}{E_C} - 1\right). \quad (28)$$

Taking into consideration the smallness of E_0, it is easy to find from this equation an upper bound of the rate of penetration of the magnetic *I–V* characteristics for all $x_0(t) \ll E_C/\dot{B}$

$$v = \frac{dB/dt}{\mu_0(J_C - 2J_\delta)}.$$

Note that in these approximate expressions the limit $J_\delta \longrightarrow 0$ corresponds to the transition to the critical state model.

In Fig. 5 for the Nb–Ti superconductor with the exponential *I–V* characteristics, the spatial distribution of the electric field strength for different instants of time is presented, calculated using Eqs. (26) and (28) and also obtained by numerical solution of the initial problems (7) and (11). In the calculations, it was taken that $dB/dt = 1$ T/s, $E_C = 10^{-4}$ V/m, $J_C = 4 \times 10^9$ A/m^2, and $J_\delta = 4 \times 10^7$ A/m^2. The convergence of the numerical solution at $\Delta x = 10^{-9}$ m with a decreasing time increment is shown in the inset in Fig. 5.

The presented analytical and numerical solutions clearly demonstrate the physical features of the process of electromagnetic field penetration into nonideal type-II superconductors induced by a varying external magnetic field. However, apart from the magnetic field variation, fluctuations in the carrier transport can also occur. Generally, these processes occur simultaneously and the final state of the current-carrying unit depends on the sequence of their variation. Among the diverse possible problems, of particular importance for the evaluation of the current-carrying capacity of wires and for the analysis of the experimental results on the superconductor *I–V* characteristics is the problem of current distribution in a superconducting composite conductor (superconductor + matrix), which has a cylindrical shape in most practical cases. We consider this problem taking into account the automodel character of the electromagnetic field propagation inside superconducting media with *I–V* characteristics of forms (2) and (4).

Consider an infinite cylindrical wire of radius a with a superconductor uniformly distributed over its cross-section and with the filling coefficient η (Fig. 6). The external magnetic field is supposed to be constant. At

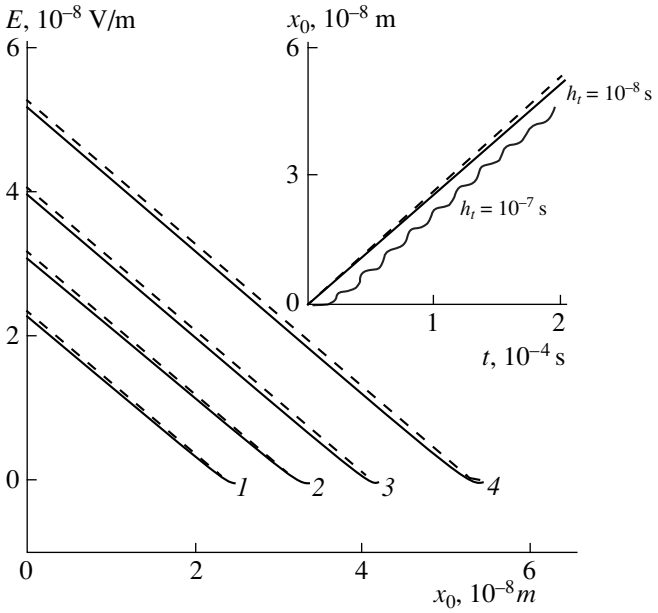


Fig. 5. The distribution of the electric field strength inside a superconductor with the exponential *I–V* characteristics. t, 10^{-4} s: (*1*) 0.888, (*2*) 1.208, (*3*) 1.537, (*4*) 2.

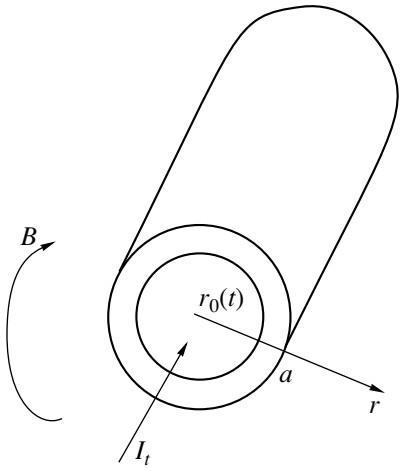


Fig. 6. Current and magnetic field distribution in the superconducting cylinder upon current injection.

the initial instant of time, the current is zero and starts to increase at a specified rate dI/dt. In a simplified problem, in which the current flowing through the matrix can be ignored [5, 6], the electric field strength induced by the varying current according to Eq. (6) in the cylindrical coordinate system can be described by a nonlinear parabolic equation of the form

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial E}{\partial r}\right) = \mu_0\eta\begin{cases}\dfrac{J_C}{nE}\left(\dfrac{E}{E_C}\right)^{1/n}\dfrac{\partial E}{\partial t} & \text{for } I\text{–}V \text{ characteristics (2),}\\ \dfrac{J_\delta}{E + E_0}\dfrac{\partial E}{\partial t} & \text{for } I\text{–}V \text{ characteristics (4)}\end{cases} \quad (29)$$

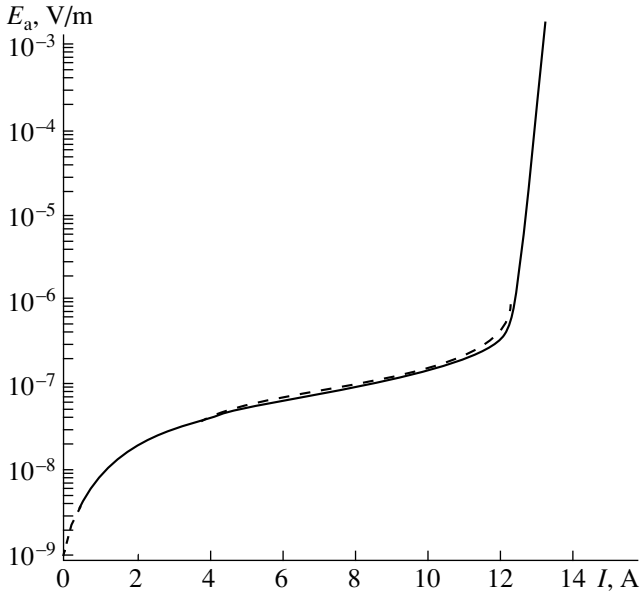


Fig. 7. Electric field strength increase at the surface of a cylindrical composite superconductor. Dashed line is a zeroth approximation (31), (32), solid line is a numerical model (33).

with initial and boundary conditions

$$E(r, 0) = 0, \quad \frac{\partial E}{\partial r}(a, r) = \frac{\mu_0}{2\pi a}\frac{dI}{dt} \tag{30}$$

and a corresponding current conservation condition

$$\int_S J ds = I_t, \quad t > 0,$$

where S is the current flow area and $I_t = tdI/dt$ is the total current injected into the composite.

For the solutions of (29) and (30), we will use the dimensionless variables in which the composite radius is used as the characteristic linear dimension introduced above. We transform the divergent part of the starting equations by substituting variables $\rho = 1 - R$, $R = r/a$, and using for $\rho \le 1$ an expansion $(1 - \rho)^{-1} \sim 1 - \rho + \ldots$. With respect to the new variable $y = \ln(1 - \rho)$, the problems (29) and (30) will take the form

$$\frac{\partial^2 e}{\partial y^2} = \begin{cases} \frac{1}{n} e^{(1-n)/n}\frac{\partial e}{\partial \tau}, \\ \frac{1}{e + e_0}\frac{\partial e}{\partial \tau}, \end{cases}$$

$$e(y, 0) = 0, \quad \frac{\partial e}{\partial y}(0, \tau) = q = \frac{\mu_0}{2\pi E_C}\frac{dI}{dt},$$

the solution to which at $q = -1$ was presented above. Therefore, it is easy to obtain the corresponding analytical expressions in a cylindrical coordinate system describing the distribution of the electric field strength induced in the superconducting composite by the injected current. In a zeroth approximation, as in the critical state model [5], we obtain

$$E(r, t) = \frac{\mu_0}{2\pi}\frac{dI}{dt}\ln\frac{r}{r_0(t)}, \tag{31}$$

where $r_0(t)$ is the moving boundary of the current flow area, which according to the current conservation condition is determined from the equation

$$I_t = 2\pi\eta J_C\left(\frac{\mu_0}{2\pi E_C}\frac{dI}{dt}\right)^{1/n}\int_{r_0}^{a}\left(\ln\frac{r}{r_0}\right)^{1/n} r dr \tag{32}$$

for the power-law I–V characteristics (2), and from

$$I_t = \pi n J_C(a^2 - r_0^2) + 2\pi\eta J_\delta\int_{r_0}^{a}\ln\left(\frac{\mu_0}{2\pi E_C}\frac{dI}{dt}\ln\frac{r}{r_0} + \frac{E_0}{E_C}\right) r dr$$

for the exponential I–V characteristics (4).

As an illustration of the validity of the these approximate expressions, the calculated results are presented in Fig. 7 for the electric field strength at the composite surface $E_a = E(a, t)$ and a power-law I–V characteristic obtained using formulas (31) and (32) along with the numerical solution of the system of the Fourier and Maxwell equations

$$c\frac{\partial T}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(\lambda r\frac{\partial T}{\partial r}\right) + EJ, \quad \mu_0\frac{\partial J}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial E}{\partial r}\right), \tag{33}$$

$$J = \eta J_s + (1 - \eta)J_m, \quad E = E_C(J_s/J_C)^n = J_m\rho_m$$

with the initial boundary conditions

$$T(r, 0) = T_0, \quad E(r, 0) = 0,$$

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

which was considered for the generality of the analysis of the results. In the calculations, the following initial parameters were used: $a = 5 \times 10^{-5}$ m, $c = 1000$ J/(m^3 K), $\lambda = 200$ W/(m K), $h = 1000$ W/(m^2 K), $T_0 = 4.2$ K, $\eta = 0.5$, $\rho_m = 2 \times 10^{-10}$ Ω m, $J_C = 4 \times 10^9$ A/m^2, $E_C = 10^{-4}$ V/m, $n = 80$, $dI/dt = 1$ A/s, which correspond to a composite conductor based on the Nb–Ti superconductor in a copper matrix.

The calculations in a zeroth approximation were performed for the states preceding the complete filling of the composite cross section by a current when its temperature is essentially the same as the coolant temperature. In the numerical model, the stage of the current injection with complete filling of the composite cross section by current was considered as well. In Fig. 7, this is seen as the initial steady increase of the electric field strength and then its avalanche buildup

with a corresponding rise of the composite temperature above the coolant temperature and subsequent transition to the normal state.

Thus, the solutions found and the analysis performed on the basis of the peculiarities of the magnetic flux penetration process into nonideal type-II superconductors with different current-voltage characteristics show that the external electromagnetic field, as in the critical state model, cannot instantly penetrate into the bulk of the superconductor. As a result, the screening currents induced by the external magnetic field flow only in a limited area of the wire cross section even though the steady voltage inside the superconductor arises well before the disruption of the stability conditions of the superconducting state. The concurrent diffusion phenomena exhibit the formation of an electromagnetic wave inside the superconductor that propagates at a finite rate. As in the critical state model for the superconducting materials that retain superconductivity at liquid helium temperatures, the spatial distribution of the electric field strength inside superconductors with the considered types of nonlinearities of the *I–V* characteristics can be approximated with good accuracy by a linear dependence, but the chosen equation of motion of the magnetization region boundary should correspond to the type of *I–V* characteristic. At the same time, in contrast to the critical state model, at the moving boundary of the magnetization region in superconducting media with power-law and exponential *I–V* characteristics, all derivatives with respect to the space coordinate are equal to zero. Therefore, in the vicinity of the magnetic flux front, a smooth transition takes place between the value of the electric field strength induced by the varying external magnetic field or the carrier transport and the corresponding nonperturbed value.

In conclusion, the nonlinear character of the *I–V* characteristics is observed in a more pronounced form in high-temperature superconductors due to their well-known physical inhomogeneity and complexity of phenomena occurring at the superconductor–normalmetal interface. Current-voltage characteristics of forms (2) and (3) are also observed in high-temperature superconductors. Therefore, the main features of the diffusion processes in superconducting media with *I–V* characteristics of forms (2)–(4) considered in this study are also valid for the high-temperature superconductors.

ACKNOWLEDGMENTS

The study was supported by the Russian Foundation for Basic Research (grant N 98-02-16046a).

REFERENCES

1. A. M. Campbell and J. E. Evetts, *Critical Currents in Superconductors* (Taylor and Francis, London, 1972; Mir, Moscow, 1975).
2. P. W. Anderson, Phys. Rev. Lett. **9** (7), 309 (1962).
3. P. W. Anderson and Y. B. Kim, Rev. Mod. Phys. **36** (1), 39 (1964).
4. C. P. Bean, Phys. Rev. Lett. **8** (6), 250 (1962).
5. A. V. Gurevich, R. G. Mints, and A. L. Rakhmanov, *Physics of Composite Superconductors* (Nauka, Moscow, 1987).
6. M. N. Wilson, *Superconducting Magnets* (Oxford Univ. Press, London, 1983; Mir, Moscow, 1985).
7. R. G. Jones, E. H. Rhoderick, and A. C. Rose-Innes, Phys. Lett. A **24** (6), 318 (1967).
8. M. Polak, I. Hlasnik, and L. Krempasky, Cryogenics **13** (12), 702 (1973).
9. Yu. K. Krasnov, V. A. Shukhman, and L. V. Matyushkina, Fiz. Nizk. Temp. **5** (2), 109 (1979) [Sov. J. Low Temp. Phys. **5**, 51 (1979)].
10. G. L. Dorofeev, A. B. Imenitov, and E. Yu. Klimenko, Cryogenics **20** (6), 307 (1980).
11. L. F. Goodrich and F. R. Fickett, Cryogenics **22** (5), 225 (1982).
12. Yu. G. Kalmakhelidze and R. G. Mints, Cryogenics **29** (11), 1041 (1989).
13. V. M. Vinokur, M. V. Feigel'man, and V. B. Geshkenbein, Phys. Rev. Lett. **67** (7), 915 (1991).
14. A. Gurevich and H. Kupfer, Phys. Rev. B **48** (9), 6477 (1993).
15. I. L. Maksimov and I. V. Shalaev, Sverkhprovodimost': Fiz., Khim., Tekh. **4** (7), 1251 (1991).
16. A. L. Kasatkin, V. M. Pan, V. V. Vysotskii, *et al.*, Physica C (Amsterdam) **310**, 296 (1998).

Translated by M. Lebedev

Technical Physics, Vol. 45, No. 12, 2000, pp. 1566–1570. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 58–62.

RADIOPHYSICS

The Effect of Friction on Nonlinear Oscillations of Interacting Domain Walls in an External Periodic Field

M. M. Solov'ev and B. N. Filippov

Institute of Metal Physics, Ural Division, Russian Academy of Sciences, ul. S. Kovalevskoĭ 18, Yekaterinburg, 620219 Russia

Received December 20, 1999

Abstract—The effect of dissipation on nonlinear oscillations in a system of domain walls experiencing an external harmonic field is studied numerically. The problem is formulated for uniaxial ferromagnet films, with the easy magnetic axis being perpendicular to the surface and with the harmonic field being aligned with the axis. Account is taken of the dynamic redistribution of magnetic poles on the film surface, which enables one to derive, in a natural way, an expression for a restoring force acting on the domain walls. The force is a nonlinear function of domain-wall displacement from the equilibrium position. It is found that the domain walls may execute complicated steady-state quasi-periodic oscillations and long-term chaotic oscillations. Attractors in the phase space of the system are determined. *© 2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Earlier, we analyzed the nonlinear behavior of a system of domain walls (DWs) interacting via magnetostatic fields that result from the redistribution of magnetic charges on the film surface when the DWs are displaced from their equilibrium positions [1, 2]. A nonlinear equation was derived to reduce the dynamics of the system to that of a unit DW in a self-consistent field relating to the motion of the rest of the DWs. Conditions for the onset of deterministic chaos were found. This may indicate, e.g., the presence of inherent noise of ferromagnetic films, in addition to that produced by the intermittent DW motion due to inhomogeneities and defects in the crystal. Also note that the derived equation differs from any of those addressed by previous studies on dynamical systems. For the first time, the description implicitly includes a nonlinear restoring force representing the action of the magnetic charges. The equation thus allows for physical properties and dimensions of the film in a natural fashion, in contrast to the approach where the restoring force is modeled by an arbitrarily constructed function [3, 4]. As a result, many more quantities can be used as bifurcation parameters. In short, our strategy makes it possible to more carefully examine the nonlinear dynamics of DWs in a magnetic film, compared with previous studies [3, 4]. On the other hand, works [1, 2] pay little or no attention to dissipation, although this factor is known to play a key role in the dynamics (see, e.g., [5]). In particular, dissipation is responsible for the existence of an attractor in the phase space of the system. Therefore, this study addresses the effect of dissipation on the nonlinear behavior of DWs in a perpendicularly anisotropic magnetic film that experiences an external harmonic magnetic field aligned with the easy magnetic axis. We employ the above-mentioned dynamic equation [1, 2], which was modified to include damping. In practical terms, the investigation may throw new light on important properties of magnetic materials, such as electromagnetic loss or magnetic noise. This issue, however, deserves separate investigation.

FORMULATION OF THE PROBLEM

Consider a plane-parallel plate of thickness L made of a uniaxial ferromagnet with the easy magnetic axis perpendicular to the surface. The coordinate axes are oriented as in Fig. 1. Let the plate have a plane-parallel domain structure as shown in Fig. 1, where arrows designate the magnetization direction of the domains. In the absence of external fields, the domain structure has an equilibrium period $2D$.

Let the plate experience an external time-varying magnetic field $\mathbf{H} = \mathbf{H}_0\sin(\omega_0 t)$, where $\mathbf{H}_0$ is the field amplitude, ω_0 is the angular frequency, and t is time. The field makes the DWs oscillate about their equilibrium positions, the distance by which they are displaced being denoted by $x_0(t)$. The displacements result

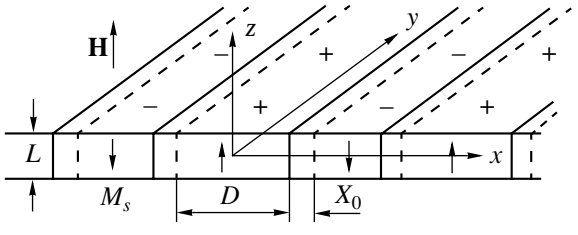


Fig. 1. Model of the domain structure in a thin uniaxial ferromagnetic film. L, the film thickness; D, the domain width in the absence of the field; and x_0, the displacement of a DW from the initial position.

1063-7842/00/4512-1566$20.00 © 2000 MAIK "Nauka/Interperiodica"

in the redistribution of magnetic poles on the surface of the plate, which, in turn, produces a magnetostatic field acting on the DWs so that they experience a restoring force. We determined this force for a periodic domain structure [1, 2]. As a result, the problem for a system of DWs was reduced to that for a single DW oscillating in a self-consistent magnetostatic field related to the displacements of the other DWs. The equation of motion for a single DW has the form

$$x''_{\tau\tau} + kx'_\tau + \left[x + \frac{2}{l}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\sin(nx)(1-\exp(-nl))\right] = \pi h\sin(\omega\tau). \quad (1)$$

Here,

$$x = 2\pi x_0/D, \quad l = \pi L/D, \quad h = H_0/4\pi M_S, \quad \tau = \Omega t, \quad \omega = \omega_0/\Omega, \quad k = k_0/\Omega, \quad \Omega = (8\pi M_S^2/mD)^{1/2}, \quad (2)$$

M_S^2 is the saturation magnetization of the plate, and m is the effective mass of the DW.

The only set of initial conditions we used is $x = 0$ and $x'_\tau = 0$. They imply that the DW cannot oscillate if no field is applied. In the presence of an external field, the forced oscillation may be a natural-frequency component (see, e.g., [6]). According to Eq. (1), the restoring force consists of a linear and a nonlinear term (see the expression in the brackets). If the latter is small, then, with $k = 0$ and $h = 0$, we easily find from Eq. (1) that the natural frequency is

$$\omega_p = \Omega = (8\pi M_S^2/mD)^{1/2}. \quad (3)$$

For example, setting $m = 10^{-10}$ g/cm^2, $M_S = 1700$ G, and $D = 10^{-2}$ cm, we have $\omega_p = 8 \times 10^9$ s^{-1}. In the relative units employed, the natural frequency is $\omega_n = 1$. It rises with the amplitude of the DW oscillation [7] and may reach a level well above unity. This nonlinear effect is more noticeable at smaller plate thicknesses.

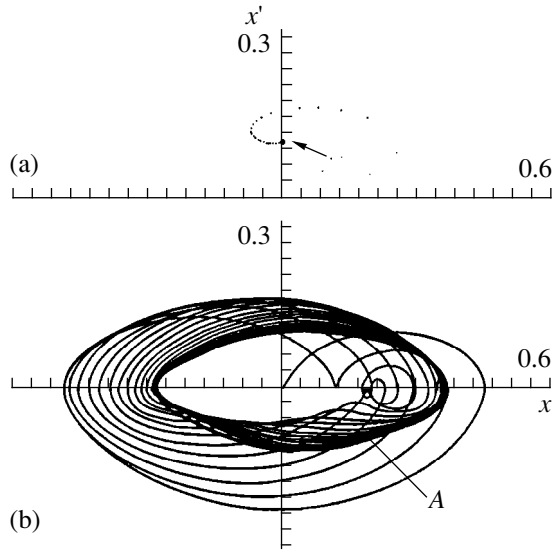


Fig. 2. DW oscillation for $k = 0.01$, $\omega = 0.3$, and $h = 0.033$: (a) the Poincaré plane and (b) the phase portrait, where A designates the attractor of the oscillation.

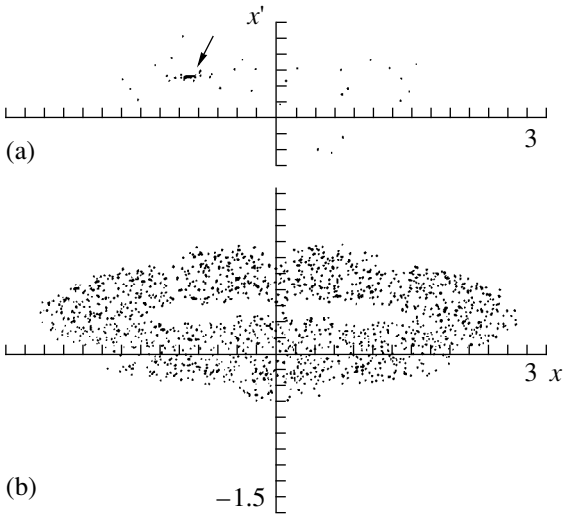


Fig. 3. Poincaré plane for $\omega = 0.3$, $h = 0.07$, and $k =$ (a) 0.01 and (b) 0.

RESULTS AND DISCUSSION

The computation was carried out for the relative film thickness $l = 1$. With this value of l, the nonlinear component of the restoring force is relatively large and becomes appreciable even at a modest x (for more details, see [1, 2]). We begin our discussion with the case of small fields $h \ll 1$. For $h = 0.033$, $\omega = 0.3$, and damping $k = 0.01$, results of the numerical experiment are shown in Fig. 2. If k were zero, h and ω being unchanged, then the forced DW oscillation would be a clear-cut quasi-periodic motion (represented by an ellipse in the Poincaré plane) and have a natural-frequency part. In the presence of damping ($k = 0.01$), we obtain a well-known linear oscillation (Fig. 2a). This is evidenced by the fact that, in the Poincaré plane, the solution tends to a point (designated by the arrow in Fig. 2a). In the phase plane, the motion is represented by an elliptic attractor (labeled A in Fig. 2b), which is symmetric with respect to the coordinate axes. The attractor corresponds to an oscillation at the driving force frequency. The elliptic shape of the attractor obviously stems from the fact that the DW oscillation excited at $h = 0.033$ has a small amplitude and is therefore almost linear.

Now, let us raise h to 0.07, with ω remaining at 0.3. If $k = 0$, the oscillation is chaotic (the corresponding Poincaré plane is shown in Fig. 3b). If damping is

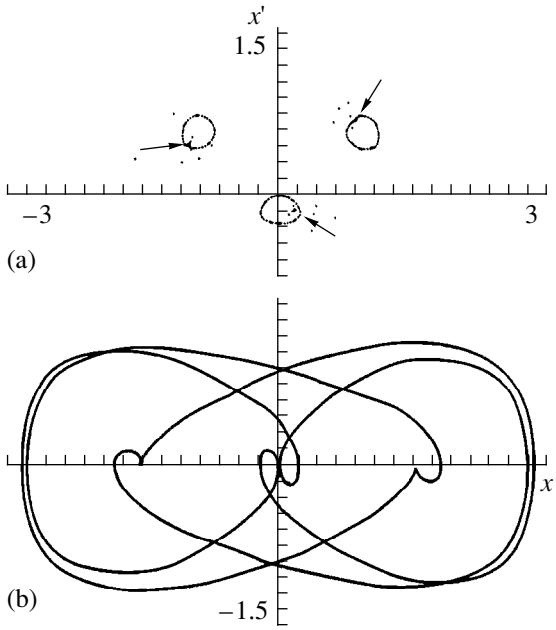


Fig. 4. DW oscillations at $\omega = 0.3$ and $h = 0.2$. (a) The Poincaré plane for $k = 0$ (three loops) and $k = 0.01$ (three points indicated by the arrows, which are approached as $t \longrightarrow \infty$). (b) The phase portrait of the steady-state oscillation for $k = 0.01$.

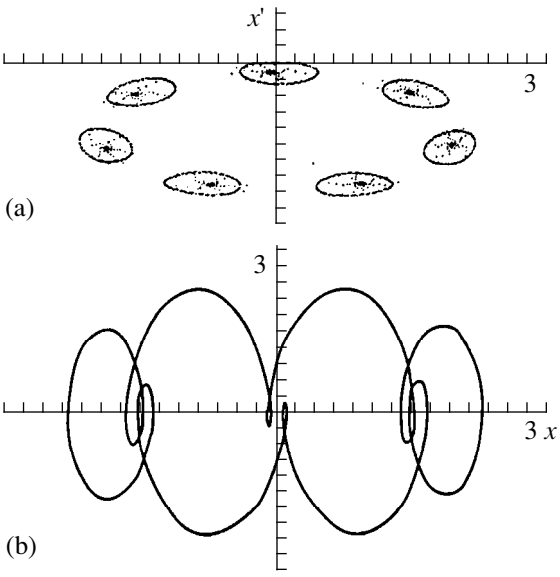


Fig. 5. DW oscillations at $\omega = 4$ and $h = 1.6$. (a) The Poincaré plane for $k = 0$ (seven loops) and $k = 0.01$ (seven points, which are approached in the limit $t \gg 1$). (b) The phase portrait of the steady-state oscillation in the presence of damping.

present, $k = 0.01$, the chaos decays (Fig. 3a) and a single-period oscillation arises. In contrast to the case $h = 0.033$, the oscillation is not harmonic, since its attractor is not an ellipse. Thus, we face effects of the problem.

If friction is reduced to $k = 0.001$ ($h = 0.07$), the transition to an attractor takes considerable time: the picture in the Poincaré plane remained chaotic even at the time $\tau_k = 20000$ [for the definition of τ, see (2)]. Indeed, the duration of the transition depends both on k and on h. For example, if h is taken slightly larger, $h = 0.09$, then the chaos disappears at a much smaller τ_k.

Friction-induced suppression of chaos may be difficult to reveal on the basis of a numerical experiment, because the necessary time interval may be very long. For example, at $h = 0.07$, $\omega = 0.3$, and $k = 0.001$, the chaotic oscillation was observed in the Poincaré plane even after $\tau_k = 40000$; an abrupt transition to the attractor occurs near the end of the interval (40000–60000).

A chaotic motion was obtained also for $h = 0.1$, $\omega = 0.1$, and $k = 0.001$. It exists at least up to $\tau_k = 60000$.

