BRIEF COMMUNICATIONS

Reflection of microwave radiation from a magnetic fluid in which mechanical vibrations are excited

D. A. Usanov, A. V. Skripal', and S. A. Ermolaev

N. G. Chernyshevski Saratov State University, 410001 Saratov, Russia
(Submitted November 19, 1996)
Zh. Tekh. Fiz. **68**, 138 (August 1998)

[S1063-7842(98)02408-8]

It was shown in Ref. 1 that the reflection of microwaves from a magnetic fluid varies with the orientation of the induction vector $\overline{B}$ of the magnetic field applied to it. This is because when $\overline{B}$ is oriented along the electric field vector $\overline{E}$ of the microwaves, the ferromagnetic particles form filaments parallel to $\overline{E}$, whereas when $\overline{E}$ and $\overline{B}$ are perpendicular, the filaments are perpendicular to the direction of $\overline{E}$. One consequence of this is an increase in the microwave reflectivity R with increasing $\overline{B}$ for $\overline{E}\|\overline{B}$ and with decreasing $\overline{B}$ for $\overline{E}\perp\overline{B}$.

The excitation of mechanical vibrations in a magnetic fluid leads to a loss of the structure established in it by the action of the magnetic field. This phenomenon has been observed experimentally by measuring the scattering of laser radiation by a magnetic liquid.[2]

It can be theorized that the excitation of mechanical vibrations and the accompanying loss of the induced structure of the magnetic described above should lead to a decrease in the microwave reflectivity when $\overline{E}\|\overline{B}$ and to an increase when $\overline{E}\perp\overline{B}$ in comparison to the case where mechanical vibrations are not excited.

This hypothesis was investigated experimentally. The measurements were carried out in the 3-cm wavelength range. The magnetic fluid completely filled a segment of a waveguide of cross section 2.4×1.0 cm and length 10 cm. Such a fairly large thickness of the liquid layer was chosen for the purpose of eliminating the influence of its rear boundary on the microwave reflectivity. The dependence of R on $\overline{B}$ was measured for $\overline{E}\|\overline{B}$ (curve *1*) and $\overline{E}\perp\overline{B}$ (curve *2*) in the absence of mechanical vibrations and in their presence (curves *3* and *4*, respectively). The results presented in Fig. 1 confirm the hypothesis that the excitation of mechanical vibrations in a magnetic fluid leads to a decrease in the reflectivity relative to the mechanically quiescent case at the same value of the magnetic field for $\overline{E}\|\overline{B}$ and to an increase in R for $\overline{E}\perp\overline{B}$.

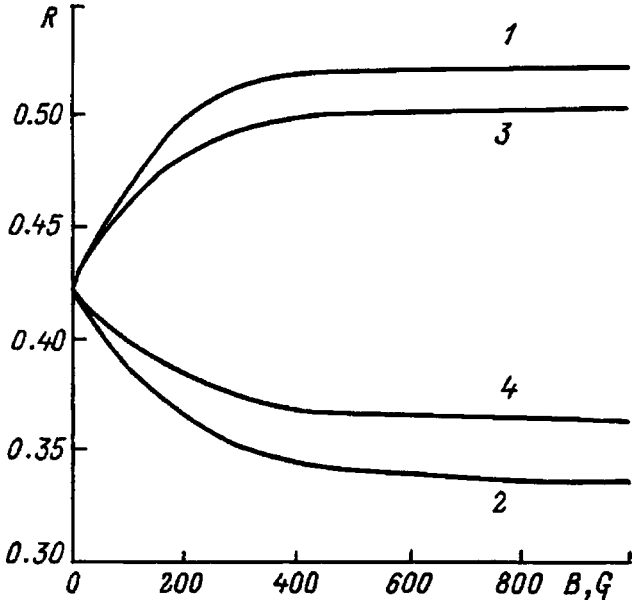

FIG. 1.