If h is raised to 0.2, with $\omega = 0.3$, then we have a quasi-periodic motion with three periods. Figure 4a illustrates this by showing three different loops in the Poincaré plane. The motion is a triple-period oscillation onto which a natural-frequency oscillation is superimposed. In the presence of damping, the latter oscillation dies away and one can see three points in the Poincaré plane, which are due to period tripling. This effect essentially stems from the nonlinearity of the restoring force.

Figure 4b is the phase portrait plotted for a time interval when the natural-frequency part is no longer existent.

If h is increased further, the oscillation takes on some new features. We will here enlarge only on those for very small ω. With $h = 0.3$, $\omega = 0.06$, and $k = 0$, the DW has time to execute several cycles during a period of the external field, since the natural frequency is much higher than that of the driving field. Figuratively speaking, the field drags the DW oscillating at the natural frequency. The number of cycles performed by the DW is maximum at the points where the DW velocity changes sign and the DW begins to shift in the reverse direction.

Damping suppresses the natural-frequency component, and eventually the DW moves more slowly, on the whole. Nevertheless, the resultant forced oscillation is not harmonic, as evidenced by computation carried out at $k = 0.1$ and even at $k = 1$, with $h = 0.43$, $\omega = 0.06$, and $\Delta\tau$ from 2500 to 3500.

In the above, ω was much lower than the linear natural frequency $\omega_n = 1$. At a higher ω, the oscillation depends on h in a similar fashion. For example, if $\omega = 1$, $h = 0.2$, and $k = 0.01$, the oscillation observed after $\tau_k = 1000$ is regular with a doubled period.

Let us examine the case of still higher ω's, e.g., $\omega = 4$. Then h must be taken much larger so as to avoid the region of linearity. Figure 5a illustrates the motion in the Poincaré plane for $h = 1.6$ and $k = 0$. Notice that we now have seven different loops, which represent a quasi-periodic solution with seven periods. If friction is

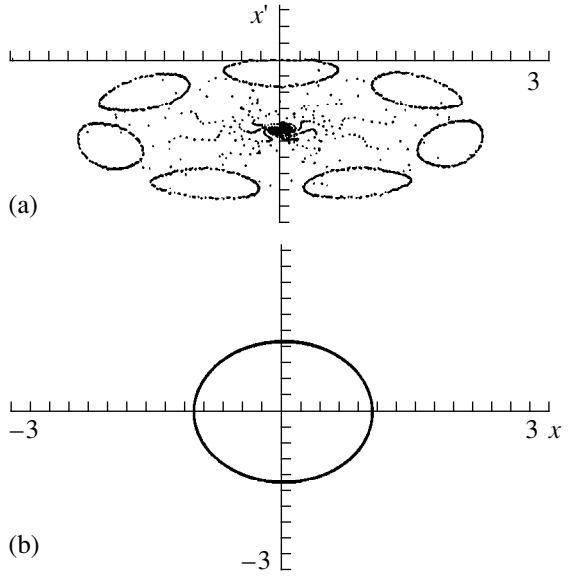


Fig. 6. DW oscillations at $\omega = 4$ and $h = 1.65$. (a) The Poincaré plane for $k = 0$ and (seven loops) $k = 0.01$ (the single point to which the loops contract). (b) The phase portrait that corresponds to a steady-state harmonic forced oscillation.

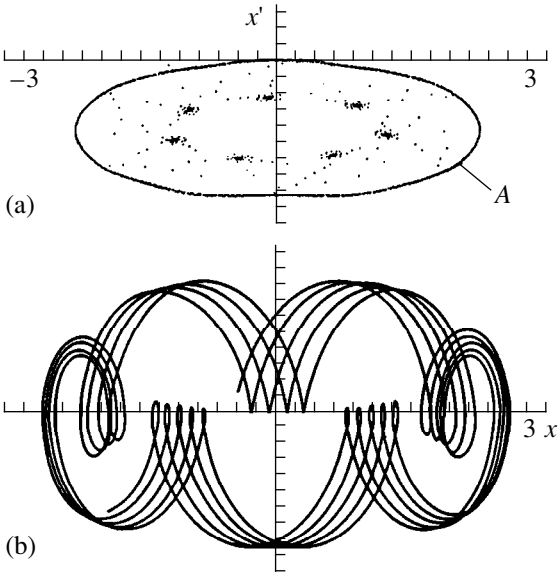


Fig. 7. DW oscillations at $\omega = 3.5$ and $h = 1.4$. (a) The Poincaré section, which is an ellipse if $k = 0$ or eventually turns into the seven points at $k = 0.01$. (b) A segment of the phase trajectory, with damping being neglected.

imparted ($k = 0.01$), the loops rapidly contract to respective interior points (Fig. 5a) for $\tau_k = 200$. Figure 5b displays the phase portrait that arises once the natural-frequency part has died away. Thus, the results are, in general terms, similar to those for $\omega = 0.3$. If, however, h is varied further, we see qualitatively different behavior in the presence of damping. For example, let h equal 1.65 with $\omega = 4$. If $k = 0$, we have seven loops in the Poincaré plane, which correspond to a seven-period motion (Fig. 6a), just as at $h = 1.6$. However, setting k equal to 0.01 changes the motion radically so that a single point appears in the Poincaré plane. Figure 6a depicts the transition to this point (situated at the center of the dark spot), which is observed up to $\tau_k = 2000$. Afterward, the phase portrait becomes an ellipse (Fig. 6b).

Comparing Figs. 5 and 6, we see that a bifurcation occurs when the value of h is changed from 1.6 to 1.65: the seven-period solution becomes a forced harmonic oscillation. It is interesting to look at the case $\omega = 3.5$ and $h = 1.4$. With zero damping, there is only one loop in the Poincaré plane (labeled A in Fig. 7a). Consequently, we are dealing with a quasi-periodic motion resulting from the superposition of a forced and a natural-frequency oscillation. The corresponding phase trajectory has an intricate shape. Figure 7b shows its fragment computed for a short interval of τ ($\Delta\tau_k = 50$). Allowing for damping, we obtain a seven-point picture in the Poincaré plane, corresponding to a seven-period solution, as in the preceding case ($\omega = 4$, $h = 1.6$). Figure 7a depicts the convergence to this regime for τ_k up to 2000. The transition to the seven-period solution is also indicated by the phase trajectory, which virtually does not change after $\tau_k = 2000$ and closely resembles the trajectory in Fig. 5b.

CONCLUSIONS

We presented the results of a numerical investigation into the effect of damping on DW oscillations (including nonlinear ones). The conclusions are as follows.

Damping may dramatically affect different types of DW oscillation. For example, it may suppress quasi-periodic or chaotic oscillations. Sufficiently strong damping, namely, that with $k \geq 0.1$, may result in a single-period oscillation that may be nonharmonic. On the other hand, with modest damping, $k = 0.01$–0.001, the DW may perform a quasi-periodic oscillation with many periods and even a chaotic oscillation. We strongly believe that the newly discovered features of nonlinear oscillations in a periodic DW system are closely related to many useful properties of ferromagnetic plates and films, especially to electromagnetic power loss. Furthermore, deterministic chaos and related types of behavior should be taken into account when dealing with noise in ferromagnets. The chaotic dynamics of DWs must result in unavoidable magnetic noise, which does not proceed from defects in the specimen but rather is inherent in the material itself.

ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research (grant no. 99-02-16279).

REFERENCES

1. M. M. Solov'ev and B. N. Filippov, Fiz. Met. Metalloved. **81** (5), 49 (1996).
2. M. M. Solov'ev and B. N. Filippov, Fiz. Tverd. Tela (St. Petersburg) **39** (11), 2036 (1997) [Phys. Solid State **39**, 1821 (1997)].
3. H. Okuno and T. Homma, IEEE Trans. Magn. **29** (6), 2506 (1993).
4. A. Sukiennicki and R. A. Kosinski, J. Magn. Magn. Mater. **129** (2–3), 213 (1994).
5. M. M. Solov'ev, B. N. Filippov, and A. N. Kalashnikov, Fiz. Met. Metalloved. **81** (5), 57 (1996).
6. I. M. Babakov, *Theory of Oscillations* (GITTL, Moscow, 1958).
7. B. N. Filippov and M. M. Solov'ev, Fiz. Met. Metalloved. **80** (2), 20 (1995).

Translated by A. Sharshakov

Technical Physics, Vol. 45, No. 12, 2000, pp. 1571–1574. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 63–66.

RADIOPHYSICS

The Energy Spectrum of an Electron Beam after Interaction with an RF Field in a Gyrotron

N. P. Venediktov, M. Yu. Glyavin, A. L. Goldenberg, V. E. Zapevalov, A. N. Kuftin, M. A. Moiseev, and A. S. Postnikova

Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhni Novgorod, 603600 Russia

Received June 16, 1998

Abstract—The electron distribution over residual energy after electron interaction with an rf field was determined using the method of decelerating electric field. It is demonstrated that, under the maximum efficiency conditions, electrons with an energy lower than 30% of their initial energy are absent in the spectrum. Hence, the efficiency of energy-recuperation gyrotrons can substantially be increased. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Gyrotrons, or high-power oscillators of mm-wave coherent electromagnetic radiation [1–3], are applied for electron–cyclotron heating of a thermonuclear plasma and high-temperature processing of materials. Typical efficiencies of modern gyrotrons are about 30–40% [4]. A simple and efficient way of improving the efficiency is recuperation—a process where electrons leaving the operating region are decelerated by the electric field near the collector (whose potential is lower than the initial potential of the beam) and transfer a portion of the kinetic energy remaining after their interaction with the rf field to the power supply. Recuperation in gyrotrons seems promising, because, after interaction, there are few (if any) electrons with an energy lower than a certain value $E_{\min}$ (that may reach 40% of the initial energy E_0). Therefore, recuperation can be efficient even in the simplest single-stage process. Thus, the energy recuperation efficiency essentially depends on the energy spectrum of electrons that have interacted with the rf field ("used-up" electrons). This spectrum, in turn, strongly depends on the interaction conditions, especially in the transition region between the resonator and the output waveguide [5]. The energy spectra have been thoroughly studied theoretically [6–8]. However, direct measurement of the energy distribution for used-up electrons is important for both analyzing gyrotrons operating in the single-stage recuperation regime and designing gyrotrons with two- and multistage recuperation. The multistage recuperation flowchart must rely on reliable information on the energy spectrum of used-up electrons and accurate energy spectra. In this work, we experimentally determine the energy spectrum of an electron flow that has interacted with the rf field at 83 GHz in a short-pulse gyrotron with an output power to 1 MW [9].

METHOD OF MEASUREMENT AND EXPERIMENTAL SETUP

The electron energy spectrum was determined using the method of decelerating field. In this method, electrons with a translational velocity sufficient to overcome the decelerating electric field produce the target current, which depends on the decelerating voltage. From this dependence, one can find the electron velocity distribution. The analyzer used in this work (Fig. 1) and the measurement technique [10] enable one to perform measurements at operating beam voltages of about several tens of kilovolts. Used-up electrons enter the analyzer placed near the collector. Approximately 1% of the total beam current arrives at the analyzer through a narrow longitudinal slot in the collector. (One can find the azimuth distribution of the electron flow using a rotating collector. However, in this work, we did not perform these measurements.) Simple estimates based on the adiabatic approximation [11, 12] demonstrate that, within the decelerating field of the analyzer, the rotational velocity of the electrons is negligibly small compared to their translational velocity, aligned with the magnetic field. In this case, the method of decelerating field makes it possible to find the electron distribution over total velocity (energy). During measurements, rectangular pulses were applied to the analyzer grid simultaneously with accelerating voltage pulses applied to the cathode. The amplitude of each subsequent grid pulse was gradually decreased from the maximum value (of about 90 KV), depending on the power source and electric strength of the analyzer insulators, to zero. The signals, proportional to the target current and decelerating voltage, arrived at an automatic data acquisition and processing system. In our experiments, this system was built on the KAMAK module and HiCom (Dec) and IBM PC computers [13, 14], which provided a high accuracy and a high

1063-7842/00/4512-1571$20.00

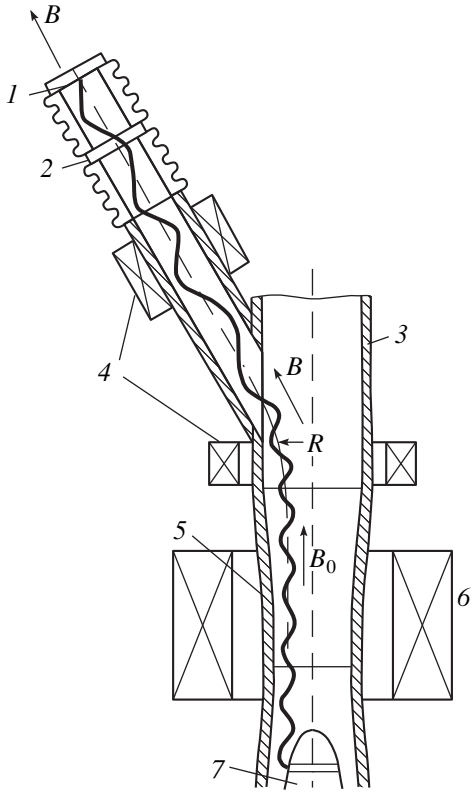


Fig. 1. Block diagram of the analyzer: (*1*) target, (*2*) grid, (*3*) collector, (*4*) additional solenoid, (*5*) gyrotron resonator, (*6*) main solenoid, and (*7*) electron gun.

speed of the measurements. In this way, errors due to long-term instability of power sources were reduced, and a large data array that improved the reliability of measurements was obtained. During data processing, interferences were suppressed by subtracting the signal arising in the measuring equipment in the absence of the oscillation current. The accuracy of measurements was determined using the dependence of the target current on the decelerating voltage in the absence of oscillation. These dependences have the form of a steeply sloping step when the grid voltage equals the beam accelerating voltage. The current fall width, which correlates with the spread in the beam electron energy, was about 3% of the accelerating voltage. This result is in good agreement with available data [15].

MEASUREMENTS AND CALCULATIONS

The purpose of the experiments described below was to find the energy spectrum of the electron flow that had interacted with the rf field in the resonator of the experimental gyrotron [9, 16]. The gyrotron's operating parameters were operating mode $TE_{11,3}$, oscillation frequency 83 GHz, output power 1 MW, pulse duration 100 μs, accelerating voltage 70 kV, beam current 40 A, and magnetic field 3 T. In the experiments, the accelerating voltage, magnetic field, and beam current were varied. We will report the basic findings derived from the electron energy spectra.

Figure 2 shows the energy distributions of electrons after their interaction with the rf field at different magnetic fields within the oscillation band. The current and voltage have nominal values (U_0 = 70 kV, I = 30 A). In the distribution corresponding to the oscillation mode with the maximum efficiency (of about 40%), at the optimal cyclotron resonance detuning $B = B_{opt}$ [17], most of the electrons have an energy that is approximately half the initial electron energy E_0 (which is governed by the total accelerating voltage U_0 before the interaction with the rf field). At the left of the maximum (at lower energies), the number of electrons is small, while at the right, the distribution function has a gentle slope and contains electrons with an energy exceeding the initial value. Such an energy distribution is in good agreement with calculations [18]. The energy spectrum expands and its maximum shifts towards the initial electron energy as the resonance detuning increases. In most cases when the beam current and voltage and the magnetic field differ from the values corresponding to optimum microwave oscillation, the spectrum changes in a similar manner.

The accuracy of measurements can be determined by estimating the gyrotron efficiency from the energy spectrum. In view of the normalization of the distribution function $\int f(E)dE = 1$, the gyrotron efficiency is given by

$$\eta = 1 - \frac{\int_0^{E_{max}} f(E)EdE}{E_0}. \quad (1)$$

For the high-efficiency (40–35%) oscillation regime, the efficiencies estimated from (1) and from the calorimetric measurement of the output power coincide. As the efficiency decreases, the greater the departure of the magnetic field from the optimum value, the greater the overestimate of the efficiency obtained from the energy spectrum becomes. Apparently, this discrepancy is attributed to a large number of electrons in the spectrum that have an energy much higher than the initial value and are not detected experimentally. This is also indicated by the behavior of $f(E)$ near the maximum decelerating voltage.

Figure 3 illustrates the evolution of the electron energy spectrum with increasing accelerating voltage U_0 at a fixed magnetic field of the solenoid and a beam current I = 20 A. In the absence of oscillation (U_0 = 16 kV), the spectrum has the form of a narrow peak with the top corresponding to the accelerating voltage. The width of the peak is several percent, which is in good agreement with earlier results [10]. When the accelerating voltage exceeds the threshold value (about 30 kV), weak microwave oscillation starts. This spec-

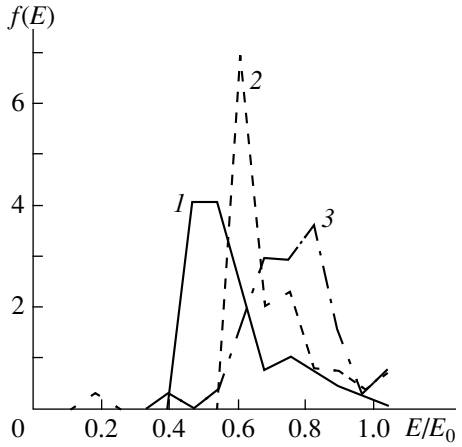


Fig. 2. Electron energy distribution within the oscillation band of one mode at magnetic field intensities (*1*) $B = B_{opt}$, (*2*) $B \cong 1.015B_{opt}$, and (*3*) $B \cong 1.02B_{opt}$. Efficiency is (*1*) 40, (*2*) 20, and (*3*) 10%.

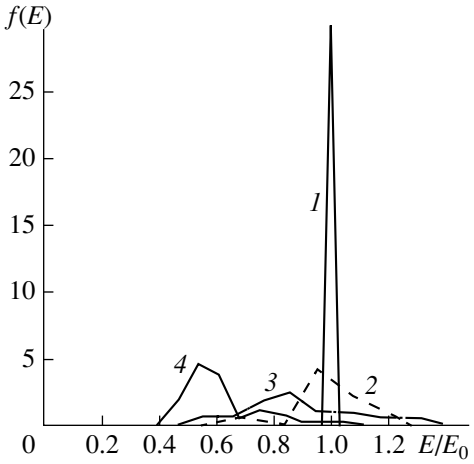


Fig. 3. Electron energy distribution at accelerating voltages U_0 = (*1*) 16, (*2*) 33, (*3*) 50, and (*4*) 70 kV. Efficiency (*1*) $\longrightarrow$ 0, (*2*) ≤5, (*3*) ≈10, and (*4*) ≈40%.

trum (U_0) = 33 KV) is substantially wider than the previous one, suggesting considerable energy exchange between the electrons and the rf field. However, if the efficiency is not high (≤5%), this interaction is linear, the numbers of decelerated and accelerated particles are almost the same, and the maximum of the spectrum is slightly lower than the accelerating voltage. As the voltage grows further (U_0 = 50 KV; the efficiency is about 10%), the maximum shifts towards lower energies. In this case, the electron energy spectrum is wide and may have several peaks. The presence of low-energy electrons may decrease the gyrotron efficiency and power in experiments on recuperation [19], because, when the collector potential lowers, low-energy electrons are reflected from the potential barrier and return to the resonator, thus deteriorating the conditions of interaction between the electron beam and the rf field of the operating mode [6]. In this case, at different beam voltages up to the optimum value U_0, the spectrally estimated efficiency is much closer to that found by directly measuring the output power. In the high-efficiency oscillation regime (U_0 = 70 KV, the efficiency is about 40%), the spectrum maximum roughly corresponds to half the accelerating voltage, and electrons with an energy lower than $0.4E_0$ escape from observation. The spectrum is much narrower than under nonoptimal interaction conditions and is in good agreement with the experimental spectrum found for another high-efficiency regime (Fig. 2). The difference is in the 3% shift of the maximum towards larger energies and in a greater number of electrons with an energy $E > E_0$.

Of interest is to compare the experimental results with the calculated electron energy spectrum. The electron–field interaction was simulated using the model described in detail in [18, 20]. The spectrum $f(\nu_\perp)$ of electron oscillatory velocities at the entrance to the operating space of the tube was approximated by a Gaussian distribution similar to $f(\nu_\perp)$ distributions observed in optoelectronic experiments [21]. Since the electrons continue interacting with the rf field (with decreasing intensity of interaction) outside the resonator (in the output waveguide at distances comparable to resonator length), in numerical calculations the domain of integration of the gyrotron equations was chosen sufficiently long so that further calculations did not change the calculated efficiency by more than 1%. The calculations were performed for the resonator profile and magnetic field distribution close to real ones.

The experimental and calculated energy spectra for the gyrotron operating in the optimum oscillation regime are shown in Fig. 4. They have differences that are essential when the gyrotron operates in the recuperation regime. For example, the calculated minimum electron energy E_{min} in the beam is about 20% of the

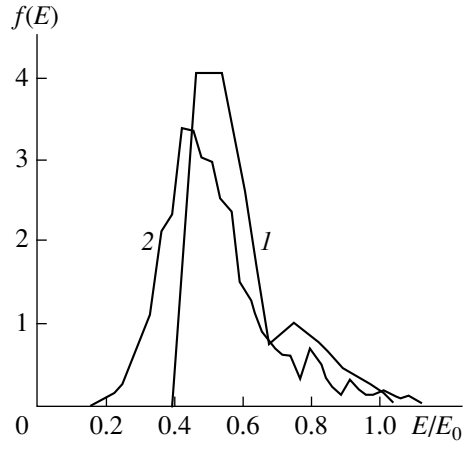


Fig. 4. (*1*) Measured and (*2*) calculated electron energy distributions in the maximum efficiency regime.

initial electron energy, while the measured value is about 40%. Previously, it was noticed that $E_{\min}$ essentially depends on the longitudinal distribution of the rf field in the resonator. Probably, these distributions differed in the calculations and experiments. In any case, the existence of spectra with $E_{\min} \approx 0.4E_0$ was indirectly confirmed by a considerable increase in the efficiency (by a factor of 1.5–1.6) in experiments with energy-recuperation gyrotons [19]. The presented spectrum is typical. It was reproduced for a number of gyrotrons with different electrodynamic structures and operating modes. In recuperation gyrotrons, the maximum difference between the collector potential $U_{\max}^{\text{col}}$ and the accelerating voltage was about 40%. As U^{col} was decreased, the output power dropped and the current to the tube case increased, which was caused by a substantial increase in the number of electrons reflected from the collector.

CONCLUSION

In the experiments, we obtained the energy spectrum of the electron flow in different regimes. It can be argued that the spectrum strongly depends on oscillation conditions. Under conditions for maximum efficiency, the spectrum originates at energies of about 30–40% of the initial beam energy. For single-stage recuperation, the efficiency can be raised by approximately a factor of 1.5, which is sometimes observed in gyrotrons with the decreased potential of the collector. The measured and calculated efficiencies agree well, but the energy distributions of used-up electrons somewhat differ. The latter fact indicates the need for improving the calculating model.

REFERENCES

1. A. V. Gaponov, M. I. Petelin, and V. K. Yulpatov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **10**, 1414 (1967).
2. A. L. Goldenberg, G. G. Denisov, V. E. Zapevalov, *et al.*, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **39**, 635 (1996).
3. *Gyrotrons. Collection of Scientific Works*, Ed. by V. A. Flyagin (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1989).
4. V. A. Flyagin, A. L. Goldenberg, and V. E. Zapevalov, in *Proceedings of the International Workshop "Strong Microwaves in Plasmas," Nizhni Novgorod, 1993*, p. 597.
5. A. L. Goldenberg, N. A. Zavol'skiĭ, V. N. Manuilov, *et al.*, in *Proceedings of the Conference "Modern Problems of Electronics and Microwave Radiophysics," Saratov, 1997*, p. 24.
6. K. Sakamoto, M. Tsuneoka, A. Kasugai, *et al.*, Phys. Rev. Lett. **73**, 3532 (1994).
7. B. Piosczyk, C. Latrou, G. Dammertz, *et al.*, IEEE Trans. Plasma Sci. **24**, 579 (1996).
8. V. L. Bratman, G. G. Denisov, and A. V. Savilov, Int. J. Infrared Millim. Waves **16**, 459 (1995).
9. N. P. Venediktov, V. E. Zapevalov, and A. N. Kuftin, in *Gyrotrons. Collection of Scientific Works*, Ed. by V. A. Flyagin (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1989), p. 12.
10. N. P. Venediktov, M. Yu. Glyavin, A. L. Goldenberg, *et al.*, Zh. Tekh. Fiz. **70** (4), 95 (2000) [Tech. Phys. **45**, 476 (2000)].
11. A. L. Goldenberg and T. B. Pankratova, Élektron. Tekh., Ser. 1: Élektronika SVCh, No. 9, 81 (1971).
12. A. L. Goldenberg, V. K. Lygin, V. N. Manuilov, *et al.*, in *Gyrotrons* (Gorki, 1981), pp. 86–101.
13. S. I. Artyuch, A. N. Kuftin, A. S. Postnikova, *et al.*, Int. J. Electron. **72**, 1145 (1992).
14. A. N. Kuftin, V. K. Lygin, A. S. Postnikova, *et al.*, in *Proceedings of the 8th Joint Russian–German Meeting on ECRH and Gyrotrons, Nizhni Novgorod, 1996.*
15. V. L. Bratman, M. Yu. Glyavin, A. L. Goldenberg, *et al.*, in *Conference Digest of the 22nd International Conference on IRMM Waves, Wintergreen, 1997*, pp. 186–187.
16. V. A. Flyagin, A. N. Kuftin, A. G. Luchinin, *et al.*, in *Proceedings of the Joint IAEA Technical Committee Meeting on ECE and ECRH (EC-7 Joint Workshop), Hefei*, pp. 355–372.
17. N. P. Venediktov, M. Yu. Glyavin, V. E. Zapevalov, *et al.*, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **41**, 670 (1998).
18. A. L. Goldenberg, V. N. Manuilov, M. A. Moiseev, *et al.*, Int. J. Infrared Millim. Waves **18**, 43 (1997).
19. N. S. Ginzburg, G. S. Nusinovich, and N. A. Zavolsky, Int. J. Electron. **61**, 881 (1986).
20. A. N. Kuftin, V. K. Lygin, V. N. Manuilov, *et al.*, in *Conference Digest of the 22nd International Conference on IRMM Waves, Wintergreen, 1997*, pp. 188–189.
21. G. S. Nusinovich and R. É. Érm, Élektron. Tekh., Ser. 1: Élektronika SVCh, No. 8, 55 (1972).

Translated by I. Efimova

Technical Physics, Vol. 45, No. 12, 2000, pp. 1575–1580. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 67–72.

SURFACE

Temperature Dependence of the Work Function of Hafnium Islands on Tungsten

O. L. Golubev, T. I. Sudakova, and V. N. Shrednik

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

Received January 14, 2000

Abstract—Two-dimensional islands of hafnium near {100} faces of tungsten and their terraces were examined by field electron microscopy. The two-dimensional sublimation energy of the islands was found to be 6.1 ± 0.5 eV. The field electron emission from the islands grows with temperature, the effect being considerably greater than theoretically expected. The negative temperature coefficient of φ, α_T, for the islands was estimated at (4–6) × 10^{-4} eV $grad^{-1}$ or more, whereas for pure tungsten α_T was found to be 2 × 10^{-5} eV $grad^{-1}$. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

When heated, hafnium atoms deposited in an amount of 0.05–0.2 monolayer on the single-crystal tungsten surface coalesce to form two-dimensional associations, or islands. They appear as bright spots on the substrate background when observed in a field electron microscope (FEM). The most contrasting Hf/W islands are formed on the {100} cubic faces and in their vicinity (Fig. 1). A pronounced temperature dependence of the work function in Zr/W islands (Zr is chemically close to Hf) was found [1]. Two-dimensional phase transitions in the Zr–W system (see, e.g., [2]) provided a great body of information on the activation energies of formation and decay of the islands and also on corresponding kinetic parameters responsible for preexponentials in the Arrhenius equations [3]. A deviation of these parameters from theoretically expected values has given rise to the ideas of "extended" adsorbate lattice and redistribution of interatomic bonds with the thermal expansion of the crystal [3]. The latter effect also had to show up as an anomaly of the temperature coefficient of the work function, which was the case [1].

The Hf–W adsorption system and the properties of Hf two-dimensional phases have been studied less comprehensively than the Zr–W system. It has been shown [4] that hafnium, like zirconium, produces contrasting emitting spots near the {100} faces of tungsten. The determination of the activation energy associated with two-dimensional phase transitions in the presence of nitrogen impurity in the adsorbed layer was described in [5]. Finally, the same parameters were found in [6] for the Hf–Mo system, which is close in properties to the Hf–W system. In all these works, carried out using a field electron microscope, the temperature dependence of the autoelectronic current was not the subject of investigation.

The main purpose of this work is to study temperature changes in the emission properties of Hf islands on W. As in the case of Zr on W [1], the discovery of a work function temperature anomaly would shed light on the nature of "nontheoretical" preexponentials in the Arrhenius equations for two-dimensional phase transitions.

EXPERIMENTAL

Experiments were carried out in a sealed-off FIM at a residual pressure of about 10^{-10} torr. The tungsten tip was cleaned at T up to 2500 K. A rounded single crystal oriented, as usual, in the [110] direction was formed on its top. Hafnium was evaporated from a small thin bar of metallic Hf welded to the molybdenum leads by the side of the tip. The bar was heated by passing the current. Usually, the deposition temperature of the Hf source was T = 2000–2100 K. The amount of deposited hafnium was controlled by the source temperature and evaporation time.

The voltage V between the tip and the anode was applied from a VS-20-10 stabilized power supply and measured by a V7-40 digital voltmeter. The emission current I was measured by an M-95 microammeter.

The tip temperature was controlled by the value of the current passing through the tip holder. The temperature was calibrated with a PYRO pyrometer and interpolated using the holder resistance.

The electric field strength F was determined from the slope of the Fowler–Nordheim current–voltage emission characteristic $\log(I/V^2) = f(1/V)$ for pure tungsten on the assumption that its work function φ = 4.4 eV. The work function of the hafnium-covered surface at a given F was also determined from the slope of the corresponding Fowler–Nordheim lines.

1063-7842/00/4512-1575$20.00

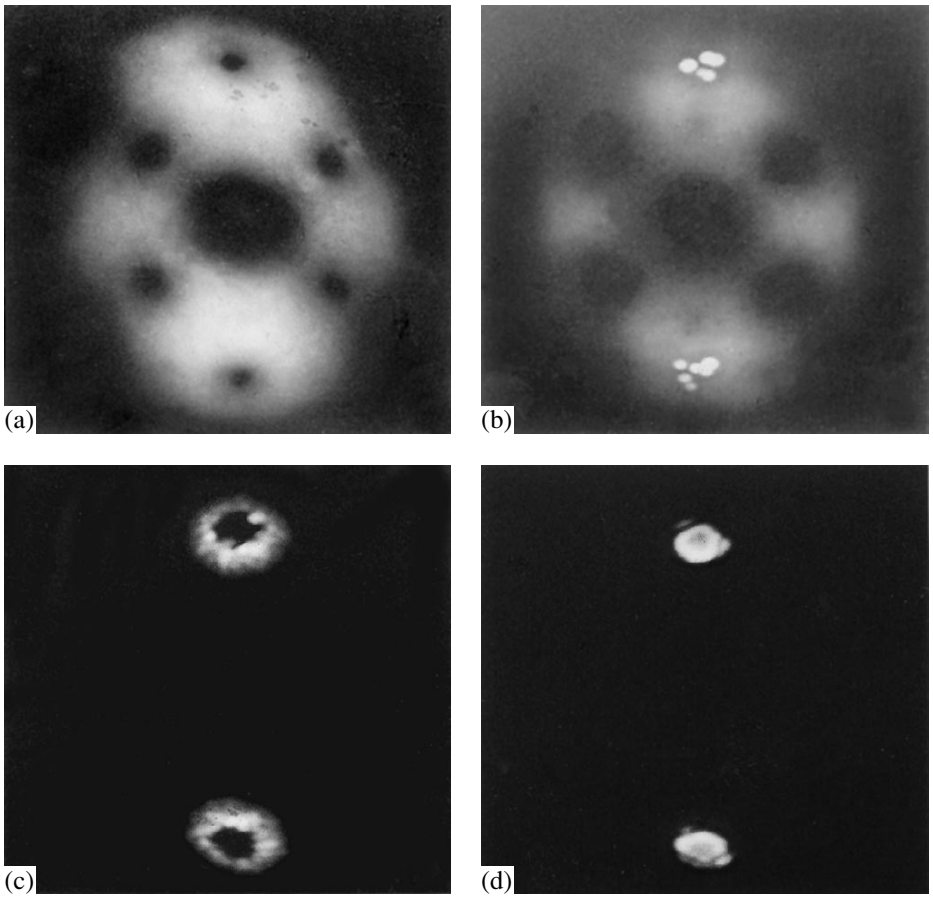


Fig. 1. FEM images of the (a) pure tungsten and (b–d) hafnium islands on its surface. The emission current is 50–70 nA. The crystal radius is about 5000 Å. The tip is kept at room temperature. (a) The imaging voltage is $V_0 = 6.5$ kV. (b) Hf islands in the {100} regions of W after depositing a small amount of Hf (about 2% of a monoatomic layer) and heating the tip to 1100 K for 2 min; $V_0 = 6.3$ kV. (c) Hf islands around the {100} poles of W after Hf deposition in an amount of about 10% of a monoatomic layer and heating the tip ($F = 0$) at 1550 K for 15 s; $V_0 = 5.4$ kV and work function $\varphi = 3.68$ eV. (d) Hf islands in the region of the {100} W poles after Hf deposition (10% of a monoatomic layer) and heating the tip ($F = 0$) at 1550 K for 1 min; $V_0 = 4.8$ kV and $\varphi = 3.3$ eV.

FORMATION OF HF ISLANDS NEAR THE {100} FACES AND THEIR THERMAL STABILITY

When deposited on pure tungsten (Fig. 1a), hafnium forms small islands in the {100} region (after thermal redistribution) even in small amounts (Fig. 1b, ≈2% of a monoatomic layer). When the amount of the deposit builds up to 10% of a monoatomic layer, the emission contrast and the island size increase as the two-dimensional hafnium phase approaches the {100} poles (Figs. 1c, 1d). The diameter of a Hf atom (3.14 Å [7]) is close to the lattice constant of W (3.1585 Å [7]). This makes for the formation of closely packed adsorbed Hf layers on the {100} W faces and on their terraces (i.e., in their vicinity). Adatoms in these layers sit in the deep potential wells and strongly interact with each other. Under migration equilibrium, Hf atoms accumulate in these regions. That the Hf islands (islands, not micropyramids) are flat is confirmed by their thermal stability and the pronounced temperature dependence of field emission [8].

A rather high energy is required to destroy two-dimensional islands (like those depicted in Fig. 1d). Two-dimensional sublimation of such islands occurs over measurable time intervals at temperatures of 1570–1700 K. Experiments on island dissolution at five temperatures allowed us to estimate the energy of two-dimensional sublimation at $Q_{sub} = 6.1 \pm 0.5$ eV from the slope of the Arrhenius plot (Fig. 2). This value is in agreement with the results obtained in [5][1], where Q_{sub} was determined as 6.4 and 6.8 eV in the case of Hf islands with nitrogen impurity near the {100} faces of *W*. We failed to obtain the Arrhenius plot for the growth of islands in this region. Early islands of the two-dimensional condensed phase were formed within 1 or 2 s of cooling the tip holder. (It is these early islands that are usually taken as a convenient reference for

[1] The preexponential determined from $\log t_0$ at $1/T = 0$ was found to be $t_0 = 7 \times 10^{-16}$ s, which also agrees with data in [5]. Data for t_0 are discussed in [3].

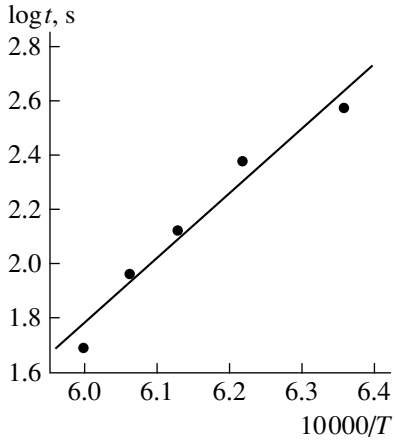


Fig. 2. Arrhenius plot for the time t of two-dimensional sublimation (dissolution) of Hf islands (like those shown in Fig. 1d). The activation energy of sublimation $Q_{sub} = 6.1 \pm 0.5$ eV, and the intercept on the ordinate axis is $\log t_0 = -15.15$, which corresponds to $t_0 = 7 \times 10^{-16}$ s.

measuring the phase growth.) In this case, we did not succeed in "quenching" the system in the state of two-dimensional gas. Nevertheless, the energy of island formation due to migration Q_m can be estimated as 1–2 eV by analogy with the Zr/W system [2]. Then the binding energy in a two-dimensional Hf crystal on {100} W will be $Q_b = Q_{sub} - Q_m = 4$–5 eV, which apparently is too high. Thus, one should assume (as in [5]) that nitrogen, which strengthens interatomic adsorbate–adsorbate bonds, is present on the surface. This makes the behavior of these bonds during crystal heating all the more interesting.

TEMPERATURE DEPENDENCE OF FIELD EMISSION AND WORK FUNCTION OF HAFNIUM ISLANDS ON TUNGSTEN

The temperature dependence of field emission was determined at temperatures that did not adversely affect the emitting surface. This was checked by emission patterns and by the return of the emission current to the initial (room-temperature) value. For Hf islands on W, this fact meant that the highest temperature in measuring I_T was much less than the temperature of island formation (1550 K). For pure W, the maximum temperature was less than that when the first signs of surface reconstruction were observed.

The field emission current as a function of temperature is presented in Fig. 3. Relative values of the current are shown: the field emission current I_T at a given temperature T is divided by the emission current from the same object at room temperature, I_R. Solid curves represent experimental results, which involve two components: the basic temperature effect of field emission (at a constant φ_R, corresponding to φ at room temperature) and an addition associated with a change in φ from φ_R

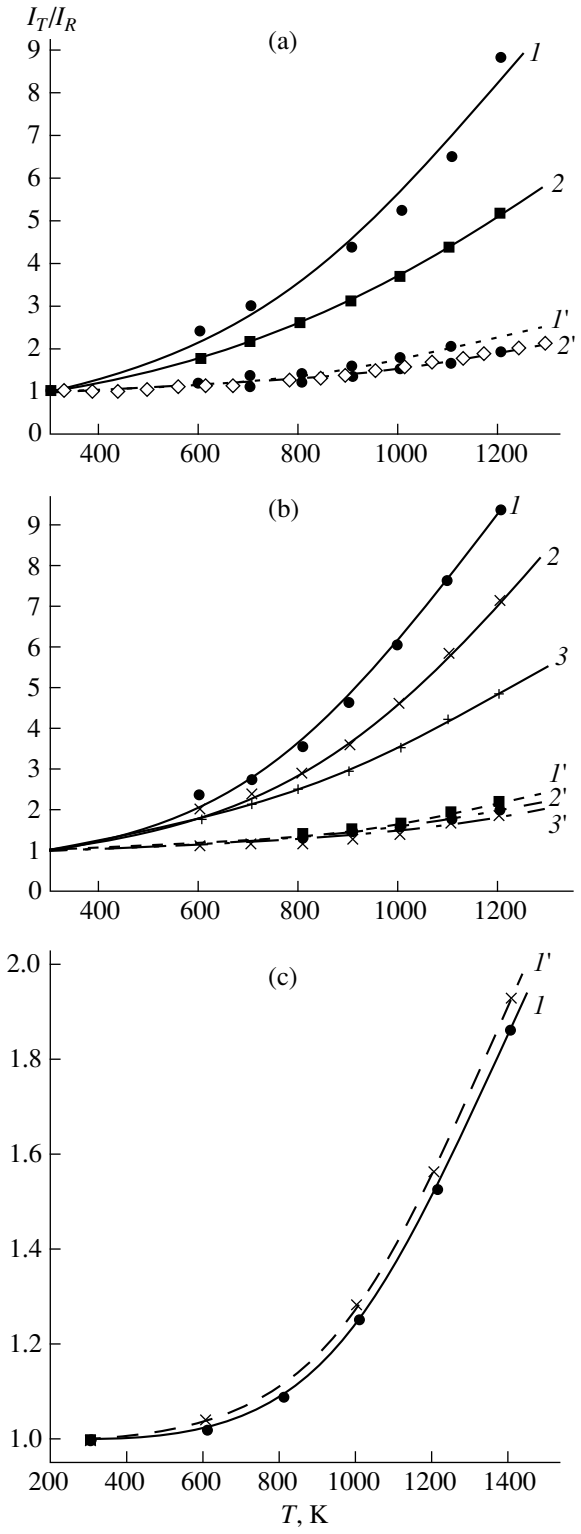


Fig. 3. Curves of the relative current increment I_T/I_R vs. temperature T, which reflect the temperature effect of field emission. (a) Hf islands like those in Fig. 1d ($\varphi = 3.3$ eV) with $F = (1)\ 2.95 \times 10^7$ and $(2)\ 3.22 \times 10^7$ V/cm; (b) Hf islands like those in Fig. 1c ($\varphi = 3.68$ eV) with $F = (1)\ 3.31 \times 10^7$, $(2)\ 3.45 \times 10^7$, and $(3)\ 3.65 \times 10^7$ V/cm; (c) pure W ($\varphi = 4.4$ eV), at $F = 4.48 \times 10^7$ V/cm.

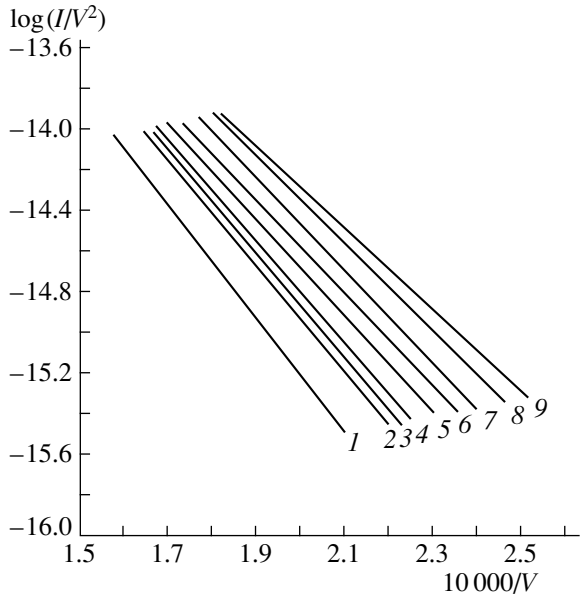


Fig. 4. A family of the Fowler–Nordheim straight lines for Hf islands like those in Fig. 1c for different temperatures (*1–9*, *T* and φ are given in Table 2).

to the value of the work function at an elevated temperature, φ_T. The first component can be accounted for in terms of the rigorous theory. Dashed curves (indicated by primed numbers) were calculated according to the Murphy–Good theory [9] (for details, see [1]). Note only that such calculations are correct at not too high I_T/I_R's. In practice, this means that I_T has to be such that I_T/I_R is less than 10. In our cases, this condition was met not only for the dashed lines (for which it must be fulfilled rigorously) but also for the solid experimental curves.

Measurements were performed for various kinds of islands (Figs. 3a, 3b) and at different fields, as well as for pure tungsten (Fig. 3c). Figures 3a and 3b show the strong temperature dependence of field emission from the Hf islands on W. This effect is much stronger than that associated with thermal excitation of electrons in the metal. The values of I_T/I_R over the dashed lines we relate to a change in $\varphi(T)$.

The dependence on F is qualitatively the same for the dashed and the solid curves in Figs. 3a and 3b: the less the value of F, the stronger the temperature dependence of emission. Physically, the dashed lines, obtained theoretically [9], mean that, for a less permeable potential barrier (smaller F), the fraction of thermally excited electrons appears to be more appreciable with respect to unexcited ones (although the absolute values of the currents, of course, diminish with decreasing F). When φ decreases, the greater relative effect at smaller F is associated with the increased emission contrast in terms of work function, which is observed more often in space. In our case, however, it is associated with temperature. The barrier height lowers (relatively) (φ decreases) more noticeably for a less permeable barrier. The physical reason for these effects is basically the same.

It is interesting to calculate the temperature coefficient of work function α_T. To do this, one should find, at a given T (or several T, s), the reduced φ (see Figs. 3a, 3b) such that, at a specified F, the solid curve runs over the dashed one. This reduced φ can be found in terms of the Fowler–Nordheim theory [10]. We will use the Fowler–Nordheim formula for the current density of field emission j in the simplified form convenient for calculations (formula (6.11) in [11]):

$$j = 1.4\times10^{-6}\frac{F^2}{\varphi}10^{4.39\sqrt{\varphi}}\times10^{-2.82\times10^7\frac{\varphi^{3/2}}{F}}. \qquad (1)$$

It is assumed that $\varphi = \varphi_R$ along the dashed curve. For the solid curve, φ_T will be a sought quantity. The quantity $\delta = I_{T\text{exp}}/I_{T\text{theor}}$ depends on the decrease in φ at a given T:

$$\delta = \frac{I_{T(\text{exp})}}{I_{T(\text{theor})}} = \frac{\varphi_R}{\varphi_T}10^{4.39(\sqrt{\varphi_T}-\sqrt{\varphi_R})}\times10^{\frac{2.82\times10^7}{F}(\varphi_T^{3/2}-\varphi_R^{3/2})}. \qquad (2)$$

Eq. (2) is transcendental with respect to φ_T and can be solved with the method of successive approximation. Recall that F and φ_R are given. Taking into account that the expected φ_T will differ little from φ_R, we assume at the first step that

$$C = \frac{\varphi_R}{\varphi_T}\times10^{4.39(\sqrt{\varphi_T}-\sqrt{\varphi_R})} \qquad (3)$$

is equal to 1. Then, we can easily calculate φ_T at a known δ. At the second step, we substitute this φ_T into (3) and the calculated C, already other than unity, into (2) and calculate the refined value of φ_T. Then, we substitute this new φ_T into (3), new C into (2), and so on. Three or four steps will typically suffice for corrections to φ_T be less than the experimental error of φ (in our case, it is not small and equals 0.1–0.2 eV). The calculated value of φ_T allows us to determine the temperature coefficient of φ as

$$\alpha_T = \frac{\Delta\varphi}{\Delta T} = \frac{\varphi_T-\varphi_R}{T-T_R}, \qquad (4)$$

where T_R is room temperature (300 K).

Such calculations are not rigorous, since φ was assumed to be equal to φ_R at any T, although it decreases with rising T according to α_T. To estimate the associated error, φ_T obtained for the greatest α_T and the highest T was taken as the initial value in the calculations of the dashed curve (obviously, such an approach gives an exceedingly high value of the error). It turned

Table 1

No.	Object	φ_R, eV	$F \times 10^{-7}$, V/cm	α_T, eV/grad	φ_T, eV (at 1100 K)
1	Hf islands on {100} faces of W	3.3	2.95	-4.0×10^{-4}	2.98
2	(Figs. 1d, 3a)		3.32	-4.0×10^{-4}	2.98
3	Hf islands around the {100} faces of W	3.68	3.31	-6.0×10^{-4}	3.2
4	(Figs. 1c, 3b)		3.45	-5.25×10^{-4}	3.26
5			3.65	-4.0×10^{-4}	3.36
6	Pure tungsten	4.4	4.48	$+2.0 \times 10^{-5}$	

out that the "refined" φ_T became 17% smaller, while α_T became 1.8 times greater. The actual error is smaller. It is clear that taking into account a decrease in φ with T only enhances the effect of large α_T. For simplicity, we carried out the calculations, ignoring the variation of φ along the dashed curves. Five approximations at every F were performed for the islands (Figs. 3a, 3b) at T = 1100 K. For pure tungsten (Fig. 3c), the one-step calculations were made. The results of calculations of α_T and φ_T at T = 1100 K are summarized in Table 1. Miscalculations of α_T associated with errors of φ_R and F measurement may reach 10%. If a decrease in φ with increasing T is taken into account when the field emission current is calculated, α_T may (in the case of Hf islands on W) increase further (but no more than twofold). Therefore, it is certain that α_T for the islands is one order of magnitude higher than for pure metals [12]. For a Hf monolayer on W, α_T is negative: φ drastically decreases with temperature. For tungsten, α_T is about 10^{-5} eV grad^{-1} and is positive, which is characteristic of porous W faces with the smallest φ and agrees well with known data [12].

In view of a low accuracy of α_T determination, one can argue that, for the five cases considered, α_T depends neither on F nor on φ or depends on them only slightly. Note also that, within the experimental accuracy, the choice of temperature for calculating α_T in the range from 600 to 1200 K is of little importance. For the selected curve in Fig. 3, the coefficients were almost the same.

Experiments on accurate determination of α_T for islands and similar two-dimensional phases are of special interest. This problem requires the determination of the initial φ_R's and acting fields F with a high accuracy and needs adequate mathematical processing with φ self-consistency in calculating solid and dashed curves (like those in Fig. 3). A correlation of α_T with F and φ_R may arise when the data are of great accuracy and reliability.

Based on our work and work [1], we can state with confidence that Hf and Zr islands on W feature anomalously high (in comparison with pure metals) negative temperature coefficients of work function α_T. This can be explained by the redistribution of interatomic bonds in the adsorbate–adsorbent system at thermal expansion of the crystal and at an increase in the interatomic distances in the already extended lattice of adatoms. In its turn, this effect can be associated with an essential temperature dependence of the bond energy in a two-dimensional crystal (the bond energy drops with increasing temperature). This fact furnishes insights into the anomalous values of the preexponentials in the Arrhenius equations describing the kinetics of two-dimensional sublimation [3].

In conclusion, let us note one more striking illustration of the pronounced temperature dependence of the work function of Hf islands on W. A family of the Fowler–Nordheim plots (straight lines) at increasing temperature (Fig. 4) was obtained for islands around the {100} poles (like the ones in Fig. 1c). The straight lines have a gradually decreasing slope and shift to the right. The effective work functions (listed in Table 2 along with corresponding temperatures) were calculated from the slopes of the characteristics. These values of φ are somewhat lower than those expected from α_T's for curves 3–5 in Table 1, since, in the latter case, the entire growth of emission is assigned to a drop in φ, whereas the effect according to the theory [9] is neglected.

Nevertheless, from a comparison between the effective φ's and those calculated through α_T, as well as from the curves in Figs. 3a and 3b, it follows that, for

Table 2

Curve no.	T, K	φ, eV
1	300	3.68
2	551	3.51
3	642	3.46
4	747	3.39
5	852	3.33
6	965	3.27
7	1079	3.21
8	1171	3.11
9	1265	2.96

the two-dimensional phases, the growth of emission due to thermal excitation of electrons is markedly weaker than the influence a decrease in φ with rising temperature has on emission. The volumetric electron effect, inherent in any metal, appears to be weaker than specific surface effects characteristic of the two-dimensional phase.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research, project no. 97-02-18066.

REFERENCES

1. E. L. Kontorovich, T. I. Sudakova, and V. N. Shrednik, Pis'ma Zh. Tekh. Fiz. **25** (10), 69 (1999) [Tech. Phys. Lett. **25**, 410 (1999)].
2. *The Problems of Modern Crystallography*, Ed. by B. K. Vaĭnshteĭn and A. A. Chernov (Nauka, Moscow, 1975), pp. 150–171.
3. V. N. Shrednik, Pis'ma Zh. Tekh. Fiz. **24** (11), 34 (1998) [Tech. Phys. Lett. **24**, 427 (1998)].
4. V. N. Shrednik and G. A. Odishariya, Izv. Akad. Nauk SSSR, Ser. Fiz. **33** (3), 536 (1969).
5. V. N. Shrednik, in *All-Union Conference on Field Ion Microscopy* (Kharkov. Fiz.-Tekh. Inst., Kharkov, 1976), p. 9.
6. O. L. Golubev, G. A. Odishariya, and V. N. Shrednik, Izv. Akad. Nauk SSSR, Ser. Fiz. **35** (2), 345 (1971).
7. B. G. Livshits, *Physical Properties of Metals and Alloys* (Mashgiz, Moscow, 1956).
8. V. N. Shrednik, Fiz. Tverd. Tela (Leningrad) **1** (7), 1134 (1959) [Sov. Phys. Solid State **1**, 1037 (1960)].
9. E. L. Murphy and R. H. Good, Phys. Rev. **102** (6), 1464 (1956).
10. R. H. Fowler and L. Nordheim, Proc. R. Soc. London, Ser. A **119** (781), 173 (1928).
11. *Unheated Cathodes*, Ed. by M. I. Elinson (Sov. Radio, Moscow, 1974), pp. 165–173.
12. L. N. Dobretsov and M. V. Gomoyunova, *Emission Electronics* (Nauka, Moscow, 1966).

Translated by M. Astrov

Technical Physics, Vol. 45, No. 12, 2000, pp. 1581–1588. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 73–80.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Detection of Positive and Negative Ions in a Wire Ionization Counter Filled with Atmospheric Air

V. K. Lyapidevskiĭ, A. B. Milent'ev, A. I. Mis'kevich, and V. N. Sysoev

Open Joint-Stock Company, Research Institute of Steel, Moscow, 127411 Russia

Received March 3, 1999

Abstract—The physical processes occurring during the detection of individual ions in a wire counter filled with atmospheric air are studied, and the amplitude distribution of electric pulses is measured. It is shown that a positive ion is detected due to electron emission from the surface of a negatively charged wire, whereas a negative ion is detected due to electron detachment near a positively charged wire. The secondary processes accompanying the neutralization of positive ions at a negatively charged external electrode and negatively charged wire are considered. It is established that, in contrast to standard proportional counters, these secondary processes result in an increase in the number of electric pulses rather than in an increase in the electric pulse amplitude. The wire counter allows the detection of both ions produced inside the detector and ions injected into the detector from the ambient air. Possible applications of air-filled wire detectors are described. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

As was shown in [1], ionizing particles in air can be detected with an open Geiger–Muller (G–M) counter located away from their paths. A cylindrical gas-discharge counter with a negatively charged wire was described in [2, 3]. The counter was exposed to α-, β-, and γ-rays. It was found that ionizing particles are detected due to both the electron emission caused by positive ions that are neutralized at the wire surface and the subsequent development of a self-sustained discharge. The characteristic feature of this operating mode is that the electric pulse amplitude and shape are independent of the type of particles and the voltage applied.

In both cases, the counters operated in the self-sustained gas discharge mode. In contrast to conventional counters filled with electropositive gases, the counter detects both the primary ions and the ions produced in the gas discharge. The separate detection of the ions generated due to ionizing radiation and the ions produced in the gas discharge has not yet been realized. The number of ions generated in a gas discharge may exceed the number of primary ions by several orders of magnitude. Under these conditions, the counter largely detects the gas-discharge ions.

In this study, we used an ionization chamber with gas amplification to detect the ions produced by ionizing radiation inside the detector or entering the detector from the ambient air [4]. The basic condition for implementing this regime consists in limiting the gas amplification so that the average number N of the ions produced in the gas discharge satisfies the inequality $\eta_I N < 1$, where η_I is the probability of the generation of a free electron when the ion approaches the wire.

We investigated wire cylindrical detectors with external cylinder diameters of 8.5, 18.5, 50, 150, and 160 mm; wire diameters of 20, 30, and 100 μm; and lengths of 18, 19, and 20 cm. Each detector was open to the atmosphere through external cylinder openings. We also explored a detector in which the cylinder was replaced with a 10-mm-diameter torus and a 20-μm-thick wire.

The main detector characteristics appear to be independent of the detector size. For a detector with a negatively charged wire, the range of operating voltages broadened as the cylinder diameter increased and the wire diameter decreased, whereas for a detector with a positively charged wire, the inverse dependence was observed.