This law also holds for the characteristic regions of saturation on the plot of the dependence of R on $\overline{B}$ and provides evidence that the former structure of the magnetic fluid is not restored at any value of the magnetic field in the presence of mechanical vibrations.

[1] D. A. Usanov, Al. V. Skripal', An. V. Skripal', and S. A. Ermolaev, Pis'ma Zh. Tekh. Fiz. **18**(23), 44 (1992) [Sov. Tech. Phys. Lett. **18**(12), 780 (1992)].
[2] D. A. Usanov, A. V. Skripal', S. A. Ermolaev, and V. V. Panov, Pis'ma Zh. Tekh. Fiz. **21**(17), 1 (1995) [Tech. Phys. Lett. **21**(9), 687 (1995)].

Translated by P. Shelnitz

Quick estimation of the resistivity of collector and the depth of the collector junction in the design of bipolar $n-p-n$ transistors

A. N. Frolov and S. V. Shutov

Kherson Industrial Institute, 325008 Kherson, Ukraine
(Submitted October 13, 1997)
Zh. Tekh. Fiz. **68**, 139–140 (August 1998)

[S1063-7842(98)02808-6]

In the design of bipolar integrated circuits and discrete $n-p-n$ transistors, it is often necessary to determine the structural parameters from the specified gain, frequency characteristics, breakdown voltage of specific junctions, etc. The well-known techniques take an inverse approach, i.e., one determines the parameters of the structural elements (thicknesses of the active regions, dopant concentrations around the $p-n$ junctions, etc.) from the technological conditions and then determines the electrical characteristics of the transistors from the known parameters of the structure.[1,2] Although this way of performing the calculation is found to give better agreement with experiment, it is not altogether convenient for practical use and requires awkward and repeated calculations.

We were thereby motivated to undertake a number of experiments to investigate the parameters of planar $n-p-n$ transistors of mass-produced integrated circuits for the purpose of establishing the interrelationships between the electrical and structural parameters of bipolar transistors. The objects of study were the planar transistors of the integrated circuits KR514AP1 (an approximate analog of the Philips SAA1060), UA03KhP2, UA03KhP1, KR1031KhA1 (control circuits for four-phase collectorless motors of different powers).

Statistical data on the depth of the base diffusion (the depth of the collector junction) χ_{jb} and on the resistivity of an epitaxial film (resistivity of the collector) $\rho_{ep.f}$, and the measured values of the breakdown voltage of the collector–base junction $U_{c.b.o}$ and the emitter–base junction $U_{em.b}$ are given in Table I. Also given are the calculated values of $\varkappa$, the coefficient of elevation of the breakdown voltages.

Analytical processing of the tabulated results suggests the following expression for calculating the depth of the collector junction:

$$\chi_{jb}=0.144\cdot(U_{c.b.o}/\varkappa)^{0.7}, \tag{1}$$

while the choice of the resistivity of the collector is made according to the formula

$$\rho_{ep.f}=4.8\times10^{-4}(U_{c.b.o}/\varkappa)^{1.92}. \tag{2}$$

The values of χ_{jb} and $\rho_{ep.f}$ are also given in Table I. In comparing the results of the calculation and the statistical data, one sees that the deviation is not more than 5%. However, the presence of the coefficient $\varkappa$ in expressions (1) and (2) necessitates making an experimental determination of the breakdown voltages of the collector junction prior to calculating $\rho_{ep.f}$ and χ_{jb} (since $\varkappa$ is by definition the ratio of the experimental value of $U_{c.b.o.}$ to the calculated value). To simplify the design process we carried out experiments to establish the relationship between the coefficient $\varkappa$ and other characteristics of planar $n-p-n$ transistors. Several batches of planar transistors were fabricated on silicon structures with epitaxial films. The resistivity of the epitaxial films doped with phosphorus were 2.4, 3.5, and 4.5 $\Omega\cdot$cm. The base regions were formed by the implantation of boron at doses of 4, 15, 75, and 450 μC/cm^2. The drive-in temperature was chosen within the range 1150–1180 °C, the drive-in time of the base impurity varied from 35 min for the implantation of boron in a dose of 450 μC/cm^2 to 650 min at a dose of 4 μC/cm^2. The emitter regions were formed in the same way for all the batches of transistors, by the diffusion of phosphorus at a temperature of 1040°. This made it possible to span a wide spectrum of values of the depth of the base diffusion, of the impurity concentration gradients in the base region, and of the breakdown voltages of the emitter junction.[3]

TABLE I. Experimental data and computational results on the depth of the base diffusion and the resistivity of the collector.