A block diagram of the experiment is shown in Fig. 1. High voltage produced by a power supply was applied directly to the detector wire. The voltage was negative when detecting positive ions, and positive when detecting negative ions. The air inside the detector was ionized by cosmic rays, environmental radioactivity, and the radioactivity of structural materials (background ionization). The air outside the detector was ionized by a ^{238}Pu α-source positioned such that the ionizing particles could not directly enter the detector. In this case, the generated ions entered the detector with the air flow produced by the pump. An electrostatic filter was used to select ions with the desired mass. Light flashes were detected with a photomulti-

1063-7842/00/4512-1581$20.00 © 2000 MAIK "Nauka/Interperiodica"

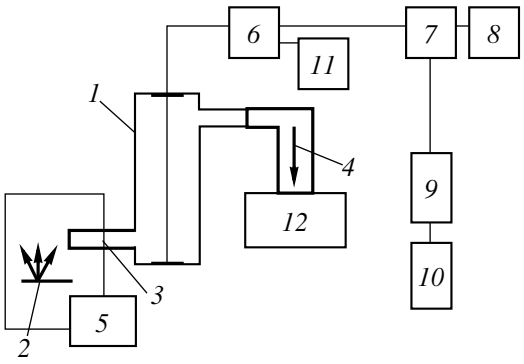


Fig. 1. Block diagram of the experiment: (*1*) gas-discharge counter; (*2*) α-particle source; (*3*) electrostatic filter; (*4*) direction of the air flow; (*5*) power supply; (*6*) preamplifier; (*7*) amplifier; (*8*) spectrometer; (*9*) discriminator; (*10*) electronic counter; (*11*) high-voltage source; and (*12*) pump.

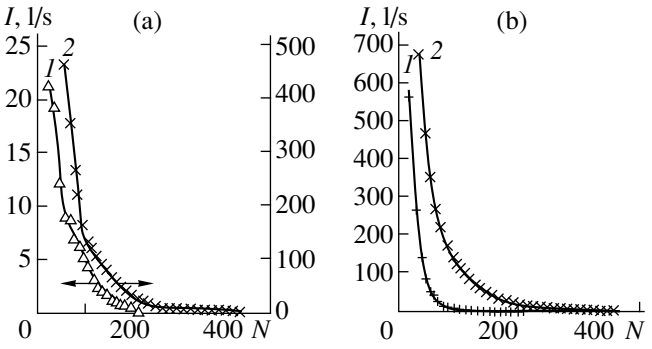


Fig. 2. (a) Amplitude distribution of electric pulses in a detector with a positively charged 100-μm-diameter wire and 8.5-mm-diameter external cylinder at a voltage of 3265 V: (*1*) ions produced in the detector and (*2*) external ions from an α-particle source; (b) external ions from an α-particle source at voltages of (*1*) 3255 and (*2*) 3265 V.

plier located directly opposite the opening in the detector case (not shown in Fig. 1).

1. DETECTION OF NEGATIVE IONS

The following characteristic regions in the air-filled detector were found [5–7]. In region *1* (adjacent to the wire), the electric field is strong (above 24.3 kV/cm under normal conditions). If a high-energy charged particle traverses this region, electrons produced due to air ionization are not thermalized and give rise to electron-impact ionization with a relevant gas amplification. In region *2* (near the external electrode), negative ions are produced due to electron attachment. The intermediate region *3* is characterized by partial electron attachment and generation of negative ions; a fraction of the electrons are not thermalized and drift from region *3* to region *1*.

When the potential difference U between the electrodes is high enough, region *4* appears near the wire. If a negative ion enters this region, an electron may be detached from the ion and an electron–photon avalanche occurs. According to [7], in air under normal conditions, this process can occur in the electric field $E > 76$ kV/cm. The produced free electrons give rise to electron–photon avalanches, which, after gas amplification, yield an electric pulse.

Thus, any negative ion produced inside the chamber can be detected after reaching region *4*. In this case, the amplitude distributions of the electric and light pulses differ significantly from those recorded when high-energy charged particles travel directly through the region near the wire where electron-impact ionization occurs [5, 6].

1.1. Amplitude Distribution of Electric and Light Pulses

The electric-pulse amplitude distribution measured for individual ions is shown in Fig. 2. As the ionizing particle traverses the main volume of the detector, the produced electrons are thermalized and attached to oxygen molecules. Negative ions are distributed in space and drift toward the positively charged wire. At a sufficiently short distance from the wire, where the electric field is high, the negative ion decays and produces an electron. This electron gives rise to the electron–photon avalanche near the wire. The number of light pulses detected by the photomultiplier far exceeds that of electric pulses. The amplitude distribution of light pulses is similar to the distribution shown in Fig. 2.

As the distance that the electron travels from the point of its origin to the wire increases, the avalanche intensity and the amplitudes of the produced electric and light pulses grow. At the same time, the probability of electron detachment from a negative ion near the wire increases due to the increase in the electric field; therefore, the number of small-amplitude electric and light pulses grows compared to the number of large-amplitude pulses. Such distributions of electric and light pulses (Fig. 2) are characteristic of the regime in which individual negative ions are detected.

It is important that the amplitude distributions obtained for ions produced inside the detector (curve *1*) and those produced outside the detector (curve *2*) are similar to each other. This fact shows that, in both cases, the counter detects primary ions rather than secondary ones, which appear in a gas discharge and yield electric-pulse distributions (Fig. 3) that are distinctly different from the distributions presented in Fig. 2.

1.2. Physical Processes Occurring inside the Detector

The rise time of the detected electric pulses is less than 10^{-8} s. This indicates that the primary electron–photon avalanche develops in a short time. The photons produced in the avalanche are partially absorbed in the

detector and generate positive ions and electrons throughout the detector volume. A fraction of photons reach the cathode and cause photoelectron emission. The number of photons that have passed through the detector (this number is measured with a photomultiplier through the openings in the cathode) far exceeds the number of electric pulses.

Unlike conventional counters filled with electropositive gases, in the detector under study, photoelectron emission plays no part in amplifying the primary electron–photon avalanche. The electrons produced inside the detector are attached to oxygen molecules, thus forming negative ions whose mobility is several orders of magnitude lower than that of electrons. Therefore, negative ions cannot take part in the avalanche amplification, because they reach the wire when the primary avalanche is already completed.

In conventional counters, the primary avalanche at the wire is localized by the trace of the high-energy charged particle, whereas the secondary electron avalanches are initiated by photoelectrons that are emitted from any point at the cathode surface. Therefore, after the detection process is completed, the gas discharge spreads over the entire wire. As the potential difference grows, the gas amplification and, hence, the space charge around the wire increase. The increase in the space charge limits the gas amplification, which results in switching the detector to the Geiger-counter mode. In contrast to conventional counters, in the detector under study, the avalanche is localized at a small section of the wire. This is clearly seen in the photographs of the discharge obtained in a condensation chamber [8].

As the potential difference U increases, the electric field also increases and the region where the electron is detached from a negative ion expands. This leads to an increase in the number and amplitudes of the detected electric and light pulses (see Fig. 2b, curve *2*, and Fig. 4).

The primary processes inside the detector end with the generation of an electron–photon avalanche near the wire: the electrons and ions of the primary avalanche reach the electrodes and the photons are partially absorbed in the detector, producing spatially distributed positive ions and electrons, which partially cross the detector and reach the cathode. The secondary processes inside the detector are governed by the photoeffect at the cathode and in the gas and by the electron emission accompanying the neutralization of positive ions at the cathode.

1.2.1. Secondary processes due to the photoeffect. The fundamental difference of the air detector from conventional counters is that the secondary electrons cannot participate in amplifying the primary avalanche. They are attached to oxygen molecules, thus forming negative ions that are distributed over the detector working space. When approaching the wire and entering a strong electric field, a negative ion initiates a new electron–photon avalanche with the probability η_I. The produced electric pulse is delayed with respect to the primary one. The maximum delay time is equal to the time τ during which a negative ion drifts from the cathode to the wire.

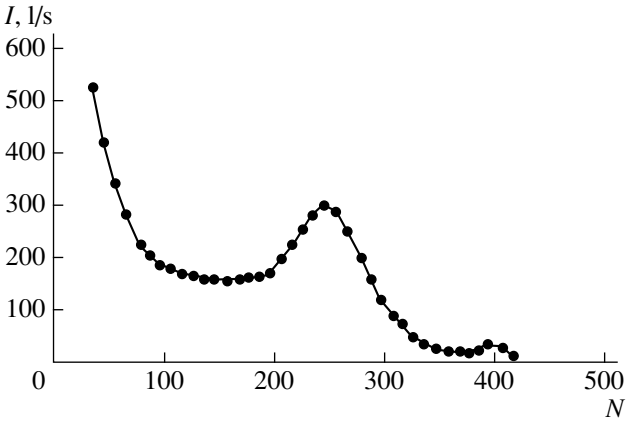


Fig. 3. Amplitude distribution of electric pulses in the regime of detecting ions produced in a gas discharge in a detector with a 50-mm-diameter external cylinder and positively charged 30-μm-diameter wire.

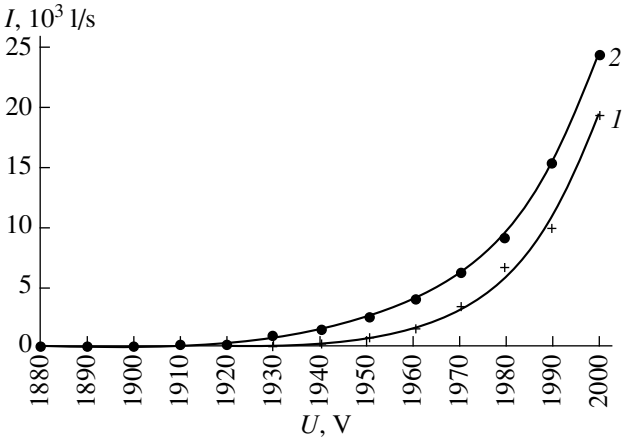


Fig. 4. The number of pulses measured per second in the detector with a 8.5-mm-diameter external cylinder and positively charged 20-μm-diameter wire as a function of the wire voltage: (*1*) background ionization and (*2*) ionization due to an external source.

Let us estimate the number of the detected electric pulses with allowance for the photoeffect. Assuming that each negative ion approaching the wire generates a free electron with the probability η_e and, hence, gives rise to the electron–photon avalanche, which yields on average N^- spatially distributed negative ions, we obtain the total number of electron–photon avalanches and, accordingly, the electric pulse from one detected ion

$$N_e = 1 + \eta_e N^- + (\eta_e N^-)^2 + (\eta_e N^-)^3 + \ldots. \qquad (1)$$

If the common ratio in series (1) is less than unity, $\eta_e N^- < 1$, then we obtain

$$N_e = \frac{1}{1 - \eta_e N^-}. \quad (2)$$

Let us assume that, without taking into account secondary processes, the number of electric pulses per second is N_{e0}. Then, with allowance for the secondary ions produced in the gas discharge, the total number of electric pulses per second is

$$N_{e\text{comp}} = \frac{N_{e0}}{1 - \eta_e N^-}. \quad (3)$$

We will refer to the ratio $N_y = N_{e\text{comp}}/N_{e0}$ as the ion multiplication factor (IMF). If the gas amplification factor in the primary electron–photon avalanche is small, then the denominator in Eq. (3) is close to one and the IMF is equal to one. The IMF increases with increasing N^-. Ion multiplication is a new phenomenon, which has not been observed in conventional counters with gas amplification.

Thus, in the standard proportional counter, the secondary processes result in an increase in the pulse amplitude. In contrast, in the detector under study, the pulse amplitude is governed by amplification in the primary avalanche, whereas the secondary processes result in an increase in the number of counts.

1.2.2. Secondary processes accompanying electron emission during the neutralization of positive ions at the cathode. The production of a false pulse in standard G–M counters is known to require only that positive discharge ions approaching the cathode induce emission of even one electron; i.e., it is sufficient that the condition $N_I^+ \gamma_I \geq 1$ be satisfied. According to the data obtained with the gas-discharge counters [9], the emission coefficient is on the order of 10^{-4}. In the air detector, the generation of an electric pulse requires that the condition $N_I \gamma_I \eta_e^- \geq 1$ be satisfied, where $\eta_e^- = 10^{-4}$–10^{-5} (see Section 1.3); i.e., the number of positive ions reaching the cathode should exceed 10^8. Therefore, for the air detector in which the gas amplification factor (and, hence, the number of positive ions) is several orders of magnitude less than the above value, the electric pulses caused by positive ions do not appear immediately after the detection of a primary ion.

1.3. Probability of Negative-Ion Detection

The probability of negative-ion detection η_I^- may be estimated by comparing the counts recorded in two experiments with the same scheme of irradiation by ^{238}Pu α-particles. In the first experiment, the detector was filled with argon, and in the second one, it was filled with air.

In the first experiment, the electric pulses produced by electrons from an α-particle track were detected. Their number N_α was almost equal to the number of α-particles in the working space of the detector. In the second experiment, the electric pulses caused by negative ions were detected, their number being N_e^-. In both cases, the ratio N_e^-/N_α appeared to be larger than 10.

Assuming that one ^{238}Pu α-particle creates 1.5×10^5 electrons on average inside the detector and that the number of negative ions formed N_I^- is equal to the number of electrons, the probability of detecting one negative ion is

$$\eta_I = \frac{N_e^-}{N_I^-} = 0.66 \times 10^{-4}.$$

Note that the probability that an electron is detached from a negative ion and, consequently, the probability η_I^- increases with increasing electric field.

2. POSITIVE-ION DETECTION

When the positive ions are neutralized, they induce electron emission from the wire surface, which is the main feature of a detector with a negatively charged wire. The electrons generate an electron–photon avalanche with subsequent amplification due to photoelectrons escaping from the wire. This results in an increase in the electric-pulse amplitude (Fig. 5b) and its strong dependence on *U*.

2.1. Amplitude Distribution of the Electric and Light Pulses

The amplitude distribution of the electric pulses is shown in Fig. 5a (curves *1* and *2* correspond to the background ionization and the additional ionization of the ambient air by an α-source, respectively). The light pulses have a similar distribution. Their number exceeds that of the electric pulses by a factor of 10^3 and more. The pronounced peak in the distribution is due to the fact that the electron–photon avalanche that originated at the wire surface decays almost at the same distance from the wire due to a decrease in the electric field. The constant shape of the amplitude distribution indicates that the pulses detected in both experiments are due to the positive ions generated in the ambient air that subsequently penetrated into the detector (Fig. 5a, curve *2*) and those produced by ionizing particles inside the detector (Fig. 5a, curve *1*). As *U* increases further, a second peak with a greater amplitude appears (Fig. 5c). The number of counts in the second peak increase with increasing *U*, whereas the relative number of counts in the first peak decreases.

At a certain value of *U*, one more peak appears; its amplitude exceeds 1 V and is independent of *U* (Fig. 6).

In character, the electric pulses in this peak are the pulses of a self-quenched and self-sustained discharge [2, 3]. The number of these pulses increases with voltage. Therefore, the number of positive ions produced in the discharge also increases and may exceed the number of primary ions by several orders of magnitude. Under these conditions, simultaneously with the large-amplitude peak, numerous pulses with smaller amplitudes appear and the total number of counts increases by several orders of magnitude (Fig. 7).

2.2. Physical Processes inside the Detector

The fundamental difference of the detector under study from conventional proportional and G–M counters is that a positive wire counter filled with electropositive gas is incapable of detecting ions and electrons separately because it is impossible to distinguish between the electrons resulting from the neutralization of positive ions at the cathode and the electrons produced in the gas. Either electrons initiate electron–photon avalanches near the wire, and the generated electric pulses do not differ from one another. In contrast, a detector with a negatively charged wire and electronegative gas is capable of separate detection. The electrons emitted from the wire surface due to the neutralization of a positive ion get into the strong field and generate an electron–photon avalanche, while the electrons produced in the gas are attached to oxygen molecules, forming negative ions, which are then neutralized at the anode.

In contrast to a detector with a positively charged wire, in which a high electric field of 70–80 kV/cm is required for a free electron to be created due to the decay of a negative ion, for the detector used in this study, a free electron appears during the neutralization of a positive ion at the wire at the expense of the ion potential energy. Therefore, for the negatively charged wire, the potential difference U starting from which the detection of electric pulses becomes possible is governed only by the gas amplification and appears to be less than that for the positively charged wire. At the same time, the dependence of the gas amplification factor on U becomes stronger. The molecule formed due to ion neutralization is in the excited state and can draw one more free electron from the wire. This electron generates an electron–photon avalanche which originates at the wire surface. Let us determine the total number of ion pairs produced in the avalanche.

We assume that the electron–photon avalanche initiated by one electron contains N ion pairs and N_{ph} photons. After reaching the wire surface, the photons create $N_{ph}\gamma_{surf}$ photoelectrons. Leaving the wire surface, these electrons generate an electron–photon avalanche containing $(N_{ph}\gamma_{surf})N$ ion pairs and $(N_{ph}\gamma_{surf})N_{ph}$ photons. The photons reach the wire and produce $(N_{ph}\gamma_{surf})N_{ph}\gamma_{surf}$ electrons, which also generate an avalanche containing $(N_{ph}\gamma_{surf})^2N$ ion pairs, and so on.

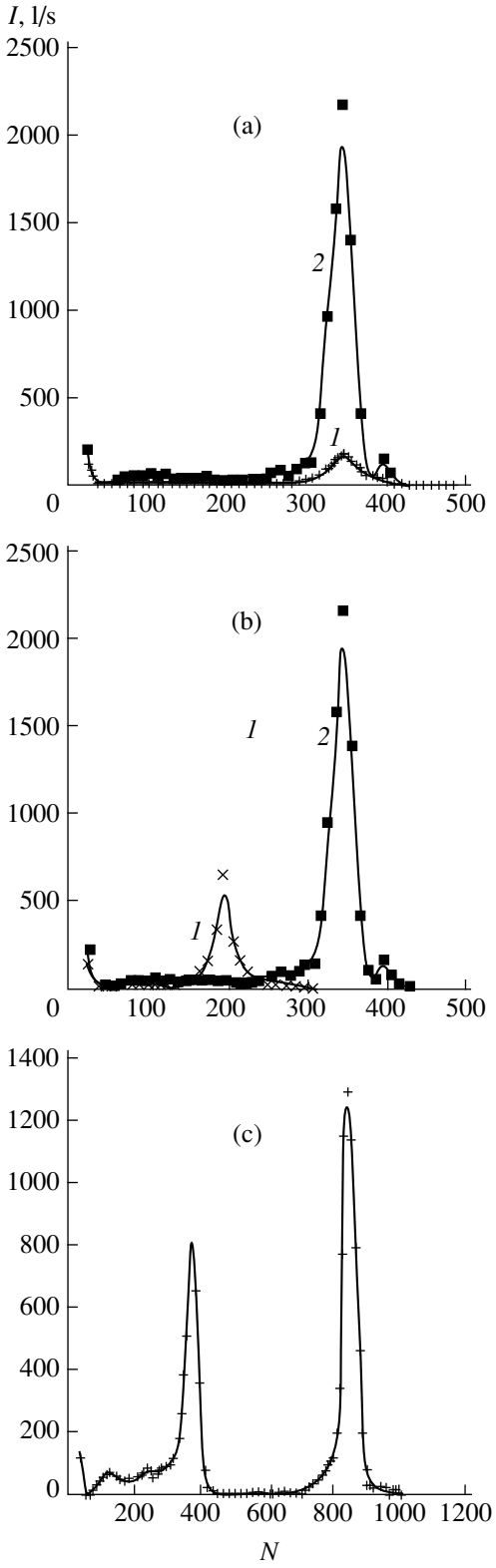


Fig. 5. Amplitude distribution of (a) electric pulses in a detector with a 18.5-mm-diameter external cylinder and negatively charged 100-μm-diameter wire at a voltage of 2290 V: (*1*) ions produced in the detector and (*2*) external ions from an α-particle source; (b) external ions from an α-particle source at voltages of (*1*) 2270, (*2*) 2290, and (c) 2310 V.

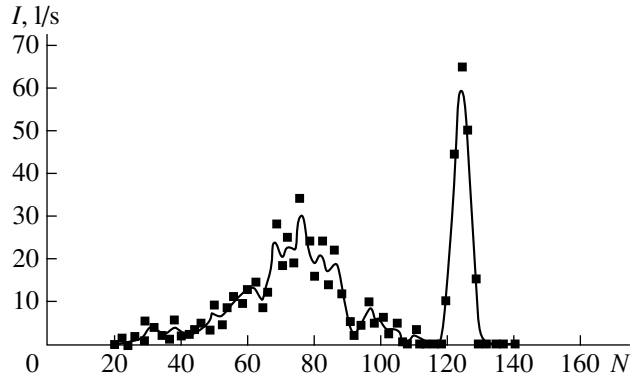


Fig. 6. Amplitude distribution of electric pulses for the intermediate operating mode of a detector with a 150-mm-diameter external cylinder and negatively charged 20-μm-diameter wire.

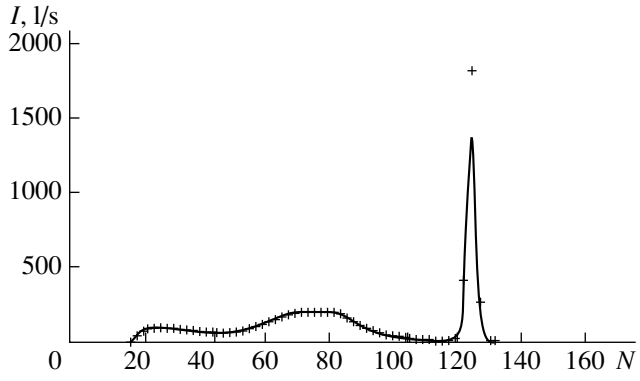


Fig. 7. Amplitude distribution of electric pulses in the regime in which a counter with a 150-mm-diameter external cylinder and negatively charged 20-μm-diameter wire mainly detects gas-discharge ions.

Summing up all successive avalanches, we find that one electron emitted from the wire produces N_p ion pairs

$$N_p = N + N(N_{ph}\gamma_{surf}) + N(N_{ph}\gamma_{surf})^2 + N(N_{ph}\gamma_{surf})^3 + \ldots. \quad (4)$$

Assuming that $N_{ph}\gamma_{surf} < 1$, instead of N ion pairs, we obtain

$$N_p = \frac{N}{1 - N_{ph}\gamma_{surf}}. \quad (5)$$

We emphasize that, unlike conventional proportional counters in which secondary electron avalanches follow one another within time intervals equal to the time during which electrons drift from the external cylinder to the wire (~10^{-6} s), in the detector under study, the electrons immediately enter a strong electric field. Therefore, the time of avalanche propagation is on the order of 10^{-9} s, which is far less than in conventional counters. The number of photons in the avalanche is very large and increases sharply with increasing voltage. In addition, as U increases, the probability that two photons are simultaneously emitted from the wire surface during the neutralization of a positive ion increases. According to Eq. (3), this results in an approximately twofold increase in the electric pulse amplitude. The appearance of the second peak in the pulse-amplitude distribution as U increases is observed experimentally (Fig. 5c). All of these processes result in a strong dependence of the number of ions on U (Fig. 8).

The strong dependence of the numbers of electrons and ions in a gas discharge on U leads to a decrease in the range of operating voltages. Under these conditions, fluctuations of the gas-amplification factor are responsible for significant variations in the counting rate. To make the detector operation more stable, it is necessary to reduce the dependence of the gas-amplification factor on U. In conventional proportional counters, this is achieved by introducing quenching admixtures into the gas, which depress the influence of the photoelectron emission from the cathode surface on the gas amplification. In the detector under study, adding admixtures is inefficient, because the photons produced in the avalanche near the wire have no time to be absorbed in the gas and cause photoelectron emission from the wire surface. Another possibility of depressing photoelectron emission is to decrease the area of the wire surface that is exposed to the photons. This is achieved by reducing both the wire diameter and the operating wire length. The operating wire length can be reduced by using a torus as the anode instead of a cylinder. The detector with a toroidal anode was shown to exhibit stable counting characteristics (Fig. 9). The duration of the detection pulse in the detector with a toroidal anode is 10^{-8} s; i.e., it is the same as in the cylindrical detector. During this time, the processes of primary-ion detection end. The positive ions created in the avalanche are neutralized at the wire, and the electrons leave the wire and are attached to air molecules. The negative ions thus formed drift towards the external electrode and are neutralized at its surface in a time τ^-.

2.2.1. Secondary processes accompanying the neutralization of positive ions at the cathode (wire). When the primary avalanche photons are absorbed in the gas, they produce photoelectrons and positive ions. The electrons are attached to oxygen molecules, and the produced negative ions drift towards the anode, where they are neutralized. The positive ions are distributed in the bulk of the detector and drift towards the wire. After reaching the wire surface with a certain probability η_I, they give rise to electron emission. We denote the maximum drift time of positive ions by τ^+. Let us find the number of the detected electric pulses, taking into account the electron emission caused by positive ions.

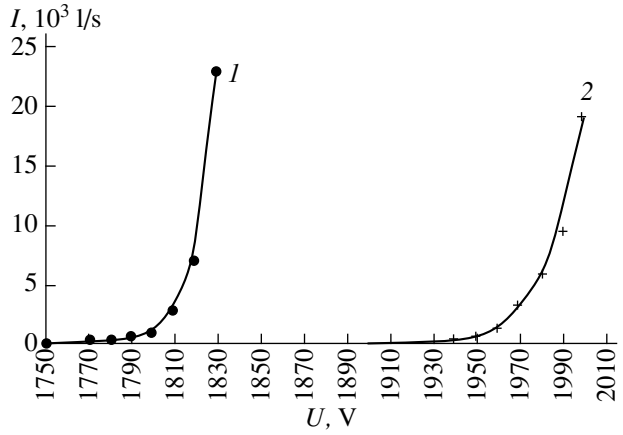


Fig. 8. The counting rate as a function of the voltage U at a 20-μm-diameter wire for background ionization in the detector with a 8.5-mm-diameter external cylinder for a (*1*) negatively and (*2*) positively charged wire.

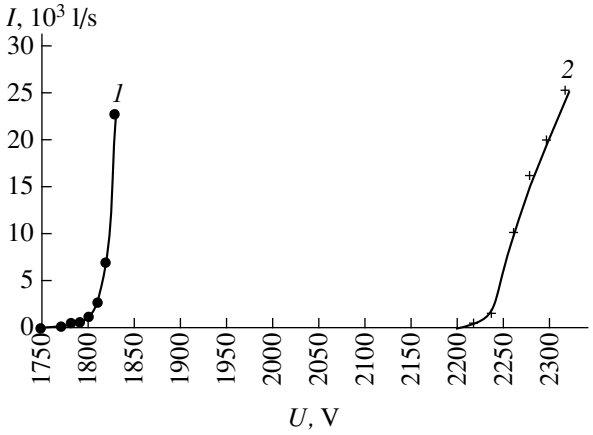


Fig. 9. The counting rate as a function of the voltage U at the negatively charged wire for (*1*) the detector with a 8.5-mm-diameter external cylinder and 20-μm-diameter wire and (*2*) the detector with a toroidal anode.

We agree that, on average, N^+ positive ions are created in the gas per one electric pulse. When approaching the wire in the time τ^+, the ions induce the emission of $\eta_I N^+$ electrons. Assuming that $\eta_I N^+ < 1$, and each electron that has reached the wire generates one electric pulse, the total number of detected electric pulses is

$$N_p = 1 + \eta_I N^+ + (\eta_I N^+)^2 + (\eta_I N^+)^3 + \ldots \quad (6)$$

i.e.,

$$N_p = 1/(1 - \eta_I N^+). \quad (7)$$

If $\eta_I N^- \ll 1$, the number of counts caused by one ion is equal to 1. The number of secondary ions in the discharge increases with increasing U; hence, the number of counts also increases. For example, for $\eta_I N^+ = 0.99$, the number of electric pulses caused by one ion is 100. For $\eta_I N^+ \geq 1$, the detector switches to the regime of a self-sustaining discharge. The number of ion pairs in the electron–photon avalanche is 10^8 [2]. In this regime, the number of secondary ions is of the same order of magnitude. Under these conditions, $\eta_I N^+ \gg 1$ and the counter primarily detects the ions generated in the gas discharge. The discharge is localized on a small wire segment; therefore, the resolving time is on the order of the duration of the detected electric pulse ($\sim 10^{-8}$ s). In this case, the resolving time is several orders of magnitude less than in G–M counters where the discharge spreads over the entire wire, causing dead time (10^{-4} s).