				Statistical data		Calculated data	
Type of IC	$U_{em.b}$, V	$U_{c.b.o.}$, V	$\varkappa$	χ_{jb}, μm	$\rho_{ep.f}$, $\Omega\cdot$cm	χ_{jb}, μm	$\rho_{ep.f}$, $\Omega\cdot$cm
KR514AP1	7.0–7.2	30–33	1.04	1.5	0.3	1.52	0.305
UA03KhP2	7.5–7.7	63–67	1.05	2.5	1.2	2.53	1.245
UA03KhP1	7.8–7.9	83–87	1.055	3.2	2.0	3.11	2.095
KR1031KhA1	8.1–8.2	125–130	1.06	4.0	4.5	4.06	4.56

A processing of the experimental data yielded a completely determined relation between the coefficient of elevation of the breakdown voltage ($\varkappa$) and the breakdown voltage of the emitter–base junction:

$$\varkappa = 1 + 3.6 \times 10^{-4} (U_{em.b})^{2.45}. \quad (3)$$

Thus for designing $n-p-n$ transistors from the specified values of the breakdown voltages of the collector–base and emitter–base junctions, the empirical expressions obtained here permit an easy calculation of the depth of the base diffusion and the resistivity of the collector region.

[1] Ya. A. Fedotov, *Silicon Planar Transistors* [in Russian], Sov. Radio, Moscow (1973), 336 pp.
[2] A. B. Greben, *Design of Analog Itegrated Circuits* [in Russian], Énergiya, Moscow (1976), 435 pp.
[3] N. A. Samoĭlov, A. N. Frolov, and S. V. Shutov, Pis'ma Zh. Tekh. Fiz. **22**(7), 36 (1996) [Tech. Phys. Lett. **22**, 281 (1996)].

Translated by Steve Torstveit

Modeling the deposition of materials on capillary walls from vapor–gas mixtures under pulsed heating

V. V. Zotov and E. P. Prokop'ev

A. Yu. Malinin Scientific-Research Institute of Materials Science, 103460 Moscow, Russia
(Submitted August 4, 1997)
Zh. Tekh. Fiz. **68**, 141–142 (August 1998)

The deposition of silicon in capillaries during the thermochemical reaction of silicon tetrachloride with hydrogen under pulsed heating is considered theoretically. It is shown that such a reaction results in a significant increase in the uniformity of the deposition rate along the length of a capillary. © *1998 American Institute of Physics.* [S1063-7842(98)02608-7]

The practical fabrication of semiconductor devices and integrated circuits sometimes calls for the deposition of materials (semiconductors, insulators, and metals) on the walls of capillary slots, holes, and depressions to make the structure planar and to form capacitors, vertical interconnections, field-effect transistors with vertical channels, etc. The known methods for such deposition are based on the performance of steady-state thermochemical reactions in vapor–gas mixtures at reduced pressure. However, steady-state processes are characterized by a higher growth rate for the layers deposited at the mouths of capillaries than for the layers deposited within them. As the pressure in the reactor is lowered, the uniformity of the thickness of the layer deposited increases, but it remains insufficient for many applications. In addition, as the pressure is lowered, the deposition rate and, hence, the productivity decrease.

The situation changes dramatically when pulsed heating of the substrates in vapor–gas mixtures is employed. The features of the deposition of silicon in capillaries in such a regime are considered below as an example.

The deposition of silicon from vapor–gas mixtures in capillaries with pulsed heating of the substrates by radiation has several valuable advantages over the steady-state regime.[1–9]

Let a silicon substrate in an atmosphere of a vapor–gas mixture (SiH_4+H_2 or, for example, $SiCl_4+H_2$) have a hole passing through it or a depression of diameter D and depth L. For simplicity in solving the problem we assume that a thermal pulse which is linear in time t acts on the substrate during irradiation, so that the temperature of the surface of the cavity, i.e., the capillary, on whose walls silicon layers are deposited from the vapor–gas mixtures, varies according to the law[1]

$$T_s(t)=300+2053t. \tag{1}$$

The job of the theory is to calculate the growth rate of the silicon layers $V_p[T_s(t),x]$ in order to determine the thickness of the layer grown during a pulse:

$$d_p[T_s(t_{\max}),x]=2\int_0^{t_{\max}} V_p[T_s(t),x]\,dt. \tag{2}$$

If the rate of the layer deposition reaction is far higher than the rate of the temperature rise, we can treat the process in the cavity in the steady-state approximation for the case of an isothermal process at each moment in time in the range from 0 to t. For a blind cavity we direct the x axis from the bottom (the origin of the x axis) into the space of the vapor–gas mixture, and for an open cavity (a hole through the substrate) we direct it from the midplane of the substrate. In this case, according to Refs. 1–9, the growth rate (A/min) of the silicon layers on the cavity walls in the isothermal case can be written in the form

$$V_p(x)=(7.23\times 10^{10})\times\frac{\dfrac{k_0P}{RT_s}\exp\left(-\dfrac{E}{RT_s}\right)\dfrac{D_0T_sP_0}{RT_0^2}\exp\left[-\dfrac{D_0}{\delta(x)}\dfrac{T_s}{DT_0\overline{\omega_0}}x\right]x_0}{\dfrac{k_0P}{RT_s}\exp\left(-\dfrac{E}{RT_s}\right)\delta(x)+\dfrac{D_0T_sP_0}{RT_0^2}\exp\left[-\dfrac{D_0}{\delta(x)}\dfrac{T_s}{DT_0\overline{\omega_0}}x\right]}. \tag{3}$$

Here P is the pressure in the cavity, k_0 is the preexponential (frequency) factor, E is the activation energy, R is the universal gas constant, T_s is the temperature of the cavity surface, D_0 is the molecular diffusion coefficient of the silicon-containing reactant, $P_0=1.013\times 10^6$, $T=273$ K, x_0 is the concentration of the silicon-containing reactant, D is the diameter of the cavity, ω_0 is the linear flow velocity of the vapor–gas mixture in the cavity, and x is the coordinate along the cavity. The thickness of the boundary layer in the cavity $\delta(x)$ can be estimated from the formula

$$\delta(x)=\frac{\sqrt{x}}{\left[\dfrac{\rho_o\omega_0}{\eta(T_s)}\right]^{1/2}}. \tag{4}$$

Here ρ_0 is the hydrogen density, and $\eta(T_s)$ is the dynamic viscosity, which equals

$$\eta(T_s)=\eta_0\left(\frac{T_s}{T_0}\right)^{1/2}\frac{1+234/T_0}{1+234/T_s}, \tag{5}$$

where η_0 is the value of $\eta(T_s)$ at $T_0=273$ K.

Depending on the diameter of the cavity, various flow regimes can be realized for the vapor–gas mixture as the

temperature rises: molecular flow, slip, laminar flow, and turbulent flow.[10,11] We assume that the pressure in the cavity varies as the temperature rises according to the law