Due to fluctuations in the number of secondary ions, at certain values of U, we observe both self-sustained discharge pulses and electric pulses corresponding to the regime of proportional gas amplification (Fig. 6).

2.3. Probability of Positive-Ion Detection

The generation of a free electron during the neutralization of a positive ion near the wire surface results (with a probability of about 1) in the development of electron-photon avalanches and the appearance of an electric pulse. Therefore, we may assume that the probability of positive-ion detection η_I^+ is approximately equal to the probability of electron emission from the wire η_e^+.

If the IMF is equal to 1, the ion-detection probability is equal to the ratio of the number of counts to the number of ions produced in the same time interval. The ionization rate (the number of ions produced in 1 cm^3 per second) inside the detector may be assumed to be equal to the ionization rate in the ambient air. The ionization rate measured with the detector was twenty ion pairs per 1 cm^3. When the IMF is equal to 1, the number of counts is almost independent of U. Such conditions were created in a detector with a 15-cm-diameter external cathode, a working volume of 4 l, and a 20-μm-diameter wire. The average counting was 80 pulses per min. Assuming that all of the ions produced were collected by the electrodes, the detection efficiency was $\eta_I^+ = 1.6 \times 10^{-5}$.

2.4. Detection of Ions from the Ambient Air

When measuring the concentration of ions with differing mobilities, the air under investigation was drawn into the detector through an electrostatic filter (Fig. 1). The filter was a rectangular channel 90 mm long and 20 mm wide and the electrodes were separated by 10 mm. By measuring the potential difference between the filter electrodes, it is possible to analyze the mobility of the detected ions (Fig. 10). At a linear air-flow rate V [l/min], only the ions with the mobility $\mu < dV/(LE)$ will pass through the flat filter (here, L is the filter length, d is the distance between the filter electrodes, and E is the electric field strength).

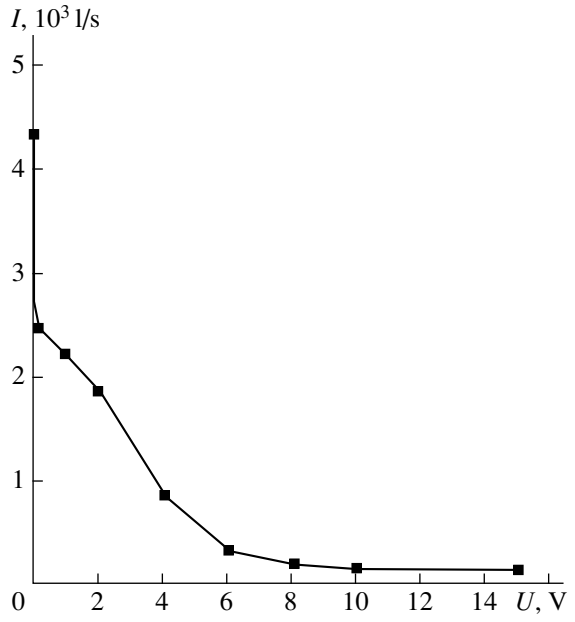


Fig. 10. The counting rate I as a function of the filter voltage U.

The air was ionized with an α-particle source located near the filter so that the direct incidence of ionizing particles onto the filter was excluded.

When the wire was positively charged, the background counting rate was 105 s^{-1}. Switching the blower on increased the counting rate to 519 s^{-1}. Bringing the α-source to the counter increased the counting rate by almost ten times (to 4353 s^{-1}). Removing the α-source restored the previous counting rate. Similar results were obtained with a negatively charged wire.

It is important to note that, in all three experiments, the shape of the amplitude distribution of the counting pulses was the same. The fact that the shape of the amplitude distribution remained unchanged shows that individual ions were detected in all cases.

3. DETECTION OF IONIZING RADIATION

The air counter detects any radiation that creates ions inside it.

At a background radiation of 20 μR/h, a detector with a 4-l working volume operating in the regime without ion multiplication detects several tens of pulses per minute, depending on the potential difference and the electrode polarity. As the dose increases (to several roentgen per minute), the counting rate increases in proportion to the dose.

REFERENCES

1. V. P. Miroshnichenko, B. U. Rodionov, and V. Yu. Chepel', Pis'ma Zh. Tekh. Fiz. **15** (12), 53 (1989) [Sov. Tech. Phys. Lett. **15**, 477 (1989)].
2. N. P. Kravchuk, V. K. Lyapidevskiĭ, and A. I. Filippov, Prib. Tekh. Éksp., No. 4, 5 (1997).
3. V. P. Kalinnikov, N. P. Kravchuk, and N. A. Kuchinskiĭ, Prib. Tekh. Éksp., No. 6, 51 (1997).
4. V. K. Lyapidevskiĭ, USSR Inventor's Certificate No. 2112994, Byull. Izobret., no. 16 (1998).
5. A. V. Ermolenko, V. K. Lyapidevskiĭ, Yu. F. Pevchev, *et al.*, Preprint No. 020-93, MIFI (Moscow Engineering Physics Institute, Moscow, 1993).
6. V. K. Lyapidevskiĭ and S. G. Pokachalov, Prib. Tekh. Éksp., No. 6, 46 (1977).
7. T. Aoyama, T. Mory, and T. Watanabe, Nucl. Instrum. Methods **212**, 261 (1983).
8. N. A. Klyachin and V. K. Lyapidevskiĭ, in *Experimental Methods and Instrumentation in Nuclear Research* (Énergoatomizdat, Moscow, 1984).
9. A. von Engel, *Ionized Gases* (Clarendon, Oxford, 1965; Fizmatlit, Moscow, 1959).

Translated by N. Goryacheva

Technical Physics, Vol. 45, No. 12, 2000, pp. 1589–1594. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 81–86.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Multipurpose Device for Nonperturbative Photoelectron Diagnostics of High-Energy H^- Ion Beams

A. S. Artemov and V. V. Antsiferov

Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141980 Russia

e-mail: artiomov@moonhe.jinr.ru

Received July 8, 1999; in final form, February 8, 2000

Abstract—A method of nonperturbative photoelectron diagnostics of a high-energy negative-ion beam proposed earlier by one of the authors is briefly discussed. A multipurpose device using the proposed method of a relativistic H^- ion beam in straight-line segments of the transportation channel is described in detail. The beam-probing photonic targets are formed by diaphragming polarized radiation of an Nd:YAG laser. The ion energy (momentum) distributions in transverse and longitudinal phase spaces are promptly determined from the spatial distributions of photoelectrons formed at the corresponding detectors after passing through a homogeneous-field 180° magnetic analyzer. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Charge-exchange methods of controlling particle flows are widely used in modern accumulating–accelerating complexes. The formation of internal charge-exchange targets in different segments of the beam transportation channel makes it possible to significantly influence both the distribution of particles in phase space and the character of their motion in an external electromagnetic field, as well as to realize a convenient method for separating high-energy beams in physical experiments [1]. The maximum efficiency of the charge-exchange method is achieved when using negative ions in the initial stage of beam transportation and employing nonperturbative diagnostics of the beam parameters. Depending on the problem to be solved, one can choose the characteristics of the internal target so that the passing beam contains the necessary number of the specific secondary particles. Both corpuscular and photonic internal targets are used as charge-exchange and beam-probing diagnostic targets (e.g., [2–11]).

Nonperturbative diagnostics imply that the beam perturbations arising over the course of measurements are negligible. The nonperturbative character and compactness of a diagnostic device are very important for the measurements in the straight-line segments of the transportation channels of intense high-energy ion beams. The use of secondary electrons as transmitters of information about the beam characteristics allows for the development of compact and multipurpose devices.

Diagnostics of a particle beam having an electron structure by a photonic target is highly promising. The monochromaticity and low divergence of laser radiation, as well as the ease of handling, make it possible to use the Doppler effect to act selectively on the quantum states of high-energy ions. Such an action is accompanied by the generation of photoionization and photodetachment electrons, which allows for beam diagnostics with a required accuracy. In 1984, one of us proposed a method of photoelectron beam diagnostics (PEBD) of negative ion beams (in particular, H^- beams; see [12, 13] for details). The method is based on the use of electrons produced due to the one-photon stripping of ions near the detachment threshold ε_t. The accuracy within which a detached electron reproduces the energy (in a. u.) and direction of the momentum of the parent ion in the laboratory frame is determined by the formula

$$\frac{\Delta E_e}{E_e} = W_e \frac{2\beta\gamma}{(\gamma-1)}\sqrt{\frac{2(\hbar\omega-\varepsilon_t)}{m_eC^2}},$$
$$\Delta\Theta_{e[\mathrm{rad}]} = W_\Theta \frac{2}{\gamma\beta}\sqrt{\frac{2(\hbar\omega-\varepsilon_t)}{m_eC^2}}, \quad (1)$$

where $E_e = E_i(m_e/M_i)$; m_e and M_i are the rest masses of an electron and ion, respectively; E_i is the ion kinetic energy; C is the speed of light; β and γ are relativistic parameters of the beam ($\gamma = 1/\sqrt{1-\beta^2}$); $\omega = \omega_0\gamma(1 - \beta\cos\eta)$; ω_0 is the photon frequency in the laboratory frame; and η is the angle between the ion and photon momentums.

In the general case, the quantities $W_e \le 1$ and $W_\Theta \le 1$ depend on ω_0, the direction of polarization of the absorbed photon, the angle η, the initial and final quantum states of the electron, and the element of the transition matrix for the ion photoionization [14]. The minimum value of $\hbar\omega - \varepsilon_t$ that can be achieved for an ion beam is mainly determined by the spread in the absolute value ($\Delta\beta$) and direction ($\Delta\Theta_i$) of the ion velocity.

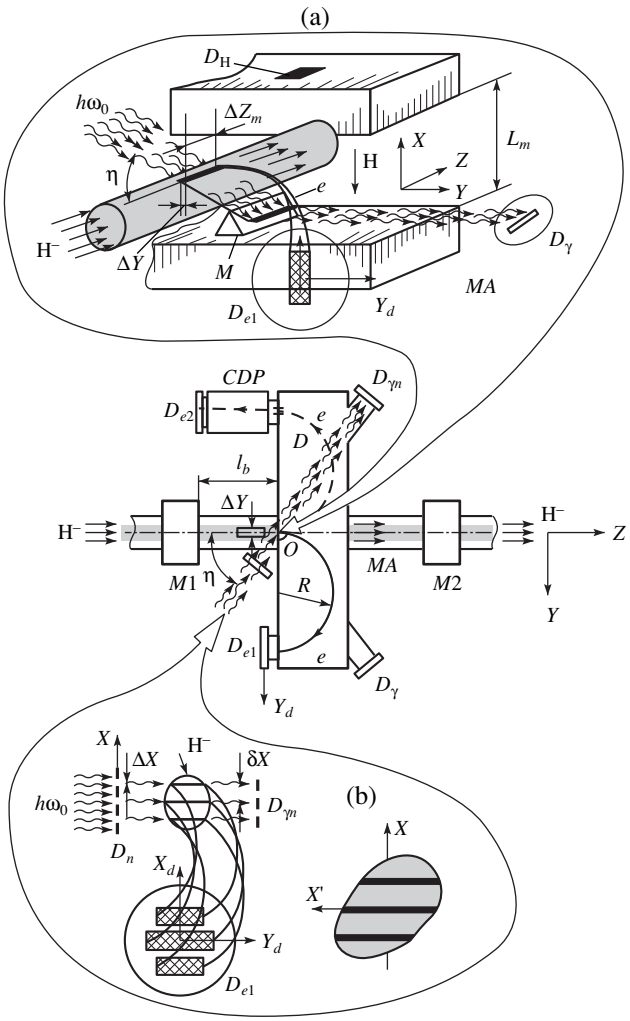


Fig. 1. A schematic diagram of the multipurpose device for nonperturbative photoelectron diagnostics of a relativistic H^- ion beam.

This difference is chosen such that the near-threshold photodetachment cross section $\sigma(\omega) \propto (\hbar\omega - \varepsilon_t)^q$ (q = 3/2 and 1/2 for the electrons in the *s* and *p* quantum states, respectively) [15] is the same for all ions:

$$(\hbar\omega - \varepsilon_t)_{\min} \gg q\gamma\hbar\omega_0\{|\Delta\beta[\beta(1-\beta\cos\eta)\gamma^2 - \cos\eta]| + \beta\sin\eta|\Delta\Theta_i|\}. \quad (2)$$

The aim of this work is to describe a multipurpose device using nonperturbative PEBD of relativistic H^- ions in straight-line segments of the transportation channel.

THE PRINCIPLE OF OPERATION AND THE SCHEME OF THE DIAGNOSTIC DEVICE

Figure 1 is a schematic diagram of the diagnostic device. A static magnetic element with a homogeneous magnetic field (*MA*) separates the photodetachment electrons from an ion beam and ensures the analysis of information transmitted by these electrons. Correcting magnets *M*1 and *M*2 completely compensate minor spatial and angular perturbations of the ion beam in front of and behind the static magnetic element. The magnet coils and the common current supply are connected in series, which ensures coordinated variations of the analyzing and correcting fields. The interpolar distance L_m is large enough to ensure both the free passage of the beam in the vacuum chamber between the magnet poles and the detection of all the electrons born in the photonic targets.

The energy spectrum and longitudinal emittance of the ion beam are varied according to the well-known layout typical of magnetic analyzers (Fig. 1a), in which a photonic target strip *O* (in the *XZ* plane) formed by laser radiation is used instead of a slit diaphragm. The target has the required spatial localization ΔY along the *Y*-axis. The energy and momentum ion spectra are determined by analyzing the spatial distribution of the electron flow density along the Y_d-axis at the detector D_{e1} with allowance for the detector spatial resolution (*d*). Using the diaphragm *D* and changing the sign and strength of the magnetic field allow us to efficiently select electrons with momentums needed for the phase analysis. By preliminarily adjusting the diaphragming system D_1 and the slit detector of photons D_γ to the analyzer, we can monitor the power and angular orientation of the target using the detector signal generated by the radiation transmitted through the interaction region. For this purpose, radiation is diverted to the *YZ* plane by a wedge-shaped mirror *M* located inside the vacuum chamber at one of the magnetic poles of the analyzer. The longitudinal emittance of the ion beam is determined by analyzing the spatial distributions of electrons arriving at the detector D_{e2} after a phase sweeping device (*CDP*) (the technique used is similar to that proposed by [16]). The main element of the phase sweeping device is a cylindrical cavity (loaded with disks), in which circularly polarized EH_{11n} electromagnetic oscillations are excited. If the polarization plane of oscillations in the phase sweeping device rotates at a frequency equal to the operating frequency of the accelerator RF field, the distribution of the electron flow density over the detector ring determines the phase spectrum of the beam ions. The degree to which the phase structure of the ion flow corresponds to that of the flow of photodetachment electrons ($\delta\Psi$) depends mainly on both the projection ΔZ_m of the photonic-target area from which the electrons arrive at the detector after diaphragming, and on the difference in the lengths of the electron trajectories inside the magnetic analyzer due to their angular spread.

Fig. 1b presents a schematic diagram of measurements of the characteristics of an H^- beam in transverse phase space. Diaphragming slits D_n form one movable (along *X*-axis) photonic target strip or several stationary

strip targets in the planes parallel to the *YZ* plane with the necessary spatial localization ΔX. The admissible intertarget distance δX is determined by the condition that the distributions of electrons arriving from the targets should not overlap at the detector D_{e1}. Such an arrangement of the targets minimizes the mutual influence of the angle and energy spectra of electrons on their distribution along the X_d-axis in the focal plane of the analyzer. Numerical simulations with allowance for analyzer geometry, adjustment accuracy, and edge fields by the optimized computer code Transport (Turtle) [12], showed that the electron distribution along the X_d-axis at the detector D_{e1} is given by the expression

$$X_d = AX + BX', \tag{3}$$

where the parameters *A* and *B* are determined only by the characteristics of the given analyzer.

The ion distribution in the $(X'X)$ phase plane, the *X*-profile of the beam, and its transverse emittance can be determined based on the electron flow distributions over the X_d-axis for the known photonic-target characteristics determining the probability of electron production and the known angular orientations and positions of the targets on the *X*-axis. An *n*-channel slit detector $D_{\gamma n}$ (where *n* is the number of the diaphragms) controls the target characteristics.

The results of numerical simulations yield the optimal parameters of the device that ensure the determination of the beam parameters with a sufficient accuracy. For example, for nonperturbative diagnostics of an H^- beam with the energy $E_{H^-} = 600$ MeV, $\Delta\beta/\beta \approx \pm 10^{-3}$ and $\Delta\Theta_i \approx \pm 10^{-3}$ rad, we developed a device with the following parameters: $R = 200$ mm, $L_m = 60$ mm, $d \approx \Delta X \approx \Delta Y \approx 0.1$ mm, and $\delta X \approx 1.5$ mm. The calculated parameters are $A = 1$, $B = 0.072$ cm/mrad, and $\Delta Z_m \leq 1$ mm (for the phase analysis with $\delta\Psi \approx \pm 1°$) or $\Delta Z_m \leq 10$ mm (for the measurements of both the distributions in the $(X'X)$ phase space and the energy (momentum) distributions [17]). The magnetic field strength of the electron analyzer ($E_e = 327$ keV) is $H = 110$ Oe. The adjustment and manufacturing inaccuracies of the analyzer lead to deviations in both the variance and the values of the parameters *A* and *B* from the calculated ones. These deviations are taken into account by calibration with the test electron beam. The distributions of ions and photodetachment electrons correspond to each other with an angular accuracy of $\Delta X' \approx 5 \times 10^{-5}$ rad and a relative energy accuracy of $\Delta E_e/E_e \approx 3 \times 10^{-4}$. This accuracy level is achieved for optimally polarized photonic targets at $\eta = 41 \pm 0.5°$. Such targets are formed by diaphragming Nd: YAG laser radiation ($\hbar\omega_0 = 1.17$ eV) with the peak intensity $I_{\gamma L} \leq 30$ kW/cm^2 in the laboratory frame. Taking into account the data from [13, 18, 19], we arrive at the conclusion that the maximum accuracy ($W_\Theta \approx 1/8$) of the transverse beam emittance measurements can be achieved for the photons polarized in the plane parallel to the intersection plane of the ion and photon beams, whereas when measuring the longitudinal emittance and ion energy (momentum) distribution, we must use the photons polarized perpendicularly to the above plane ($W_e \approx 1/5$). Note that the angle between the momentums of ions and photons is close to $\eta_b = \arccos(\beta) \approx 38°$, which corresponds to photodetachment within the minimum phase interval (the laser beam is perpendicular to the ion beam in the frame of the center of mass). The laser developed in [20] operates in the quasi-continuous regime with the duration of constant-amplitude lasing phase $\tau_\gamma \approx 250$–500 μs, the repetition rate $f_\gamma \approx 1/5$ Hz, and the angular divergence $\alpha_\gamma \approx 2 \times 10^{-4}$ rad. A special electronic appliance synchronizes the laser with the ion beam ($\tau_i \approx 100$ μs, $f_i = 1, 100$ Hz) to ensure pulse overlapping during the interval in which the radiation power is nearly constant. Optimal time matching of the laser pulses is ensured by controlled laser firing. The laser radiation can be synchronized with any source having a service timer with a frequency in the range $f \leq 2 \times 10^5$ Hz and a pulse duration multiple of 10 μs. The relaxation time of the residual thermal deformations in the garnet crystal, which result in defocusing of the laser beam and a decrease in the radiation intensity in the interaction region, determines the laser repetition rate [21]. Placing the laser in a basement under the accelerator and delivering its radiation to the diagnostic device by mirrors makes it possible to diminish the radiation action on the active element of the laser. This also simplifies the access to the laser components when the accelerator is in action. Remote mechanical rotation of the mirrors makes it possible to deliver the polarized light to one of the diaphragming systems (D_1 or D_n) to form a target of the necessary spatial orientation. A stack of wedge-shaped glass plates placed inside the cavity at the Brewster angle to the cavity axis provides almost 100% polarization of the output beam. If necessary, a half-wave quartz plate rotates the polarization of the output radiation by an angle of $\pi/2$. An analyzer, a photodiode, and an oscillograph are used to periodically monitor the direction and the degree of polarization of the output radiation. To switch between the operating regimes (see Fig. 1), we can use two alternately activated lasers producing beams with the required spatial orientations and polarizations.

The power of a single strip target $P_{\gamma L}[W]$ (in the laboratory frame) is sufficient for the measurements, provided that we can distinguish between the electrons resulting from a single-photon detachment and background electrons with the same mean energy. The latter are generated due to the interaction of H^- ions with the residual gas components (see, e.g., [22]), two-photon detachment [19], and the Lorentz dissociation of ions in the magnetic field of the dipoles.

One can neglect the influence of the residual gas if

$$\frac{P_{\gamma|L}\sigma(\omega)(1-\beta\cos\eta)}{4.8\times10^{-9}\hbar\omega_0\sin\eta\beta\Delta l}\geq 3\times10^{16}\nu_e\nu_r\sigma_{-1b}P_b l_b, \quad (4)$$

where ν_e is the ratio of the flow densities of the detached and background electrons at the detector; ν_r is the coefficient of linear broadening of the photonic-target image in the focal plane of the analyzer due to angular and momentum spread of electrons (this coefficient depends on the type of analyzer); $\sigma_{-1b} = \sigma_{-10|b} + \sigma_{-11|b}$ [cm^2] is the mean cross section of stripping H$^-$ ions by the residual gas components; P_b [torr] is the residual gas pressure; l_b [cm] is the effective distance from which the background electrons arrive at the detector; Δl [cm] is the transverse (with respect to the intersection plane of the photon and H$^-$ ion beams) size of the target (ΔX or ΔY); and $\sigma^{(1)}(\omega) = 2\times10^{-16}(\hbar\omega - \varepsilon_t)^{3/2}(\hbar\omega)^{-3}$ [cm^2] is the cross section of the one-photon stripping of H$^-$ ions near the threshold [23–25]. For $\nu_e\nu_r \approx 10^2$, $\sigma_{-1b}\approx 6\times10^{-19}$ cm^2, $l_b \approx 25$ cm (the distance between the magnet $M1$ and the analyzer MA), $\Delta l \approx 10^{-2}$ cm, and $P_b \approx 5\times10^{-8}$ torr, we obtain $P_{\gamma|L}\geq 130$ W. This means that the measurements of the profile, transverse emittance, and ion energy (momentum) distribution by strip photonic targets with the length $l_\gamma \approx 6$ mm require a laser intensity of $I_{\gamma|L}\geq 20$ kW/cm^2. For a probing target cross section of 0.1×0.5 mm, the phase analysis and the measurements of the longitudinal emittance require an intensity of $I_{\gamma|L}\geq 260$ kW/cm^2. Such an increase in the intensity can be achieved by a three- to fourfold decrease in the transverse size of the photon beam (telescoping). Although the laser angular divergence increases in this case, it is still several times smaller than $\Delta\Theta_i$.

Using the results from [19] and the given values of the parameters η, $\Delta\beta$, and $\Delta\Theta_i$, the laser intensity (in the frame related to the ion beam) at which the information regarding the beam parameters is substantially distorted by the two-photon detachment can be estimated as $L_{\gamma|\max}\approx 10^{10}$ W/cm^2. Note that when this background channel for the electron production in the diagnostic target is dominant, the accuracy of the measurements of the beam parameters decreases to $\Delta X'_{\rm cr} \approx 3\times10^{-4}$ rad and $\Delta E_e/E_e|_{\rm cr}\approx 2\times10^{-3}$.

When H$^-$ ions move in the magnetic field **H**, their characteristic lifetime in the rest frame can be estimated by the expression [26, 27]

$$\tau(F)[s]\approx(8\times10^{-14}/F)\exp(42.6/F), \quad (5)$$

where F[MV/cm] = $0.3\gamma|[\boldsymbol{\beta}\times\mathbf{H}]|$ and **H** is in kOe.

It can be easily shown that the magnetic fields of the elements $M1$, $M2$, and MA are insufficient to induce the Lorentz dissociation of the ion beam. However, they influence the stripping of high-energy H$^-$ ions via the motion-induced electric field F [28]. In particular, this influence changes the near-threshold behavior of the one-photon stripping cross section $\sigma^{(1)}(\omega, F)$, its value at $\hbar\omega = \varepsilon_t$, and the appearance of oscillations in the functional dependence on the photon energy above the threshold energy [28–31]. These effects are most pronounced when the laser polarization is parallel to the electric field (π-polarization) [30]:

$$\sigma^{(1)}(\omega, F)$$
$$\approx\begin{cases} 0.16\times10^{-17}\dfrac{F}{(\hbar\omega)^3}\exp[-68.3(\varepsilon_t-\hbar\omega)^{3/2}/F] \\ \text{at}\quad \hbar\omega<\varepsilon_t, \\ \dfrac{21.5\times10^{-17}}{(\hbar\omega)^3}\{(\hbar\omega-\varepsilon_t)^{3/2} \\ +0.015F\cos[68.3(\hbar\omega-\varepsilon_t)^{3/2}/F] \\ \text{at}\quad \hbar\omega>\varepsilon_t, \end{cases} \quad (6)$$

where $\hbar\omega$ and ε_t are in eV and F is in MV/cm.

A tabulated function yields intermediate results at $\hbar\omega = \varepsilon_t \pm 0$, which provides a smooth transition between the above dependences and more accurately describes the cross section $\sigma^{(1)}(\omega, F)$ at $\hbar\omega \sim \varepsilon_t$. The diagnostic target is π-polarized in the detection scheme shown in Fig. 1a and exhibits the corresponding projection of the polarization vector on the field **F** when operating in the regime shown in Fig. 1b. For the given values of η and **H**, expression (6) yields a 20% increase in the cross section. If we assume that the photodetachment threshold for H$^-$ ions is equal to the photon energy $\hbar\omega_F = \varepsilon_t - \Delta\varepsilon_{t|F}$, at which $\sigma^{(1)}(\omega_F, F)\approx 0.1\sigma^{(1)}(\varepsilon_t/\hbar, F)$, then we arrive at $\Delta\varepsilon_{t|F}$[eV] ≈ 0.1 (F[MV/cm])$^{3/2}$. The decrease in the threshold by this value with allowance for $W_\Theta \approx 1/8$, $W_e \approx 1/5$, and beam photodetachment in the edge field of the analyzer (see Fig. 1) worsen the experimental accuracy by $\leq$25%. This worsening can be compensated by correcting the angle η. At present, there is no information about the influence of external electric and magnetic fields on angular and energy distributions of electrons in the case of the one-photon stripping of H$^-$ ions in the laboratory frame. This is related to difficulties in the description of the process in the ion rest frame at different spatial orientations of the fields and the photon polarization vector. However, in our case, the direction of the field **F** with respect to the plane of the ion–photon interaction meets the condition of the optimal photon polarization. Thus, we may expect that the edge field of the analyzer will not deteriorate the ultimate diagnostic parameters.

The above analysis shows that the photonic target with $I_\gamma/I_{\gamma|L} = (\omega/\omega_0)^2 \approx 0.44$ allows one to reliably distinguish between background electrons and the detached electrons generated due to one-photon stripping of H$^-$ ions. Such targets also ensure that electrons

transmit the information about the ion beam with a fairly high accuracy. In experiments, the relative loss of ions passing through the target area is estimated to be $\approx 2 \times 10^{-6}$.