$$P = P_0 \frac{T_s(t)}{T_0}. \tag{6}$$

The pulsed application of an incoherent light source to a silicon substrate in a $SiCl_4 + H_2$ atmosphere is accompanied by a fairly sharp click, which is characteristic of a weak shock wave.[10] The ratio between the intensity of a shock wave in a capillary and its intensity in the reactor volume can increase sharply as a function of the diameter of the capillary. However, according to Ref. 11, the propagation rate of a shock wave in a capillary can be assumed to be equal in order of magnitude to the speed of sound, i.e., $\omega_0 \approx 10^5$cm/s. With this value of ω_0 the thickness of the boundary layer in an open cavity of length 175μm at the point $L/2$ would be $\approx 8\times 10^{-4}\text{cm} = 8\mu\text{m}$. Thus, under real experimental conditions δ can be taken equal to the diameter of the cavity D, i.e., $\delta \approx D$. Estimates for $D < 10\mu$m and $L \approx 175\mu$m show that the first term in the denominator of formula (3) is far smaller than the second term, and formula (3) takes the following form:

$$V_p = 7.23\times 10^{10} \frac{k_0 P}{R T_s} \exp\left(-\frac{E}{R_T T_s}\right) x_0, \tag{7}$$

i.e., the growth rate does not depend on x, and high uniformity of the thickness of the silicon layers grown is ensured. According to Refs. 1–9, the thickness of the layers grown can be estimated in this case from the formula

$$d_p(t_{\max}) \approx 2V_p(t_{\max})/a_0, \tag{8}$$

where

$$a_0 = \ln V_p(t_{\max})/t_{\max}. \tag{9}$$

Let us estimate the values of $V_p(t_{\max})$ and $d_p(t_{\max})$ using formulas (7)–(9) for the standard conditions of the tetrachloride process: $k_0 \approx 10^7$cm/s, $P = 1.013\times 10^6 T_s/T_0$ dyn/cm^2, $E = 38\,000$ cal/mol, and $x_0 = 10^{-2}$. For $T_s(t_{\max}) = 1473$ K, i.e., at $t_{\max} = 0.571$ s, we have $V_p(t_{\max} = V_p(0.571\text{ s}) = 7.423 \times 10^5$Å/min and $d_p(t_{\max} = d_p(0.571\text{ s}) = 1.499\times 10^3$Å $= 149$ nm at $P_{\max} = P_0 T_s(t_{\max})/T_0$. At atmospheric pressure $P_{\max} = P_0 = 1.013\times 10^6$dyn/cm^2, and we get $d_p(t_{\max}) = 26.3$ nm, i.e., roughly 7–35 pulses of type (1) are needed to fill in a cavity of diameter $D \approx 1\mu$m, in agreement with our experimental data.[1]

[1] E. P. Prokop'ev, V. V. Zotov, and V. N. Statsenko, Vysokochist. Veshchestva, No. 5, 165 (1993).
[2] E. P. Prokop'ev, Vysokochist. Veshchestva, No. 6, 63 (1994).
[3] E. P. Prokop'ev, V. V. Zotov, and V. N. Statsenko, Vysokochist. Veshchestva, No. 1, 39 (1994).
[4] E. P. Prokop'ev, V. N. Statsenko, and V. V. Zotov, Vysokochist. Veshchestva, No. 4, 54 (1995).
[5] E. P. Prokop'ev, Vysokochist. Veshchestva, No. 3, 57 (1995).
[6] E. P. Prokop'ev, Vysokochist. Veshchestva, No. 3, 66 (1995).
[7] E. P. Prokop'ev, Fiz. Khim. Obrab. Mater. No. 1, 76 (1996).
[8] E. P. Prokop'ev, S. V. Petrov, and V. S. Belousov, Peterb. Zh. Elektron., No. 1, 29 (1996).
[9] E. P. Prokop'ev, Zh. Tekh. Fiz. **65**, 22 (1995) [Tech. Phys. **40**, 128 (1995)].
[10] E. P. Prokop'ev, V. V. Zotov, and V. N. Statsenko, Zh. Fiz. Khim. **67**, 503 (1994).
[11] Ya. B. Zel'dovich and Yu. P. Raĭzer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* [Academic Press, New York, 1966–1967; Nauka, Moscow, 1986, 686 pp.].

Translated by P. Shelnitz

ERRATA

Erratum: Optical bistability and modulation of light in a thin-film resonator based on total internal reflection [Tech. Phys. 42, 1406–1410 (December 1997)]

A. D. Soler and Yu. P. Udoev

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
Zh. Tekh. Fiz. **68**, 143 (August 1998)

[S1063-7842(98)02708-1]

The label on the abscissa in Fig. 2 on page 1407 should read *i*, deg.
In the caption to Fig. 3 on page 1408, it should read $n_3 = 1.50$.
In the caption to Fig. 4 on page 1409, it should read: Δ_i, deg: ... *3*—0.004, ...

Translated by Steve Torstveit

Properties of the phase space and bifurcations in the complex Lorenz model

A. G. Vladimirov

Scientific-Research Institute of Physics at St. Petersburg State University, 198904 St. Petersburg, Russia

V. Yu. Toronov and V. L. Derbov

Saratov State University, 410071 Saratov, Russia

(Submitted January 22, 1997)
Zh. Tekh. Fiz. **68**, 1–9 (August 1998)

The geometrical structure of the phase space and bifurcations in the complex Lorenz model are investigated. It is shown that the hierarchy of bifurcations in a single-mode laser with detuning of the resonator frequency from the frequency of a spectral line is similar to the hierarchy of bifurcations of the logistic map. © *1998 American Institute of Physics.* [S1063-7842(98)00108-1]

INTRODUCTION

The complex Lorenz equations, or complex Lorenz model

$$\dot{x} = -\sigma(x-y),$$

$$\dot{y} = -(1-i\delta)y + (r-z)x,$$

$$\dot{z} = -bz + \frac{1}{2}(x^*y + xy^*), \tag{1}$$

was first introduced by Gibbon and McGuinness[1] as a generalization of the standard Lorenz model.[2,3] The complex Lorenz model differs from the latter in that x and y are complex. Formally, this complexity stems from the presence of the real parameter δ and the complex parameter $r = r_1 + ir_2$, which are not a part of the original Lorenz model. The complex Lorenz model has important bearing on nonlinear dynamics because it is a universal finite-dimensional approximation for the class of distributed systems that exhibit so-called dispersion instability of a steady-state solution at a point of parameter space $\mu = \mu_c$ (Ref. 1). For such systems it has been shown[1] that the expansion of the vector representing the perturbation of the steady-state solution in powers of the small parameter $\varepsilon = (|\mu - \mu_c|)^{1/2}$ in a certain approximation leads to a system of equations, equivalent to (1), for the coefficients of the expansion. As an example, the model of a baroclinic instability in the atmosphere[4,5] is investigated in Ref. 1.

The variables in Eqs. (1) are the perturbation amplitudes relative to a spatially homogeneous solution of partial differential equations and, as such, do not admit a clear-cut physical interpretation in application to a baroclinic instability. However, there are systems for which the variables of the complex Lorenz model are observable quantities. These systems are lasers and masers, for which x and y represent the slowly varying complex amplitudes of the electric field and polarization of the medium, respectively, and z denotes the population difference between the energy levels of the working transition. In reality Eqs. (1) appeared in quantum electronics long before their ''discovery'' by Gibbon and McGuinness[1] (see, e.g., Ref. 6). The direct observability of the variables makes lasers the most suitable object for experimental implementation of the dynamics associated with the complex Lorenz model. Indeed, the results of experiments with single-mode, far-infrared lasers have been found to best match the results of numerical integration of the system (1) (Ref. 7). And even though Eqs. (1) provide the most realistic description of this type of laser exclusively, they are the simplest model reflecting such fundamental laser properties as the threshold character of lasing, frequency pulling, and the capability of generating complex wave modes.

However, despite a wealth of papers on the complex Lorenz model,[1,8–17] to this day it has not received the attention that it deserves. The investigations of this model are concerned primarily with the analysis of particular regimes, usually by numerical methods. In our opinion, the most incisive results have been obtained in Refs. 10 and 11 in regard to the stability analysis of simple periodic solutions corresponding to steady-state lasing. The objective of the present study is to investigate the geometrical structures of the complex Lorenz model and certain global properties of its solutions in connection with distinctive features of this structure.