Precise operation of the device requires determining the spatial positions of strip photonic targets relative to the *X*- and *Y*-axes with accuracies of $\delta_x \approx \delta_y \leq 0.1$ mm and $\alpha_x \approx \alpha_y \leq 3$ mrad, screening the background magnetic fields (H_b), and high accuracy (δH) in setting the analyzer magnetic field $H_b \approx \delta H \leq 3 \times 10^{-4}$H. A sensor based on amorphous elements (similar to the Kolpits generator [32]) makes it possible to monitor the magnetic field at the required accuracy level.

Diagnostic efficiency depends on the power of the photonic targets and on the beam parameter to be measured. In particular, the estimates show that, for $\tau_i \approx 100$ μs, $I_i \approx 50$ mA, and $I_{\gamma L} \approx 20$ kW/cm^2, the nonperturbative measurements of the ion energy (momentum) distribution or distribution in the (*X*′*X*) phase plane can be carried out within a time interval equal to the beam pulse duration. In this case, a large pulsed flux of photodetachment electrons (about 10^9–10^{10} e/s from one strip photonic target) allows one to read the data in less than 10 μs. If the electron detector is activated for this amount of time in serial measurements (e.g., with a time shift multiple of 10 μs relative to the leading edge of the beam pulse), then one can trace the changes in the above parameters along the pulse over a time interval of about one minute (10 target pulses). On the other hand, for $I_{\gamma L} \approx 300$ kW/cm^2, the information about the ion distribution in the ($\Delta P_i/P_i$, Ψ) phase space or the longitudinal beam emittance can be obtained within a longer time interval of about eight minutes (~100 laser pulses). In this case, the flux of background electrons can be neglected if the detector is activated only during the lifetime of the photonic target. With allowance for the expected level of background radiation from the H^- beam, the flux can be measured using a controlled radiation-stable detector that determines the spatial distribution of the pulsed flux of relativistic electrons [33]. A substantial decrease in the time required to measure the distribution of H^- ions in the longitudinal phase space makes it necessary to develop alternative active laser elements allowing for an increase in the repetition rate and generation of radiation with the required wavelength, power, pulse duration, and beam divergence necessary for target formation.

CONCLUSION

The described multipurpose device can be readily applied to the nonperturbative diagnostics of high-intensity pulsed relativistic H^- ion beams with a current of hundreds of milliamperes (e.g., in the newly developed high-intensity neutron sources based on linear accelerators; see [4–7]). This conclusion is based on the results of theoretical considerations [34] showing the influence of the negative ion beam current on the information about the beam characteristics carried by photodetachment electrons in the homogeneous magnetic field after the probing target. In the above analysis, the beam is represented as an infinite cylinder with a homogeneous distribution of the ion flow density through the cross section.

ACKNOWLEDGMENTS

We are grateful to all of our colleagues from the Sukhumi Physicotechnical Institute and the Troitsk Institute for Nuclear Research of the Russian Academy of Sciences for help and support in the development of the diagnostic device.

REFERENCES

1. G. I. Dimov and V. G. Dudnikov, Fiz. Plazmy **4**, 692 (1978) [Sov. J. Plasma Phys. **4**, 388 (1978)].
2. R. C. Isler, Plasma Phys. Controlled Fusion **36**, 171 (1994).
3. J. J. Burgerjon, Nucl. Instrum. Methods Phys. Res. B **10/11**, 951 (1985).
4. S. Martin and C. W. Planner, in *Proceedings of the 3rd European Particle Acceleration Conference, Berlin, 1992*, Vol. 1, p. 435.
5. A. J. Jason and R. Woods, in *Proceedings of the 4th European Particle Acceleration Conference, London, 1994*, Vol. 3, p. 2684.
6. M. Mizumoto, J. Kusano, *et al.*, in *Proceedings of the 18th International Linear Acceleration Conference, Geneva, 1996*, Vol. 2, p. 592.
7. T. P. Wangler, J. Billen, *et al.*, in *Proceedings of the 18th International Linear Acceleration Conference, Geneva, 1996*, Vol. 2, p. 749.
8. W. B. Cottingame, G. P. Boicourt, *et al.*, IEEE Trans. Nucl. Sci. **NS-32**, 1871 (1985).
9. V. W. Yuan, R. C. Connolly, *et al.*, Nucl. Instrum. Methods Phys. Res. A **329**, 381 (1993).
10. Stephen L. Kramer and D. Read Moffett, IEEE Trans. Nucl. Sci. **NS-28**, 2174 (1981).
11. R. C. Connolly and D. P. Sandoval, in *Proceedings of the IEEE Particle Acceleration Conference, San Francisco, 1991*, Vol. 2, p. 1237.
12. A. S. Artiomov, N. G. Vaganov, *et al.*, in *Proceedings of the IEEE Particle Acceleration Conference, San Francisco, 1991*, Vol. 3, p. 1573.
13. A. S. Artiomov, A. A. Avidzba, and A. S. Vartazarian, Nucl. Instrum. Methods Phys. Res. A **328**, 408 (1993).
14. A. S. Artemov, Author's Abstract of Doctoral Dissertation, No. 13-97-353 (Joint Institute for Nuclear Research, Dubna, 1997).
15. S. N. Dixit and P. Lambropoulos, Phys. Rev. A **27**, 861 (1983).
16. É. S. Zlunitsyn, A. I. Zykov, G. D. Kramskoĭ, *et al.*, Vopr. At. Nauki Tekh., Ser.: Tekh. Fiz. Éksp. **2** (28), 37 (1986).
17. A S. Artemov and V. A. Kalyagin, in *Test Unit (Engineering Data no. 1.248.00.000)* (Sukhumi Fiz.-Tekh. Inst., Sukhumi, 1991).

18. A. S. Artemov and A. A. Avidzba, Preprint No. SFTI -90-17 (TsNIIatominform, Moscow, 1990).
19. A. S. Artemov, A. A. Avidzba, and A. S. Vartazaryan, Preprint No. P9-96-146, OIYaI (Joint Institute for Nuclear Research, Dubna, 1996).
20. V. V. Antsiferov, A. S. Artemov, G. F. Astrakharchik, *et al.*, *Nd:YAG Laser (Technical Specification and Operating Instruction)* (Sukhumi Fiz.-Tekh. Inst., Sukhumi, 1991).
21. A. V. Mezenov, L. N. Soms, and A. I. Stepanov, *Thermal Optics of Solid-State Lasers* (Mashinostroenie, Leningrad, 1986).
22. A. S. Artemov, in *Proceedings of the 14th Conference on Charged-Particle Acceleration, Protvino, 1995*, Vol. 2, p. 36.
23. B. H. Armstrong, Phys. Rev. **131**, 1132 (1963).
24. N. C. Bryant, K. B. Butterfield, *et al.*, in *Proceedings of the 7th International Conference on Atomic Physics, London, 1981*, p. 29.
25. K. R. Lykke, K. K. Murray, and W. C. Lineberger, Phys. Rev. A **43**, 6104 (1991).
26. G. M. Stinson, W. C. Olsen, *et al.*, Nucl. Instrum. Methods **74**, 333 (1969).
27. G. F. Drukarev and B. S. Monozon, Zh. Éksp. Teor. Fiz. **61**, 956 (1971) [Sov. Phys. JETP **34**, 509 (1971)].
28. I. I. Fabrikant, Phys. Rev. A **43**, 258 (1991).
29. J. E. Stewart, H. C. Bryant, *et al.*, Phys. Rev. A **38**, 5628 (1988).
30. M. L. Du and J. B. Delos, Phys. Rev. A **38**, 5609 (1988).
31. Bo Gao and A. F. Starace, Phys. Rev. A **42**, 5580 (1990).
32. T. Uchiyama and K. Mohri, J. Magn. Soc. Jpn. **19**, 469 (1995).
33. A. S. Artemov, Report of Joint Institute for Nuclear Research, No. P13-97-247 (Joint Institute for Nuclear Research, Dubna, 1997).
34. A. S. Artemov, Report of Joint Institute for Nuclear Research, No. P9-97-94 (Joint Institute for Nuclear Research, Dubna, 1997).

Translated by A. Chikishev

Technical Physics, Vol. 45, No. 12, 2000, pp. 1595–1603. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 87–97.

EXPERIMENTAL INSTRUMENTS AND TECHNIQUES

Effect of C_{60} Fullerene, Fullerene-Containing Soot, and Other Carbon Materials on the Sliding Edge Friction of Metals

B. M. Ginzburg*, M. V. Baidakova, O. F. Kireenko*, D. G. Tochil'nikov*, and A. A. Shepelevskiĭ***

** Institute of Problems of Mechanical Engineering Russian Academy of Sciences, St. Petersburg, 199178 Russia*

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

e-mail: ginzburg@tribol.ipme.ru

Received October 15, 1999; in final form, February 3, 2000

Abstract—A comparative study of different carbon materials (C_{60} fullerene; soot, both with and without fullerenes; graphite; and industrial carbon black) as additives to industrial lubricating oils has been carried out for copper–steel and steel–steel sliding couples. The soot containing fullerene and the powder of pure fullerene produce a noticeable improvement in the antifriction and antiwear properties of steel–steel and steel–copper couples, especially under heavy loads and pressures at the contact. The greatest improvement was observed for the steel–steel couple. Structural-mechanical studies were carried out for copper riders and it has been demonstrated by several methods that the addition of the C_{60} fullerene (pure fullerene or as a fullerene-containing soot) creates a fullerene-polymer film on the frictional surface about 1000 Å thick, which has a protective effect. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

In studies [1–5], it has been found that the addition of C_{60} fullerene to industrial lubricating oil has a positive effect on the antiwear and antifrictional properties of steel-steel and copper-steel friction units. The purpose of this work is to investigate these properties of C_{60} as an additive to industrial lubricating oils in comparison with other carbon materials under a wide range of pressures at the frictional contact and also to study the mechanism of their action with the use of various physical methods. Special emphasis is placed on the study of fullerene soot both prior to the extraction of fullerenes and after extraction. This study of fullerene soot has two purposes. First, it makes it possible to vary the C_{60} fullerene concentration in the additive while leaving its other properties intact and second, to estimate the prospects of replacing pure C_{60} fullerene additive with less expensive fullerene-containing materials.

TECHNIQUE OF TRIBOLOGICAL TESTS

Tribological tests were carried out in a 2070-CMT-1 roller machine with four different set-ups shown in Fig. 1 and described earlier [1–5] that differ in either the type of frictional contact or in the way in which the frictional unit is loaded.

The linear velocity of sliding was 1 m/s, and the duration of tests using techniques I–III was 300 s for every load value. One drop of oil (~0.05 g) was used to lubricate the apparatus before every test. The average frictional coefficient f_{av} over the test duration, the friction coefficient at the end of test f_e, volumetric wear V_w, and the linear wear intensity I_h were determined.

According to regime I, the initial contact was linear for every chosen value of the load F_n, which was maintained throughout the test. With the initial linear contact, high contact pressures (up to 200–800 MPa) typically encountered in frictional units, like the practically important tooth and worm gears, could be achieved. This test regime provided the most expeditious means of qualitative comparison of different frictional units. There are, however, a number of drawbacks in the regime, the most serious being a drastic reduction of

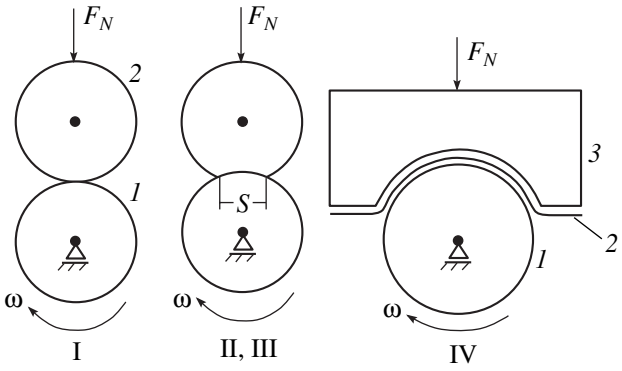


Fig. 1. Tribological test regimes: (*1*) moving roller; (*2*) fixed specimen being tested (roller in regimes I–III, foil in regime IV); (*3*) brushing with partial foil liner.

pressure in the frictional contact in the initial moments, which can cause unregistered changes of the friction and wear mechanisms [6]. Therefore, a modification in the regime I was proposed in which the load to the contact was applied in a stepwise manner after the preliminary working of the wear groove (scheme II) [6]. In contrast to regime I, the pressures applied to the contact were lower but still high enough (up to 50 MPa). The results of the tests under regime II were used to determine the dependence of frictional characteristics on load data for constant or weakly varying pressures. Regime III was used to determine the dependence of frictional parameters on pressure in the contact under constant load. In the specimens, wear grooves were first worked out and then tests under constant load were performed.

In tests using regimes I–III, the frictional contact areas were 1.5–25 mm^2 (wear groove width 0.2–2 mm). Despite the limited quantity of oil, the small frictional contact areas provided sufficient lubrication and maintenance of a semifluid frictional regime over the entire test period. Only for the highest loads was an onset of semidry or dry friction indicated by a sharp increase of the frictional torque.

Tests using regime IV lasted a total of 10–20 min and used a foil pressed between a conformal pad and a roller. One drop of oil was used in the beginning of the test, during which a series of increasing load steps was applied without stopping the frictional machine. This scheme was used only for studies of the frictional characteristics under low nominal pressures at the contact (a few MPa). One advantage of this method is the maintainance of constant pressure at the contact for a long period of time, which provides more accurate pressure measurements and the convenience of foil for subsequent structural studies. Because of the rather large contact area and longer test time, less oil was available per unit area and there was a transition from semifluid friction to semidry or dry friction with scouring at higher loads.

Thus, with the test techniques used, comparative studies were possible of the tribological characteristics of various carbon-containing additives in a wide range of loads and pressures in the contact, with the regime of varying lubrication from semifluid to dry friction. Each test was repeated 3–15 times, depending on scatter in the data, and the arithmetical mean of the measured quantities was then calculated. To determine the reproducibility of the data, a control test was carried out after 10–12 tests with a stock oil, I-40A (I-G-A as per GOST 17479-87).

TECHNIQUES OF STRUCTURAL STUDIES

X-ray diffraction patterns from the frictional surfaces were taken with a single-crystal Dmax-RC diffractometer manufactured by Rigaku Corp. using CuK_α radiation filtered with Ni. Microscopic observations were made with a CamScan 4-88DV-100 scanning electron microscope at an accelerating voltage of 15 kV or with a MIM-8 metallographic microscope. Hardness was measured by a PMT-3 microhardometer.

Analysis of thermal degradation was made in an MX-1320 mass-spectrometer. Samples were heated at a rate of 5 deg/min. In order to avoid the destruction of the thermal degradation products under electron impact in the ion source, a low-energy (18 eV) electron beam was used to ionize the thermal degradation products.

MATERIALS

The additives were mechanically mixed into the lubricating oil as a fine powder in amounts of 5 wt. % (if not specified otherwise). As a rule, the resulting mixtures were inhomogeneous suspensions. The additives used were the C_{60} fullerene; a fullerene soot produced by the Nuffman-Krätschmer method [7], designated hereinafter as (S + FL); a fullerene soot designated (S–FL), from which ~8 wt. % of the mixture of C_{60} and C_{70} fullerenes were extracted with toluene; the extracted mixture of C_{60} and C_{70} fullerenes in the ratio ~3 : 1; 6 wt. % of polystyrene engrafted onto C_{60} fullerene (6 chains of average molecular weight 5000 per fullerene molecule, so that the total content of fullerene in the oil is 0.2 wt. %); grade GS graphite powder (GOST 8295-73); and industrial soots (GOST 7885/E, grades P-514, T-900, and P-324). All additives except graphite, industrial soots, and the (S–FL) soot, are considered to be fullerene-containing additives (FCA).

THE RESULTS OF TRIBOLOGICAL STUDIES

Detailed results of the tribological studies will be published later and here we provide only the most characteristic results. The tribological characteristics were determined for the semifluid, semidry and dry friction regimes. Improvement of the antifriction and antiwear properties of the stock oil as a result of FCA additions was observed in all frictional regimes. However, under low and medium loads, the positive effect of FCA on wear in the semifluid regime differed little from the effect of the graphite additive. Under high loads and pressures, especially at the end of tests when the small initial quantity of oil was used up, and conditions arose for semidry or dry friction, the efficiency of FCA was by far better than that of graphite.

Plots of the frictional coefficient f_{av} versus the load on the friction unit obtained with regime I are shown in Fig. 2. For a steel–steel friction couple under low and medium loads the reduction of the f_{av} coefficient due to the addition of FCA amounted to ~20%; at high loads, 25–45%. For the graphite additive, f_{av} values at small loads differed insignificantly from those for FCA,

whereas at medium and heavy loads f_{av} for graphite was much higher than for FCA.

For a copper–steel friction couple, addition of FCA reduced f_{av} by 10–15%; at high loads the reduction in f_{av} reached 25–30%. Figures for different types of FCA are close and can be considered to be identical within experimental error. Yet for the copper-steel friction couple as well, f_{av} is much lower with FCA than with graphite. The advantage of FCA over graphite in regimes close to semidry is illustrated in Table 1, where values are given of the frictional coefficient f_e obtained at the end of regime I tests under heavy loads and considerable depletion of the lubricating materials: for FCA, f_e values are significantly lower. In the regime of semidry friction, considerable advantages of FCA have been determined regarding antiwear properties (Figs. 3–5), while in the semifluid regime (beginning of tests, low loads) graphite additives also give good results.

Figure 3 shows the dependence of the volumetric wear on a stepwise increasing load (regime II). The initial nominal pressures in the contact were varied over 40–50 MPa for the steel–steel frictional couple and 60–90 MPa for the steel–copper frictional couple. The most significant reduction in volumetric wear was observed for C_{60} additives (a factor of 1.5–3) and graphite (a factor of 1.3–2.5). In the case of the fullerene soots, the volumetric wear also decreased, although to a much lesser degree (by 40–80% with the (S + FL) soot and by 25–60% with (S–FL) soot). Relative values of the frictional coefficients f_{av} for various additives in tests II were about the same as in tests I under small and medium loads.

Note should be made of the additive with engrafted polystyrene; this additive is slightly more effective than others, although the fraction of added fullerene is very small (0.2 wt. %). This suggests that C_{60} acts as a catalyst of tribopolymerization with the formation of a protective film.

In Figs. 4 and 5, dependences of I_h on pressure in the contact obtained in tests by regime III with semifluid friction are shown. Only four types of additives were tested in this regime. The best antiwear properties are with the addition of C_{60}. Additions of the (S + FL) soot and graphite also produced good antiwear properties. Determinations of the energy of wear intensity as a function of contact pressure in [8, 9] provide data for similar conclusions.

Note that the advantage of FCA is more obvious for the steel–steel frictional couple and for the copper–steel frictional couple, their effect does not differ very strongly from that of carbon additives. The mechanism underlying the effect of FCA, however, is totally different and is discussed later in the text.

In tests under regime IV, the friction of a steel roller sliding over a copper foil was studied. The contact area was constant (~200 mm^2). First of all, note that the

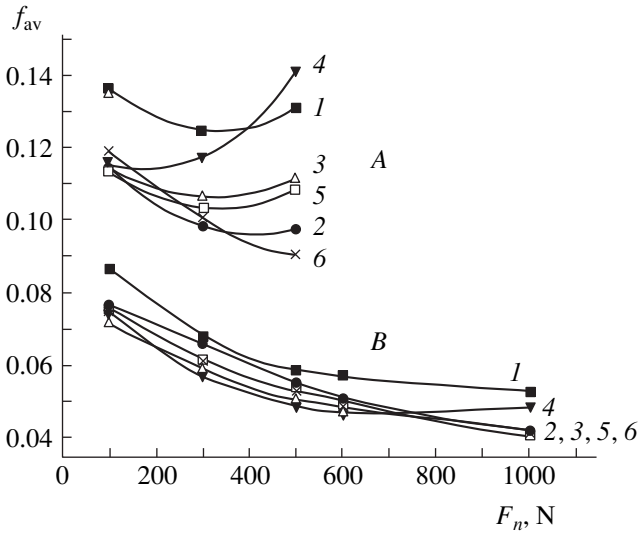


Fig. 2. Plots of the frictional coefficient f_{av} versus load on the friction unit measured in tests according to regime I. Curves of group *A* correspond to the steel–steel frictional couple and curves of group *B* correspond to copper–steel couple; (*1*) lubrication with I-40A oil; (*2*) I-40A + 5% (S + FL) fullerene soot; (*3*) I-40A + 5% (S–FL) fullerene soot; (*4*) I-40A + 5% graphite; (*5*) I-40A + 5% C_{60}; (*6*) I-40A + 6% polystyrene engrafted onto C_{60} fullerene (6 chains of average molecular weight 5000 per fullerene molecule).

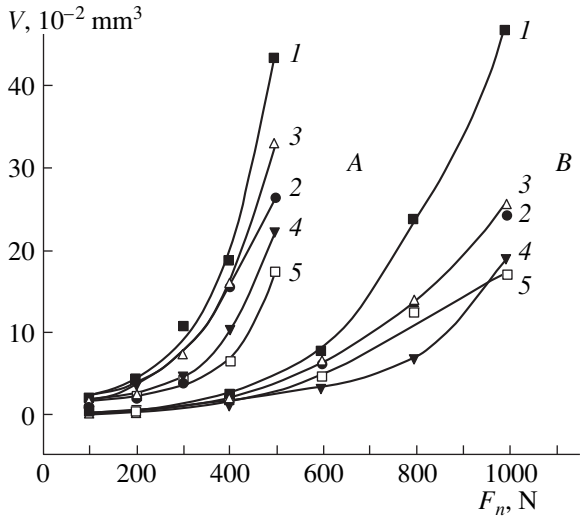


Fig. 3. Plots of volumetric wear against load on the friction unit obtained in tests according to regime II. *A*, *B*, (*1–5*) are the same as in Fig. 2.

samples from various batches of the extract (C_{60} + C_{70}) that were used as an additive were mostly or completely dissolved in I-40A oil (probably because of the formation of soluble C_{60} and C_{70} complexes) and this resulted in greater wear of the copper specimens. As shown earlier [1], C_{60} introduced by oil into a solution has little effectiveness, because to have a positive effect on the antifriction or antiwear properties of a lubricating oil, C_{60} should be in the form of a rather coarse suspension

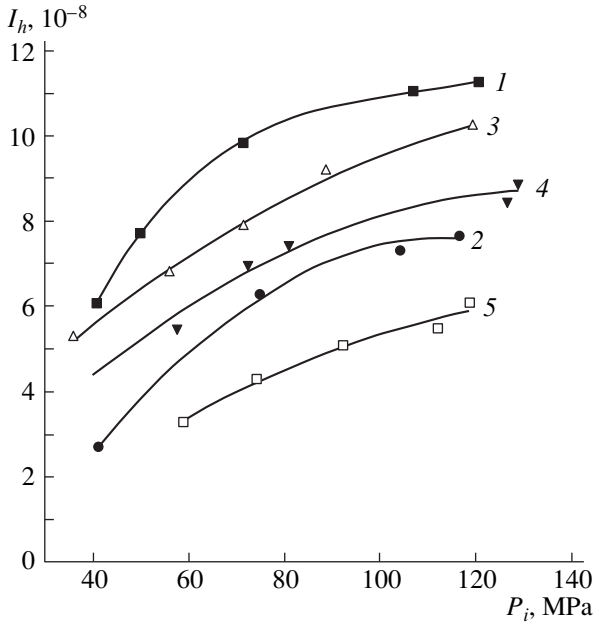


Fig. 4. Plots of the linear wear intensity against initial pressure on the tribocontact for grade St-45 steel at a load of 500 N (regime III). (*1–5*) are the same as in Fig. 2.

with macroscopic aggregations. Dispersion down to a molecular or permolecular level does not produce positive antifriction or antiwear effects. The mechanism of this effect of fullerene-containing additives is not yet known. Comparative tests of other additives to the I-40A oil (graphite powders of different origin, fullerene soot after extraction of fullerenes with toluene, industrial soots) have shown that no effect has been produced that is comparable to that of (S + FL) soot, or of C_{60} added directly into the oil in powder form (Fig. 6). For all the above materials, the rise of the frictional torque that leads to scouring begins at low loads (300–600 N; Fig. 6, curves *3–5*), whereas a frictional unit with oil containing C_{60} or (S + FL) can operate at loads in the range 800–1200 N (Fig. 6, curves *2*). In this case, the frictional coefficient stabilizes at a very low level, 0.015–0.030, which is unattainable with other carbon additives. The use of the (S + FL) soot makes the application of fullerene-containing additives to lubricating oils more simple and less expensive compared to the use of the (C_{60} + C_{70}) extract or pure C_{60}. As mentioned above, studies of various carbon materials were undertaken in order to uncover the specific effect of C_{60}, apart from it being a carbon. For this reason in particular, a number of industrial soots that contained no fullerenes were studied. However, the structure and properties of soots largely depend on their production technology and, in principle, it could be assumed that the negative effect of industrial soots is not related to the lack of fullerenes in them, but instead to properties arising from different production technologies. Therefore, although all industrial soots not containing fullerenes produced no positive tribological effects observed for C_{60} and the (S + FL) fullerene soot, we paid special attention to the (S–FL) fullerene soot, i.e., the soot produced under the same technological conditions but with most of the fullerenes extracted. The total fraction of residual fullerenes in the (S–FL) that was determined by additional prolonged extraction does not exceed 0.2–0.3 wt. %, but the determination method allows one to suppose that the fraction of residual of fullerenes is larger. In [10], it was shown that using reaction-type extraction, about 12 wt. % of soluble products can be additionally extracted from type (S–FL) soot, which represents a mixture of various fullerenes (from C_{60} up to C_{350}–C_{418}). It is still important to note that the use of the (S–FL) soot as an additive drastically changes the tribological effects. So, for example, in 25% of the cases where (S–FL) was used as an additive at loads above 400–800 N, a sudden drop in the frictional torque *M* was observed on friction diagrams; i.e., after rising to high values typical of scouring, *M* sharply dropped to values characteristic of the wear-free regime obtainable with C_{60}, peaks on the friction diagram disappeared, and the service life of the friction unit increased considerably (Fig. 6, curve *7*). In the majority of cases, the friction diagrams revealed the instability of the friction process at relatively low loads, which ultimately led to scouring (Fig. 6, curve *6*). However, it is shown below that in all cases, the effect of the (S–FL) additive is quite different from that of C_{60} or the (S + FL) soot.

RESULTS OF THE STRUCTURAL-MECHANICAL STUDIES OF THE COPPER FOIL SPECIMENS

1. Microscopic Studies

Methodology of the structural-mechanical studies was generally the same as in our previous work on additives based on pure C_{60} [2]. Of the carbon materials under study, except for the pure C_{60}, only the (S + FL) soot produced a thin film (of thickness about 1000 Å)

Table 1. Effect of different additives on the frictional coefficient in the semi-dry frictional regime

Friction couple	Load, N	I-40A	I-40A + 5% graphite	I-40A + 5% C_{60}	I-40A + 5% (S + FL)	I-40A + 5% (S–FL)	I-40A + 5% (PS–FL)
Steel–steel	500	0.147	0.187	0.097	0.106	0.114	0.095
Copper–steel	1000	0.072*	0.049	0.036	0.036	0.038	0.037

* In 40% of the tests, scouring occurred in the friction unit after 220–270 s of operation.

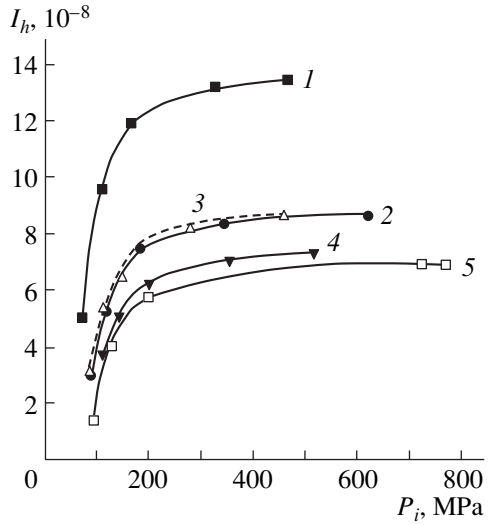


Fig. 5. Plots of the linear wear intensity against initial pressure on the tribocontact for copper at a load of 1000 N (regime III). (*1–5*) are the same as in Fig. 2.

on the surface of the copper rider observed under the microscope. In contrast to previous observations on C_{60}, in specimens subjected to friction with the (S + FL) additive alongside the bright surface covered with the film, dark areas were observed.

2. Microhardometry and X-ray Spectroscopy

The microhardness of both surface areas, bright and dark, was higher than that of the starting copper (cf. curves *1*, *3*, *4* in Fig. 7). In the same figure, data are given for the pure C_{60} additive. It is seen that the curves for C_{60} and the (S + FL) soot (Fig. 7; curves *2*, *3*) nearly coincide. This does not contradict the opinion that in both cases the properties and, consequently, the mechanism of the protective layer formation are identical. The very fact that the microhardness depends on the load means that the properties of the surface layer are different from those of the bulk material and the drooping of curves *2–4* suggests that the microhardness of the lower layers monotonically decreases away from the surface.

As in previous studies, the results of microhardness suggest two version of the structure of surface layers: it is either a thin surface film (which agrees with the scanning electron microscopy data) or a strengthened copper layer (the result of cold-hardening or the implantation into the surface of harder particles, for example, steel riders).

To choose between these two alternatives, we used, as before [2], the method of wide-angle X-ray diffraction. Details of the experimental technique are given in [2]. In Table 2, quantitative relationships of the intensities of the main X-ray reflections for specimens subjected to friction are given and are compared with the reflections for the reference specimen of pure polycrystalline copper [11], as well as for the surfaces of the starting foils on both sides (specimens I-1 and I-2) which had different textures because of differences in rolling conditions. In addition, one of the series of diffraction patterns is presented in Figs. 8 and 9, which show the general forms, positions and shapes of individual reflections.

Comparison of the results shows that all specimens tested for friction with oil containing C_{60} in a pure form (specimen F-1) or in the (S + FL) soot have similar diffractional patterns with corresponding types of the starting texture evident in conservation of the relations of the main reflections. Adding (S–FL) soot (even if the

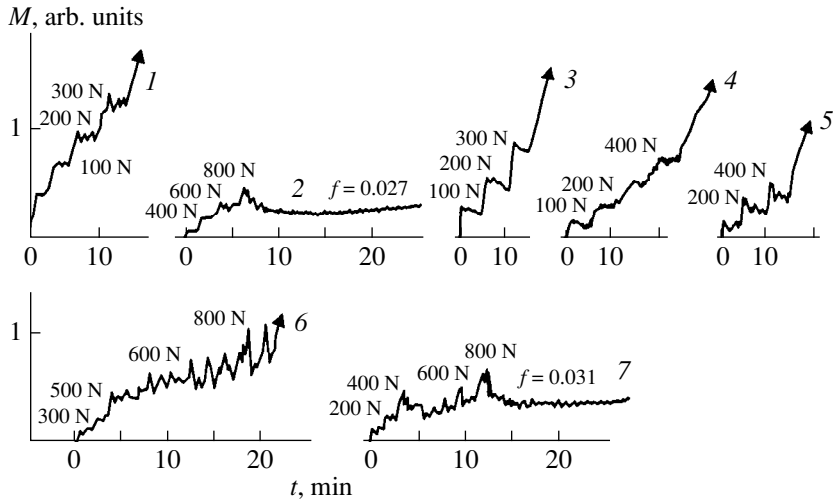


Fig. 6. Typical friction diagrams for steel roller and copper foil in regime IV (stepwise loading). (*1*) lubrication with I-40A oil; (*2*) same oil with 3–5% of C_{60} or the (S + FL) fullerene soot; (*3*) 5% of grade GS graphite; (*4*) 5% of industrial soot, grade P-324; (*5*) 5% of the mixture of C_{60} and C_{70} fullerenes in the ratio ~ 3 : 1; (*6*) 3-5% of the (S–FL) soot (typical regime ~75% of the tests); (*7*) same but with attainment of the fullerene regime, ~25% of the tests.

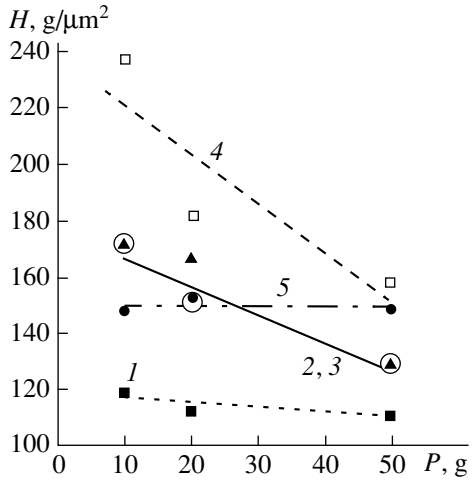


Fig. 7. Plots of microhardness H of the frictional surfaces of copper foil specimens against load on indentor P. (*1*) Initial copper surface; (*2*) frictional surface with oil containing 5% C_{60} powder (circles); (*3*) frictional surface with oil (bright regions covered with film) containing 5% of the (S + FL) soot (triangles); (*4*) same for the frictional surface regions with a dark deposit; (*5*) frictional surface with oil containing 5% of the (S–FL) soot without C_{60}.

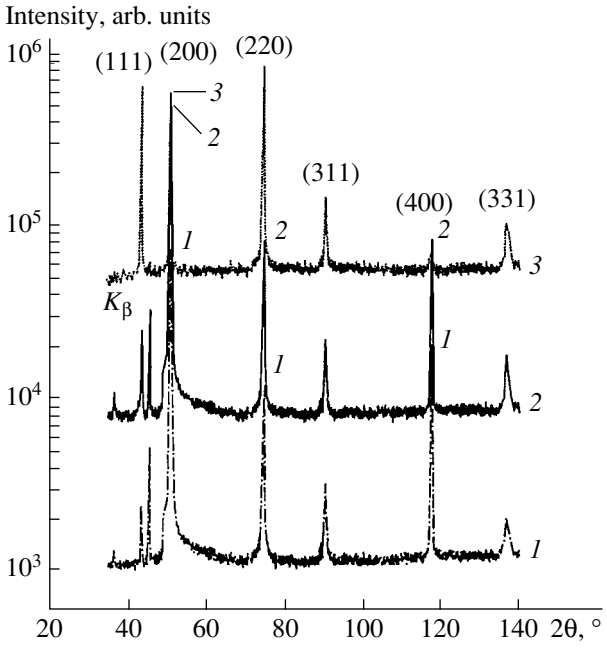


Fig. 8. Transformations of the wide-angle X-ray diffraction patterns of the copper foil as a result of friction: (*1*) initial copper foil, texture-1; (*2*) same foil after friction against steel with I-40A industrial oil + 5% C_{60}; (*3*) same foil after friction against steel with I-40A industrial oil.

tribological effect produced is quantitatively comparable to that of C_{60}) to the lubricating oil drastically changes the diffraction pattern, as seen upon comparison with patterns for the starting specimens (Table 2 and Fig. 9). Thus, enhancement of the 111 peak and even 311 peak for both specimens in this series is evidence of the tendency of the textured structure to return to the reference state in a fairly thick near-surface volume (the same is confirmed by the growth in intensity of the long range ordering peaks), though the large intensity of peaks 200 and 220 is a deviation of the observed pattern from the reference. Thus, the effect of the (S–FL) soot on the frictional microprocesses is not analogous to the effect of C_{60} and (S + FL) additives as suggested from the results of the tribological tests. Judging by the X-ray diffraction data, this mechanism is more similar to the formation of a self-organized surface structure and to the destruction of initial texture and the tendency of the intensity of X-ray diffraction peaks towards the reference pattern of polycrystalline copper [2].

Thus, if a suspension of pure C_{60} or the (S + FL) soot is present in a lubricating oil, friction, even in long tests lasting several hours, does not produce extensive changes in the type of initial surface structure. Moreover, for the main reflections, no changes have been observed in their position and width. It follows from these results that the second version of the protective layer structure is inadequate. Thus, the X-ray data supports the microscopic observations of the formation of a protective film. Also, the conclusion made in a previous work [2]—that this film not only protects the surface from the mass transfer of unlike materials in the friction couple, but also screens the initial material structure from high shearing tensions—is confirmed. Some assumptions about the structural mechanisms capable of providing energy dissipation in this very thin layer have also been made in [2].

With the addition of 5 wt. % (S + FL) to the oil, the content of C_{60} will be 0.5–0.75 wt. %. This quantity is enough to form ~4 monolayers of closely-packed C_{60} molecules on the friction contact. The thickness of this film, even if the entire mass of C_{60} present in the friction gap is involved, will be not more than 40 Å, whereas the electron microscope data provide a value of 1000 Å for the protective film thickness. Therefore, the protective film cannot consist only of the pure C_{60} fullerene.

This conclusion if confirmed by two types of experiments. Foil (with the protective film formed on it) held for many hours in toluene, which is the best solvent for C_{60}, will not cause the film to be dissolved and removed from the surface. Comparative mass-spectrometric analyses of the products of thermal degradation of tribofilms formed with and without C_{60} fullerene in the oil have also shown that the film contains hydrocarbon polymer chains with covalent bonds with the C_{60} molecules. The results of this analysis have been considered in detail in [12], and only some results are presented here.

3. Mass-Spectrometric Studies

The lubricating oil consists mainly of the products of the naphthenic and paraffin series. The results for both series are qualitatively similar; therefore, we will present only the results for the naphthenic series. In

Fig. 10a, mass-thermograms corresponding to thermal degradation of the naphthenic component of the starting I-40A oil are shown. The basic yield of the products is found in the temperature range from 100 to 300°C. In Fig. 10b, mass-thermograms obtained from a heated copper foil after friction with pure I40-A oil are shown. Only traces of the peaks corresponding to the exit of ions are observed below 250°C, and these can be attributed to traces of oil on the foil surface. The intensive exit of all ions is observed in the high-temperature region. Peaks occur at 280–290 and 324°C, as well as the two most intense peaks at 390 and 480°C.

The appearance of the high-temperature peaks not present in the mass-thermograms of the starting oil and the general shift of the main peaks to considerably higher temperatures provide evidence of the change in the state of starting oil components as a result of friction. Obviously, the products formed on the foil are no longer in a free state, but are instead bound, which requires greater activation energy for thermal degradation. The most plausible explanation of this is the formation of a tribopolymer film on the frictional surface, which agrees with data obtained by other methods. The phenomenon of tribopolymerization is widely known [13–15]. In this process, cross-linked systems are formed [15], although linear chains can also form. In addition, the radicals that form as a result of friction can react chemically with the substrate [15]; this is confirmed by the fact that the tribofilm formed on the foil cannot be mechanically separated from the substrate or by cooling of the foil with the film by liquid nitrogen (separation of the film from the foil might be expected because of the difference in the thermal expansion coefficients).

Ultimately, the thermal degradation temperature of the majority of the carbon-chain polymers is above 300°C [16, 17]. It can be assumed that the structure that formed as a result of friction is a spatial polymer hydrocarbon network and the temperatures characterizing its degradation correspond to the two groups of well-pronounced maxima around 395–400 and 475–480°C.

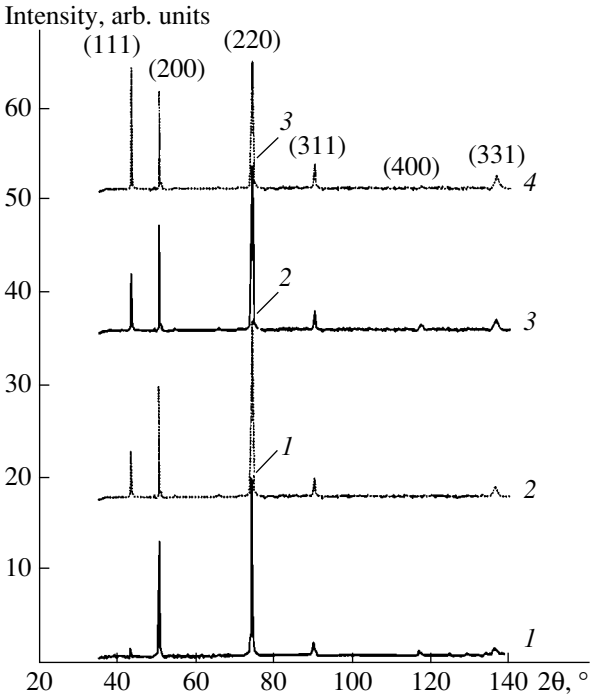


Fig. 9. Transformations of the wide-angle X-ray diffraction patterns of the copper foil as a result of friction: (*1*) initial copper foil, texture-2; (*2*) same foil after friction against steel with I-40A industrial oil + 5% C_{60}; (*3*) same foil after friction against steel with I-40A industrial oil +5% (S + FL) soot; (*4*) same foil after friction against steel with I-40A industrial oil +5% (S-FL).

In order to clarify the role of C_{60} in the process of tribopolymerization, mass-thermograms have been measured on a copper foil with oil, to which the fullerene has been added (Fig. 10c). In this instance as well, the dominant mass of the exiting product falls into the high-temperature region, where two distinct groups of peaks are seen around 320 and 430–435°C. As in the case of pure oil, we take into consideration that this is a polymer product of the tribochemical reaction, i.e., a

Table 2. Intensities of the diffraction peaks from different crystallographic planes of copper in arbitrary units (with intensity of the strongest reflection for a given specimen equal to 100)

Specimen	Texture type	Intensities of reflections, arb. units					
		111	200	220	311	400	331
Reference specimen	–	100	46	20	17	3	9
I-1	1	1	100	5			
I-2	2	9	60	100	7	2	5
F-1	2	21	67	100	10		6
(S+FL)	2	27	61	100	9		4
(S+FL)	1	4	100	3	<1	<1	<1
(S-FL)	–	100	54	75	16	2	10
(S-FL)	2	97	75	100	22	<1	9

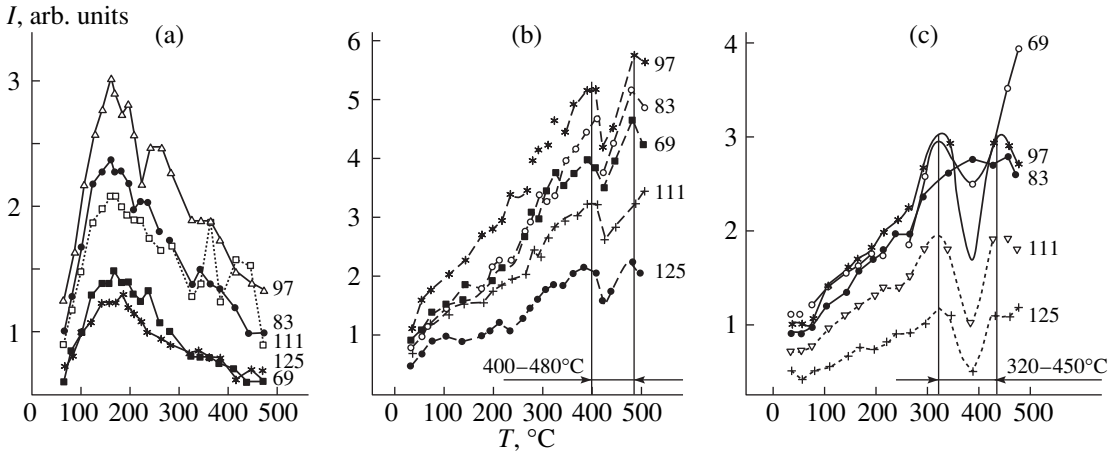
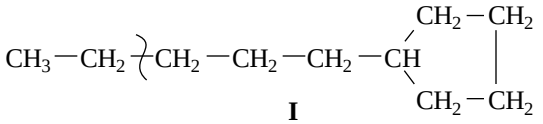


Fig. 10. Mass-thermograms for weights corresponding to thermal degradation of naphthenic compounds: (a) I-40A oil; (b) tribopolymerization products of the I-40A oil on a copper foil; (c) tribopolymerization products of the same oil + 5% C_{60} on a copper foil

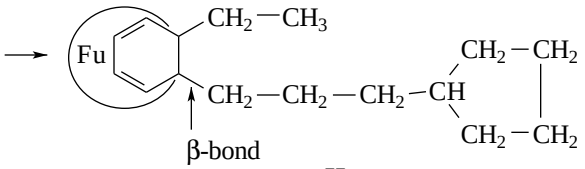
tribopolymer film formed on the surfaces of rubbing riders.

However, the positions of characteristic high-temperature maxima have shifted from the regions of 400 and 480°C (Fig. 10b) to the regions of 320 and 435°C, respectively. Lowering of the thermal stability by such a large value (~60°C on average) can be explained by the emergence in the structure of the formed product of a weaker bond. Based on the known correlation between the energy of the C–C bond and the temperature that corresponds to the loss of one-half the weight during thermal degradation [18], the drop of the thermal stability by 50–60°C corresponds to a reduction of the bond energy by ~20–21 kJ/mol. Since it is the presence of C_{60} fullerene in the oil that leads to the tribopolymer structure with the weakened C–C bonds, the cause should be related to it. It is known that the C_{60} fullerene is highly reactive [19, 20]. In particular, its molecules act as traps for the free radical in polymers [19] and stimulate formation of polymer networks, where they occupy point positions [20]. Attachment of the polymer radicals to the fullerene causes formation of a weak β-bond in the cyclodiene ring of C_{60} [19–21].

Consider this process in the instance of some molecule of the naphthenic series.



I



II

Suppose that in the friction process, mechanical degradation of this molecule occurred (the possible break in the chain is shown with a wavy line). The two radicals thus formed can react with the C_{60} fullerene to form covalent bonds (II). In the thermal degradation process during thermal analysis, the weakest β-bond will break first. As a result of the broken bond, radicals will be produced, and these fragments are detected by the mass-spectrometer. Consequently, the shift of high-temperature peaks in Fig. 10c to lower temperatures, compared with similar peaks in Fig. 10b may indicate that molecules of the C_{60} fullerene are predominantly found in points of the tribopolymer network.

DISCUSSION

The results can be used to build a model of the mechanism of antifriction and antiwear effect of FCA in the following way. In the process of tribopolymerization of the mineral oil, a coating is formed on the contacting surfaces that is a spatial tribopolymer network bonded to the substrate (by covalent bonds). This coating protects the rubbing surfaces from immediate contact, which prevents mass transfer between the rubbing riders and, being a spatial polymer network, retains the mineral oil in its cells, thus ensuring a low-wear friction regime and low frictional coefficient. This process is possible without C_{60}, but without the fullerene, formation of the tribopolymer coating requires more time and specific conditions. The presence of the highly reactive fullerene greatly accelerates the tribopolymerization process, causing formation of a tribopolymer network with predominantly C_{60} fullerene molecules in the point positions. This completely changes the frictional process, because the tribopolymer protective coating forms much sooner and the low-wear antifriction regime is established before scouring occurs. Finally, if the tribopolymer coating is damaged (for example, because of increased load) and the rubbing bodies

come into direct contact, the tribopolymerization process will be enhanced and the tribofilm restored. If the maximum possible thickness of the continuous layer of protective C_{60} film in the network is considered, the maximum volumetric fraction of the fullerene in the network will be about 4 and 40% for a 5% additive of the (S + FL) fullerene soot and the pure fullerene, respectively. This amount, especially for the C_{60} fullerene, is too high. Besides, the presence of only the fullerene–polymer network on the frictional surface cannot explain the greater microhardness of the surface layers. Therefore, modification of the C_{60} fullerene in the frictional zone with the formation of a film of modified fullerene should be considered. Numerous data on the structural transformations of the C_{60} fullerene under various high-energy impacts (high pressure, knock, irradiation, etc. [22, 23]) support this conclusion. So, under high pressure, even without additional shear stress, structures similar to amorphous carbon [22] or diamond [23] can be formed in C_{60}. Friction in the tribocontact zone, where the shear stress is high, at high enough local pressures, is accompanied by the same processes of structural modification that occur at high pressures [24]. Observations of the formation of intermediate layers consisting of ultrafine particles of the rubbing materials and experiencing unusual interactions with the environment are well known in the frictional process [24–26]. Therefore, the tribocontact constitutes a rather powerful reactor of various structural transformations and in the friction zone, fullerene modifications can be produced that have increased hardness and determine the properties of the surface films that form in the friction process. The surface film is probably a composite material of microscopic thickness (~1000 Å) that is comprised of layers based on a modified C_{60} fullerene with increased hardness and a fullerene-polymer film. For a deeper knowledge of the surface film structure, additional studies are required.

REFERENCES

1. B. M. Ginzburg, O. F. Kireenko, D. G. Tochil'nikov, *et al.*, Pis'ma Zh. Tekh. Fiz. **21** (23), 38 (1995) [Tech. Phys. Lett. **21**, 966 (1995)].
2. B. M. Ginzburg, O. F. Kireenko, M. V. Baĭdakova, *et al.*, Zh. Tekh. Fiz. **69** (11), 113 (1999) [Tech. Phys. **44**, 1367 (1999)].
3. B. M. Ginzburg, D. G. Tochil'nikov, O. F. Kireenko, *et al.*, Pis'ma Zh. Tekh. Fiz. **21** (22), 62 (1995) [Tech. Phys. Lett. **21**, 933 (1995)].
4. B. M. Ginzburg, D. G. Tochil'nikov, and V. P. Bulatov, Trenie Iznos **18** (2), 235 (1997).
5. D. G. Tochil'nikov and B. M. Ginzburg, Zh. Tekh. Fiz. **69** (6), 102 (1999) [Tech. Phys. **44**, 700 (1999)].
6. M. M. Khrushchev, *Run-In of Bearing Alloys and Trunnions* (Akad. Nauk SSSR, Moscow, 1946).
7. W. Krätschmer and D. R. Huffman, Philos. Trans. R. Soc. London, Ser. A **343** (1667), 33 (1993).
8. G. Fleisher, Schmierungstechnik **4**, 9 (1973).
9. Yu. P. Kozyrev, B. M. Ginzburg, N. D. Priemskii, *et al.*, Wear **171**, 71 (1994).
10. F. Beer, A. Gügel, K. Martin, *et al.*, J. Mater. Chem. **7** (8), 1327 (1997).
11. Joint Committee on Powder Diffraction Standards (JCPDS)—International Centre for Diffraction Data (Swarthmore, 1993), no. 4–836.
12. A. A. Shepelevskiĭ, L. A. Shibaev, B. M. Ginzburg, *et al.*, Zh. Prikl. Khim. (St. Petersburg) **72**, 1198 (1999).
13. Yu. S. Zaslavskiĭ and R. N. Zaslavskiĭ, *Friction Polymers: A New Principle of Lubrication* (TsNIITÉ Neftekhim, Moscow, 1976).
14. Yu. S. Zaslavskiĭ and R. N. Zaslavskiĭ, *Antiwear Oil Additives: Principle of Operation* (Khimiya, Moscow, 1978).
15. A. A. Polyakova, L. G. Nekhamkina, L. O. Kogan, *et al.*, *Tribopolymer-Forming Lubricants* (Nauka, Moscow, 1979), pp. 56–66.
16. H. H. Jelinek, in *Aspects of Degradation and Stabilization of Polymers* (Elsevier, Amsterdam, 1978).
17. N. Grassie and G. Scott, *Polymer Degradation and Stabilisation* (Cambridge Univ. Press, Cambridge, 1985; Mir, Moscow, 1988).
18. I. Mita, *Aspects of Degradation and Stabilization of Polymers* (Elsevier, Amsterdam, 1978), pp. 247–294.
19. L. A. Shibaev, T. A. Antonova, L. V. Vinogradova, *et al.*, Pis'ma Zh. Tekh. Fiz. **23** (18), 81 (1997) [Tech. Phys. Lett. **23**, 730 (1997)].
20. L. A. Shibaev, T. A. Antonova, L. V. Vinogradova, *et al.*, Pis'ma Zh. Tekh. Fiz. **23** (18), 87 (1997) [Tech. Phys. Lett. **23**, 732 (1997)].
21. L. A. Shibaev, T. A. Antonova, L. V. Vinogradova, *et al.*, Zh. Prikl. Khim. (St. Petersburg) **71** (5), 835 (1998).
22. H. Hirai, K. Kondo, N. Yoshizawa, *et al.*, Appl. Phys. Lett. **64** (14), 1797 (1994).
23. O. G. Epanchintsev, A. S. Zubchenko, Yu. D. Tret'yakov, *et al.*, Dokl. Akad. Nauk **340** (2), 201 (1995).
24. V. A. Zhorin, L. D. Livshits, and N. S. Enikolopyan, Dokl. Akad. Nauk SSSR **258** (1), 110 (1981).
25. D. Rigney, Trenie Iznos **8** (6), 973 (1987).
26. D. H. Buckley, *Surface Effects in Adhesion, Friction, Wear, and Lubrication* (Elsevier, Amsterdam, 1981; Mashinostroenie, Moscow, 1986).

Translated by B. Kalinin

Technical Physics, Vol. 45, No. 12, 2000, pp. 1604–1607. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 98–101.

BRIEF COMMUNICATIONS

Fractal Transition Radiation

V. N. Bolotov

Institute of Electromagnetic Research, Kharkov, 310022 Ukraine

e-mail: renic@iemr.vl.net.ua

Received October 21, 1999

Abstract—Phenomena related to fractal properties of moving emitting objects and electrodynamic structures are of considerable interest now. Electromagnetic radiation accompanying fractal transition processes is considered. © *2000 MAIK "Nauka/Interperiodica".*

INTRODUCTION

Not only fractal specific beauty has aroused considerable interest in recent years but also a wide range of new phenomena in fractal physics [1]. Nowadays, the existence of electrodynamic fractal structures is beyond question. Moreover, it has become clear that electrophysical properties of such structures are of a special nontrivial character. This refers to both stochastic and regular fractal objects.

In this paper, electrodynamic radiation problems involving fractal structures are considered. Regular fractals obtained from the Cantor function are used, which does not mean loss of generality.

Transition processes result in time or space variations of system parameters. In this paper, we consider fractal transition radiation (FTR) induced by variations of these parameters according to a fractal law. As is known [2], a particle bunch moving uniformly and rectilinearly may emit electromagnetic waves while crossing the boundary of two or more media with different permittivities. FTR results from the fact that the boundary is a fractal. Such a boundary configuration brings about the special features in the radiation spectra; namely, they are wideband and self-similar.

ELECTROMAGNETIC RADIATION ACCOMPANYING VARIATION IN THE STATE OF AN EMITTING SYSTEM BY A FRACTAL LAW

If a system of electric charges rearranges (the system parameters change) in a time T, one can argue that the system emits electromagnetic waves during this rearrangement. Let, at the initial time instant, a system of charges have a dipole moment $\mathbf{P}_1$ and after the lapse of rearrangement time T, $\mathbf{P}_2$. Let also the system rearrange (i.e., transit from one stationary state into another) by a fractal law. We use one of the well-studied fractal functions—the Cantor function $\alpha_\xi(t)$, which is related to the Cantor set [1]. Here, ξ is an interval removed at each iteration ($0 < \xi < 0.5$). The Cantor function is constant within intervals removed from the range [0, 1] and jumps at the points of the Cantor discontinuum (Fig. 1).