An important geometrical property of the complex Lorenz model is its invariance under the transformation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} xe^{i\psi} \\ ye^{i\psi} \\ z \end{pmatrix}, \tag{2}$$

where ψ is an arbitrary phase constant.

This transformation corresponds to the group $U(1)$ acting in that subspace C^2 of the total phase space $\mathcal{H}$ which refers to the variables x and y. To picture the role of $U(1)$ — the symmetry for the structure of the limit sets in the phase space of the complex Lorenz model, it is sufficient to consider the case $\delta = r_2 = 0$. It can be shown[17] that in this case any trajectory in $\mathcal{H}$ is attracted to the invariant three-dimensional hyperspace

$$\frac{\mathrm{Re}(x)}{\mathrm{Im}(x)} = \frac{\mathrm{Re}(y)}{\mathrm{Im}(y)} = \mathrm{const}, \tag{3}$$

where the constant on the right side depends on the initial conditions.

On this hypersurface the real and imaginary parts of x and y vary synchronously with time, satisfying the equations of the original Lorenz model. Even under these conditions, however, the attractors of the original Lorenz model and its complex generalization are not identical. Indeed, if a given trajectory is attracted to a given limit set situated on the hypersurface defined by Eqs. (3), it follows from the symmetry property (2) that a trajectory that differs from the given trajectory by a certain common phase of x and y will be attracted to a set that is the image of the given set under the action (2) with the corresponding phase ψ. Consequently, an attractor of the Lorenz model in the given situation represents the direct sum of an infinite number of sets isomorphic to an attractor of the original model.

This example graphically demonstrates the more complex structure of the phase space of the complex Lorenz model in comparison with the original model and the role of symmetry in this greater complexity. However, it will be shown below that symmetry is not a tool that can be used to simplify the problem. Here we note an interesting analogy between the role of symmetry for physical systems described by the complex Lorenz model and for quantum systems described by a wave function. For neither system does the general phase of the state vector (which would be the vector with components x and y in the case of the complex Lorenz model) carry any information about the physical state; it merely characterizes the result of interference of a given state with certain other states. In quantum mechanics ''spurious'' information can be filtered out by application of the density matrix formalism. If a given state vector in a Hilbert space is described by N complex numbers, the corresponding density matrix is characterized by $2N-1$ real numbers.

The indicated analogy permits this approach to be applied to the complex Lorenz model. In Sec. 1. we introduce a special projective space, in which states differing by the common phase of x and y are treated as equivalent, and we derive equations of motion for the complex Lorenz model in this space. We also show how all information on the physical state of a system and phase evolution can be reproduced by means of these equations. In Sec. 2. we use these equations to analyze the boundedness properties of the limit sets of the complex Lorenz model in the projective space, and we show how these properties are related to prominent features of the phase dynamics and homoclinic bifurcations. In particular, the well-known bifurcation of generation of a homoclinic ''butterfly'' in the original Lorenz model is found to have codimensionality 2 in the complex model. On the basis of these results, in Sec. 3. we construct and analyze a one-dimensional map in the vicinity of a homoclinic bifurcation point for the domain of parameters of the complex Lorenz model. The results of our analysis of the map are compared with the results of a direct numerical investigation of bifurcations.