In this study, the interval $[0, 2\pi]$ is used. The fractal dimension of the Cantor set is $D_f = \ln 2/|\ln\xi|$. Thus, in our model, the dipole moment of the emitting system has the form

$$\mathbf{P}(\mathbf{r}, t) = \left[\mathbf{P}_1 + \alpha_\xi\left(\frac{2\pi t}{T}\right)(\mathbf{P}_2 - \mathbf{P}_1)\right]\delta(\mathbf{r}). \qquad (1)$$

The system size is assumed to be small and can be neglected. Mathematically, this fact is reflected by introducing the Dirac delta function $\delta(\mathbf{r})$. Consider radiation of frequency ω that occurs when the dipole moment obeys law (1). The Fourier component of the vector potential is defined by the formula [2]

$$\mathbf{A}_\omega = \int \mathbf{j}_\omega(\mathbf{r})\frac{\exp(i\omega/c|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}'. \qquad (2)$$

Here, $\mathbf{j}_\omega$ is the Fourier component of the current $\mathbf{j}$,

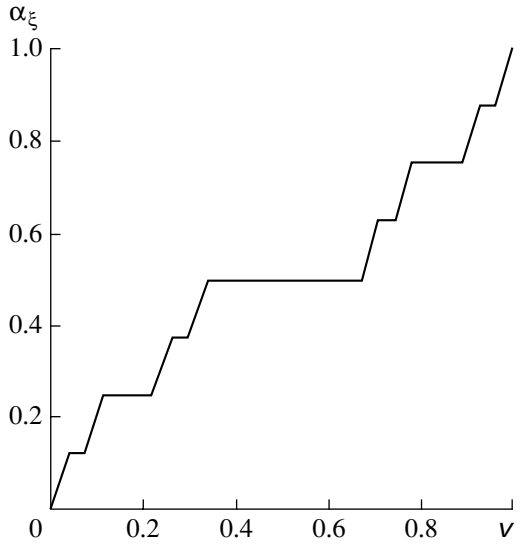


Fig. 1. Cantor function $\alpha_\xi(x)$ (ladder).

1063-7842/00/4512-1604$20.00

which is associated with a change in the dipole moment:

$$\mathbf{j}(\mathbf{r}, t) = \frac{\partial \mathbf{P}}{\partial t} = (\mathbf{P}_1 - \mathbf{P}_2)\delta(\mathbf{r})\frac{d}{dt}\alpha_\xi\left(\frac{2\pi t}{T}\right). \quad (3)$$

From this equation, one can easily find the Fourier component of the current density by going from the Fourier transform for $\mathbf{j}(\mathbf{r}, t)$,

$$\mathbf{j}_\omega(\mathbf{r}) = \frac{1}{2\pi}\int \mathbf{j}(\mathbf{r}, t)\exp(i\omega t)dt, \quad (4)$$

to the Fourier–Stieltjes transform for the Cantor function using the relation

$$\frac{d}{dt}\alpha_\xi(t)dt = d\alpha_\xi(t). \quad (5)$$

Thus, following [3],

$$\mathbf{j}_\nu(\mathbf{r}) = \frac{(\mathbf{P}_2 - \mathbf{P}_1)}{2\pi}\delta(\mathbf{r})\exp(-\pi i\nu) \times \prod_{k=1}^{\infty}\cos(\pi\nu\xi^{k-1}(1-\xi)), \quad (6)$$

where $\nu = \omega T/2\pi$.

The spectral radiation intensity per frequency interval $d\omega$ and solid angle $d\Omega$ is given by

$$dW_{\mathbf{n},\omega} = c|\mathbf{H}_\omega|^2\mathbf{r}^2 d\Omega d\omega. \quad (7)$$

At large distances r from the origin, the magnetic field $\mathbf{H}_\omega$ is expressed in simple terms through the vector potential $\mathbf{A}_\omega$:

$$\mathbf{H}_\omega(\mathbf{r}) = i\frac{\omega}{c}[\mathbf{n}, \mathbf{A}_\omega], \quad (8)$$

where $\mathbf{n}$ is the unit vector in the radiation direction.

Thus, for given $\mathbf{j}_\omega$, $\mathbf{A}_\omega$, and $\mathbf{H}_\omega$, we can obtain the radiation intensity $I(\omega)$ in the wave zone by integrating expression (7) over the solid angle:

$$I(\omega) = \frac{(\mathbf{P}_2 - \mathbf{P}_1)^2}{3\pi c^3}\omega^2 S_\xi(\omega), \quad (9)$$

$$S_\xi(\omega) = \prod_{k=1}^{\infty}\cos^2\left(\frac{\omega T}{2}\xi^{k-1}(1-\xi)\right). \quad (10)$$

Expression (9), unlike those obtained earlier, involves the structure factor $S_\xi(\omega)$ containing the product of squared cosines. Such a feature in the radiation spectrum is inherent in fractal structures [4]. The fractal dimension of a function describing the transition process has an influence on the dipole radiation. This dimension governs the spectrum shape. In the case of instantaneous rearrangement, $\omega T \ll 1$, one can easily reproduce the spectrum obtained in [5]:

$$I_0(\omega) = \frac{(\mathbf{P}_2 - \mathbf{P}_1)^2}{3\pi c^3}\omega^2. \quad (11)$$

Thus, it becomes clear that, during transition processes, fractal properties may show up only in the case of a finite rearrangement time. Figures 2–5 show the structure factors S_ξ for radiation spectra for several fractal dimensions.

One can easily obtain the functional equation for the radiation intensity from Eqs. (9) and (10):

$$I(\omega) = \frac{\cos^2[\omega T(1-\xi)/2]}{\xi^2}I(\xi\omega). \quad (12)$$

The analysis of this equation shows the self-similarity of the FTR spectra. This fact immediately follows from Figs. 2 and 3.

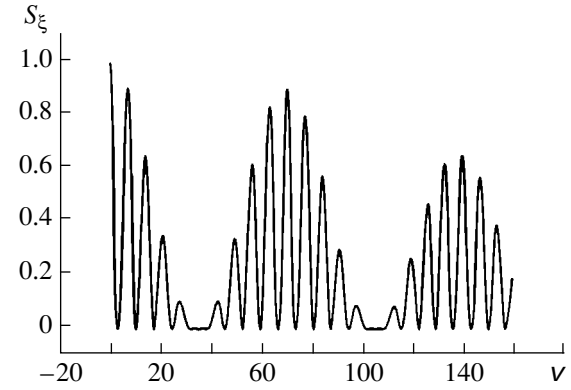


Fig. 2. $S_\xi(\nu)$ for $D_f = 0.3$.

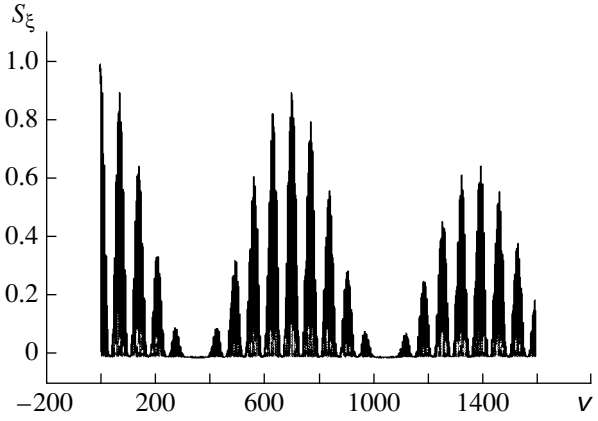


Fig. 3. $S_\xi(\nu)$ for $D_f = 0.5$.

TRANSITION RADIATION AT THE FRACTAL BOUNDARY BETWEEN TWO MEDIA

Recent studies have established that rough solid surfaces may have fractal properties (i.e., the Hausdorff–

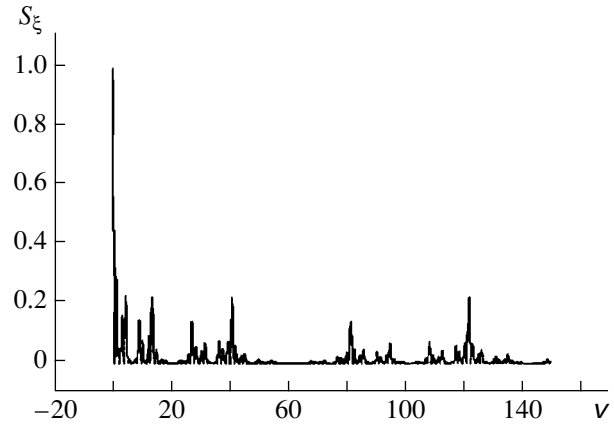


Fig. 4. $S_\xi(\nu)$ for $D_f = 0.62$.

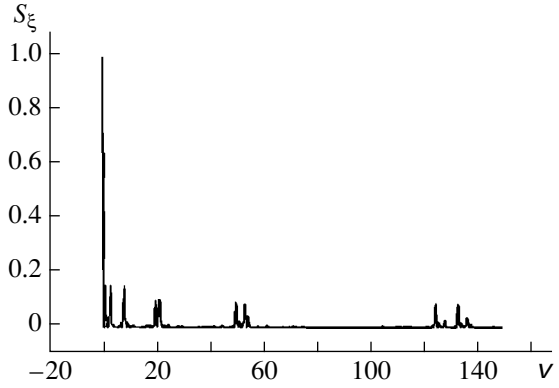


Fig. 5. $S_\xi(\nu)$ for $D_f = 0.75$.

Bezikovich fractional dimension). This statement is valid for interfaces in phase transitions as well.

To solve the problem posed in the title to this section, we take the following model of the fractal boundary between two media. Let the space $z < 0$ and $z > L$ be filled by a material with a permittivity ε_1. In the transition layer ($0 \le z \le L$), the permittivity varies according to the fractal law for alternate dielectric layers with ε_1 and ε_2:

$$\varepsilon(z) = \varepsilon_1 + (\varepsilon_2 - \varepsilon_1) L \frac{d}{dz} \alpha_\xi(z/L). \tag{13}$$

We consider the electromagnetic radiation of a charge uniformly moving with a velocity **v** through the transition layer. As a result, the transition radiation arises because the dynamic parameters of the medium (in our case, permittivity) change along the charge trajectory. We assume that

$$\frac{\Delta\varepsilon}{\varepsilon_1} = \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_1} \ll 1. \tag{14}$$

In [2], the formula for the spectral angular energy density of transition radiation at a diffuse boundary between two transparent media was obtained subject to condition (14). For backward radiation, this formula is read as

$$W_2(\omega, \theta_2) = \frac{q^2\omega^2[1 - \varepsilon\beta^2 + \beta\sqrt{\varepsilon}\cos\theta_2]^2 \sin^2\theta_2}{4\pi^2\varepsilon^{3/2}c^3[1 - \varepsilon\beta^2\cos^2\theta_2]^2} \times |\delta\varepsilon_k|^2\big|_{k = -\omega/v[1 + \beta\sqrt{\varepsilon}\cos\theta_2]}; \tag{15}$$

for forward radiation,

$$W_1(\omega, \theta_1) = \frac{q^2\omega^2[1 - \varepsilon\beta^2 + \beta\sqrt{\varepsilon}\cos\theta_1]^2 \sin^2\theta_1}{4\pi^2\varepsilon^{3/2}c^3[1 - \varepsilon\beta^2\cos^2\theta_1]^2} \times |\delta\varepsilon_k|^2\big|_{k = -\omega/v[1 - \beta\sqrt{\varepsilon}\cos\theta_2]}, \tag{16}$$

where $\beta = v/c$, q is the particle charge, $\varepsilon = (\varepsilon_1 + \varepsilon_2)/2$, $\cos\theta_1 = \mathbf{kv}/kv$, $\cos\theta_1 = -\cos\theta_2$, **k** is the wave vector in the line of radiation, and

$$\delta\varepsilon_k = \int dz \exp(-ikz)\delta\varepsilon(z). \tag{17}$$

Using formula (6) for the Fourier–Stieltjes transform of the Cantor function, one can now easily obtain the spectrum of the transition radiation of charges crossing the fractal boundary. To do this, we substitute the expression for $|\delta\varepsilon_k|^2$,

$$|\delta\varepsilon_k|^2 = \Delta\varepsilon^2 S_\xi(kL), \tag{18}$$

into the expression for the radiation probability.

For $\beta \ll 1$ and $W_1 = W_2 = W$,

$$W(k, \theta) = \frac{q^2\beta^2\Delta\varepsilon^2\sin^2\theta}{4\pi^2\varepsilon^{3/2}c} S_\xi(kL)k^2. \tag{19}$$

Thus, we may conclude that the fractal properties of the medium or transition processes induce the singularity of transition radiation through the structure factor S_ξ.

It is worth noting that the Cantor discontinuum is the simplest perfect nowhere dense set [3]. Therefore, fractal media obtained from the Cantor set are special cases of perfect nowhere dense sets.

CONCLUSIONS

In this paper, radiation spectra arising when fractal structures are involved were calculated. The calculations are based on the Cantor functions and the Fourier–Stieltjes transform.

An ordinary transition radiation occurs when charged particles cross the boundary between two media with different permittivities. This paper has elucidated the influence of the fractal properties of the transition layer on the spectra of fractal transition radiation. The spectra become wideband and self-similar. There appears the structure factor, which bears information on the fractal properties of the boundary between the media. Similar structure factors arise in

other types of radiation under appropriate electrodynamic structures (e.g., in dipole radiation accompanying a fractal transition process, scattering by fractal lattices, fractal generalization of the Smith–Parcell radiation, and also when the permittivity varies in time by a fractal law).

REFERENCES

1. J. Feder, *Fractals* (Plenum, New York, 1988; Mir, Moscow, 1991).
2. V. L. Ginzburg and V. N. Tsytovich, *Transition Radiation and Transition Scattering* (Nauka, Moscow, 1984).
3. A. Zygmund, *Trigonometric Series* (Cambridge Univ. Press, Cambridge, 1959; Mir, Moscow, 1965), Vol. 1.
4. C. Allen and M. Clouatre, in *Fractals in Physics*, Ed. by L. Pietronero and E. Tosatti (North-Holland, Amsterdam, 1986; Mir, Moscow, 1988).
5. B. M. Bolotovskiĭ, V. A. Davydov, and V. E. Rok, Usp. Fiz. Nauk **126** (2), 311 (1978) [Sov. Phys. Usp. **21**, 865 (1978)].

Translated by M. Fofanov

Technical Physics, Vol. 45, No. 12, 2000, pp. 1608–1611. Translated from Zhurnal Tekhnicheskoĭ Fiziki, Vol. 70, No. 12, 2000, pp. 102–105.

BRIEF COMMUNICATIONS

Experimental Investigation of a 25-MW Microwave (3-cm Range) Compressor Prototype

S. N. Artemenko, V. A. Avgustinovich, V. L. Kaminskiĭ, P. Yu. Chumerin, and Yu. G. Yushkov

Research Institute of Nuclear Physics, Tomsk Polytechnical University, pr. Lenina 30, Tomsk, 634050 Russia

Received February 16, 2000

Abstract—An experimental study of a resonance microwave (3-cm-range) compressor with gas insulation and energy output through a superdimensional coaxial interference switch is reported. The rated parameters of the compressor are output signal power ~25 MW, signal duration 2 ns, gain 26 dB, and efficiency ~57%. A gain of 20 dB was achieved at a peak output signal power of 12 MW, signal duration of 2 ns, efficiency of ~24%, and traveling wave power of 24 MW. © *2000 MAIK "Nauka/Interperiodica".*

(1) The output signal power of conventional resonance microwave compressors with identical gas-insulation storage cavities and energy output through interference switches built around T-junctions is 1–3 MW in the 3-cm range and 50–200 MW in the 10-cm range. The respective gains are 13–15 and 15–17 dB at a signal duration of 2–5 ns [1]. In such compressors, higher performance cannot be reached because of a low intrinsic quality factor Q_0 of the single-mode cavities and their small cross-section area S. This follows from expressions for the ultimate gain M^2 of resonance compressors

$$M^2 = Q_0/\omega T \quad (1)$$

and the power P of an electromagnetic wave transmitted through the waveguide

$$P = AE^2S, \quad (2)$$

where ω is the cyclic operating frequency of the cavity, T is the doubled-travel time of a wave in the cavity, A is a coefficient defined by the waveguide geometry and the propagating mode, and E is the intensity of the electric component of the wave.

One can see from formulas (1) and (2) that, for given ω, T, and E, the only way of increasing the gain and power of the compressor is to use cavities with a higher quality factor and a larger cross-section area. This means that superdimensional large-volume cavities should be used. In this case, for cylindrical cavities, the gain and the power will increase linearly and quadrantically with the cavity radius, respectively. Calculations have shown that, in the 3-cm range, superdimensional cylindrical cavities with $H_{01(p)}$ modes may have the quality factor $Q_0 = (1–1.5) \times 10^5$ for a cross-section area of 50–100 cm^2. Therefore, for T = 2–5 ns, the gain of a compressor equipped with such cavities may reach ~30 dB, and the output signal power may range up to ~10–100 W or higher. However, the design of such high-performance systems is limited by the lack of efficient means and devices for fast energy output from superdimensional cavities.

Earlier [2], we reported on the possibility of fast energy output from large-size (multimode) axisymmetric cavities through interference switches built around superdimensional coaxial lines. In these lines, the degree of cavity–load coupling is controlled with the help of a commutator that governs the interference of waves radiated into the load from the line and a radial gap, i.e., a break in the inner conductor of the coaxial line. Early experimental results confirmed the efficiency of such switches.

In this work, we provide new data on energy output from a multimode cylindrical cavity with the $H_{01(p)}$ mode through a superdimensional coaxial interference switch with the H_{01} operating mode. Our aim was to develop a prototype of a 3-cm-range microwave compressor whose gain and output signal power are one order of magnitude higher than similar characteristics of conventional compressors.

(2) The prototype is designed for operation with gas electrical insulation of the storage volume, which simplifies switching the system from the storage to output mode. The compressor output signal power is ~25 MW at a pulse duration of ~2 ns. Since the limiting power flux density of an electromagnetic wave transmitted through a gas-filled waveguide is ~1–5 MW/cm^2 [3], the cavity cross-section area must be no less than ~25 cm^2. Therefore, in the experiments, we used a cavity

1063-7842/00/4512-1608$20.00

with a diameter of 9 cm (a cross-section area of ~65 cm^2) and a switch with the ~50-cm^2-area transmitting part of the coaxial line. In order for the doubled travel time of the wave in the cavity to be comparable to the required duration of output signals, the cavity length was ~25 cm.

In order to lower the mode coupling at windows used to couple the cavity with the feeding circuit, the energy was delivered to the cavity through an *E*-matched T-junction and the operating mode was excited through two windows placed at the midpoints of radii belonging to the same diameter of the cavity end face. For such an excitation, the quality factor of the $H_{01(p)}$-mode cavity (without the switch) equals (0.9–1.2) $\times 10^5$ in a frequency range of 8.2–11.2 GHz. The spread in the quality factor is most probably due to the nonuniformity of the oscillation spectrum density in this frequency band and, correspondingly, with the different effect of mode coupling on the operating mode.

To reveal how the switch changes the quality factor, we studied the frequency characteristic of this parameter for a cavity with the inner conductor of the switch inside. The conductor was fixed coaxially with the cavity on its end face free of input coupling windows. The conductor length and diameter were 4.5 and ~9 cm, respectively. These values are optimal for a switch operating at 9.12 GHz. It was found that the operating-mode quality factor increases monotonically with frequency from 4×10^4 at 8.2 GHz to 7×10^4 at 11.2 GHz. This, most probably, is caused by frequency dependence of the attenuation constant for the H_{01} mode propagating in the coaxial line. At frequencies close to the cutoff frequency (6.8 GHz), the attenuation constant is large and the quality factor lowers substantially. When the frequency increases, the attenuation constant decreases and the quality factor increases. Such an influence of the attenuation constant means that we should use switches that are as short as possible and frequencies substantially exceeding the cutoff frequency.

The replacement of the cavity end face by the switch did not result in a noticeable decrease in the cavity quality factor, approximately equal to 5×10^4. In this case, the lengths of the input and output switch arms were 4.5 cm, the height of the radial gap was ~2.45 cm, and the diameter of the outer conductor (9 cm) was equal to the cavity diameter. Oscillation modes excited in the cavity, switch input arm, and radial gap were $H_{01(12)}$, $H_{01(2)}$, and $H_{01(1)}$, respectively. As follows from calculations, for the quality factor attained, the system gain may be ~26 dB.

The transition loss of the switch strongly depends on the radial gap height and nonparallelism of the switch walls. With a change in the gap height by 5–10% or in a wall misalignment by 2–3°, the cavity becomes almost completely open. Figure 1 depicts the cavity quality factor versus the gap height. Note also that, due to radiation losses, the switch removed almost all spurious modes and thus made the cavity spectrum radially sparse.

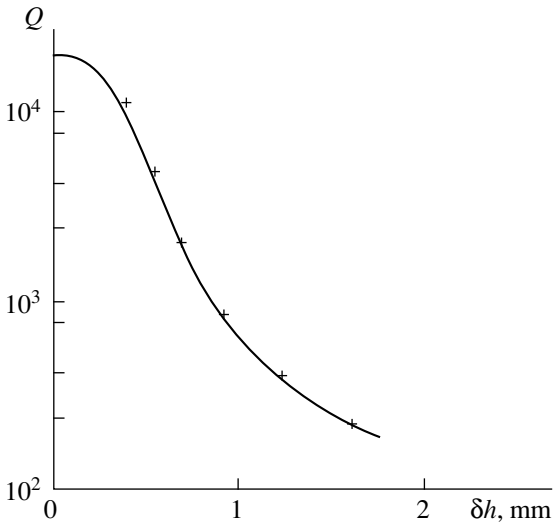


Fig. 1. Cavity quality factor vs. height of the switch radial gap.

(3) Investigations at a high power level were performed with the setup presented in Fig. 2. The setup comprises microwave oscillator *1*, circulator *2*, directional couplers *3*, phase shifter *4*, *E*-plane *T*-junction *5*, input windows *6*, cavity *7*, superdimensional coaxial interference switch *8* (with input arm *8a*, radial gap *8b*, and output arm *8c*), commutator *9*, commutator dischargers *10*, waveguide junctions *11*, matched loads *12*, and detectors *13*.

The system was fed from a 120-kW pulsed magnetron with a pulse duration of ~1 μs. The stored microwave energy was extracted through a switch with a "packet" of four identical gradient waveguide junctions at its end (Fig. 2), which were arranged in 90° intervals and transformed the H_{01} operating mode of the switch into the H_{10} mode passing in the rectangular waveguide of each of the junctions. The elements of the waveguide paths (directional couplers, loads, etc.) required for measuring the radiation characteristics (power, duration, and mode type) were connected to the outputs of these junctions. The power was extracted through the packet in order to ensure accurate determination of the output signal parameters.

Storage-to-output regime switching of the system was performed with a trigatron commutator. This commutator initiates a microwave discharge at the maximum of the φth component of the electric field in the radial gap. This discharge, developing parallel to field lines, causes a phase inversion of the wave radiated through the gap and in-phase summation of this wave with the cavity one; thus, the cavity opens. The discharge was initiated either in air or in argon filling the cavity and the switch at atmospheric pressure. For switching in air, the pulse duration was 10–20 ns,

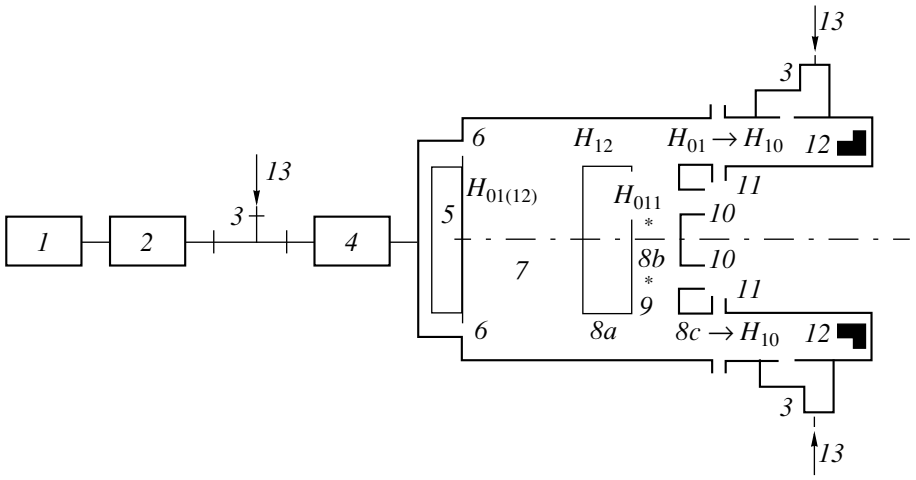


Fig. 2. Experimental setup for testing the microwave compressor prototype at a high power level.

which is 5–10 times longer than in argon. Correspondingly, the output pulse amplitude in air was much (approximately by one order of magnitude) smaller than that in argon. These differences are associated with specific features of microwave discharges in air and argon.

At the maximum power of input pulses (~120 kW), that of the traveling wave in the cavity reached 24 MW. In this case, at argon-switching energy extraction, pulses with a duration of ~2 ns (at a level of 0.5) and a power of 12 MW (~3 MW per junction channel) were formed at the output of the coaxial line with the H_{01} mode. The system efficiency under such conditions reached 24%. A typical oscillogram of the output signal envelope is shown in Fig. 3.

The effective diameter of the cavity formed in the radial gap was estimated at ~5.6 cm from the system operating frequency and the gap height. The cross-section area of this cavity equals ~20 cm^2. Therefore, for gas insulation, the power of the traveling wave in this cavity may reach 20–200 MW. The ultimate level of the output power in this system will be of the same order of magnitude (this level depends on the gas type).

In conclusion, note some problems concerning the stabilization of the microwave discharge along the φth electric field component in the radial gap. The instability of the plasma discharge channel often causes a substantial spread in the amplitude and duration of output signals. At this stage of our investigations, these problems seem to be the most serious.

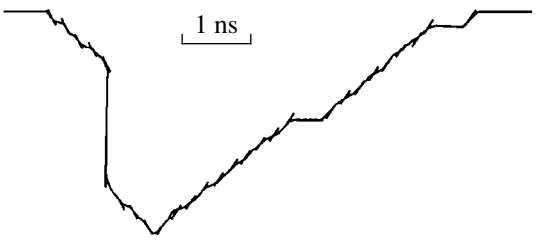


Fig. 3. Envelope of the output microwave pulses.

(4) Thus, in this work, we demonstrated the possibility of designing a 3-cm-range microwave compressor with an output signal power exceeding 10 MW. It was shown that, at a high power level, a superdimensional coaxial interference switch with the H_{01} operating mode can be used as an efficient element ensuring fast energy output. Such a switch can be used both to form high-power short (nanosecond) microwave pulses and to generate high-power energy-intensive (long) microwave pulses when the energy is extracted from large-size cavities, including combined systems, e.g., built around conventional coaxial or relativistic coaxial magnetrons.

It seems realistic that, by increasing the oscillator power to 250 kW and improving the cavity electrical strength by raising the argon pressure, the signal power of the compressor can be raised to 25 MW. If the system is filled by SF_6 gas and the input power is increased to 2–3 MW, the signal power can be raised still further (approximately by one order of magnitude). If the system is used as an energy output in large-size cavities with volumes of 10–30 l, it could be fed from (10–30)-MW oscillators; in this case, 100–200-MW pulses with a duration of 50–100 ns seem to be a possibility.

Moreover, it seems possible to bring the output signal power of the system to 500 MW by replacing gas insulation by vacuum insulation and performing switching in a gas-filled tube, as was proposed in [4]. This possibility stems from the fact that the switch becomes almost completely open when the height of the radial gap changes insignificantly (Fig. 1). This allows the discharger to be displaced from the maxi-

mum of the electric component of the gap field. In the 10-cm range, the system power may be one order of magnitude higher than in the 3-cm range.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 97-02-16461.

REFERENCES

1. A. N. Didenko and Yu. G. Yushkov, *High-Power Microwave Nanosecond Pulses* (Énergoatomizdat, Moscow, 1984).
2. S. N. Artemenko, V. L. Kaminskiĭ, and Yu. G. Yushkov, Zh. Tekh. Fiz. **63** (2), 105 (1993) [Tech. Phys. **38**, 111 (1993)].
3. A. L. Fel'dshteĭn, L. R. Yavich, and V. P. Smirnov, *Components of Waveguide Technology: A Handbook* (GÉI, Moscow, 1963).
4. A. L. Vikharev, A. M. Gorbachev, O. A. Ivanov, *et al.*, Pis'ma Zh. Tekh. Fiz. **24** (20), 6 (1998) [Tech. Phys. Lett. **24**, 791 (1998)].

Translated by A. Kondrat'ev