1. NORMAL FORM OF THE EQUATIONS AND THE PROJECTIVE SPACE

It will be advantageous below to use another system of variables in addition to x, y, and z. Let

$$\sigma(r_1-1)-\frac{\delta^2}{4}\equiv\eta>0, \quad b<2\sigma. \tag{4}$$

We introduce the change of variables

$$x'=\eta^{-3/4}ax, \quad y'=\eta^{-5/4}\sigma a\left(y-\left(1+\frac{i\delta}{2\sigma}\right)x\right),$$

$$z'=\eta^{-1}\sigma\left(z-\frac{xx^*}{2}\right), \quad t'=t\sqrt{\eta}, \tag{5}$$

where

$$a=e^{-i\delta t/2}\sqrt{\frac{2\sigma-b}{2}}.$$

After this substitution, which is similar to one proposed in Refs. 18 and 19 for the Lorenz model, Eqs. (1) assume the form

$$\frac{dx'}{dt'}=y', \quad \frac{dy'}{dt'}=(1+i\nu)x'-\mu y'-x'z'-\varrho x'|x'|^2,$$

$$\frac{dz'}{dt'}=-\beta z'+|x'|^2. \tag{6}$$

Here

$$\nu=\frac{2r_2\sigma+\delta(\sigma-1)}{2\eta}, \quad \mu=\frac{1+\sigma}{\sqrt{\eta}},$$

$$\varrho=\frac{\sqrt{\eta}}{2\sigma-b}, \quad \beta=\frac{b}{\sqrt{\eta}}. \tag{7}$$

The Jacobian of the substitution (5) is equal to $|a|^4\sigma^3/\eta^5$. Consequently, for $\eta>0$ and $b<2\sigma$ the change of variables (5) specifies a one-to-one continuous map (diffeomorphism) of the phase space of the system (1) $\mathcal{H}$ onto the phase space of the system (6) $\mathcal{H}'$ and vector fields, specified by the systems (1) and (6), which are topologically equivalent. In other words, when the conditions imposed on the values of the parameters are satisfied, the dynamics of the system (6) is equivalent to the dynamics of the complex Lorenz model.

If we ignore the term $\varrho x|x|^2$ in (6), we obtain a complex generalization of the Shimizu–Morioka equations, which have been investigated in detail in Ref. 20. It has been shown[21,22] that in a certain approximation the Shimizu–Morioka equations are also the normal form of the equations describing the chaotic dynamics near a bifurcation point with a triple zero eigenvalue having geometric multiplicity 2. We note that in (6) all the coefficients are real when $\nu=0$, i.e., when condition (33) holds (see below).

We consider the map Π: $\mathcal{H}(x,y,z)\to\mathcal{P}$, where $\mathcal{P}$ is a projective space with Cartesian coordinates u,v,w,z:

$$u=(|x|^2-|y|^2)/2, \quad v=\mathrm{Re}(x^*y), \quad w=\mathrm{Im}(x^*y). \tag{8}$$

The idea of using this map is based on the analogy of the complex Lorenz model with a two-level quantum-mechanical system. If we regard the variables x and y as components of a Schrödinger state vector, we can express the corresponding density matrix in terms of a linear combination of Pauli matrices with u, v, and w as the coefficients of this expansion. Note that

$$|x|^2=R+u, \quad |y|^2=R-u, \quad x^*y=v+iw, \tag{9}$$

where

$$R=(u^2+v^2+w^2)^{1/2}=(|x|^2+|y|^2)/2. \tag{10}$$

It is evident from Eqs. (9) that the variables u, v, w, and z contain the sum-total of information about the state of the system in the sense discussed in the Introduction. The map Π associates with each point of $\mathcal{H}$ that differ only by the common phase of x and y the map Π associates the same point in $\mathcal{P}$, whereas the images of states that differ by the amplitudes and/or the phase difference of x and y differ. We shall borrow the term *ray* from geometry[23] to designate the set of points of $\mathcal{H}$ corresponding to the same point of $\mathcal{P}$, and we shall refer to $\mathcal{P}$ as ''ray space.''

Differentiating Eqs. (8) with respect to the time and making use of Eqs. (1), we obtain the equations of motion in ray space

$$\begin{aligned}
&\dot{u}=-(\sigma+1)u+(\sigma-r_1+z)v-r_2w-(\sigma-1)R,\\
&\dot{v}=-(\sigma+1)v-\delta w-(\sigma-r_1+z)u+(\sigma+r_1-z)R,\\
&\dot{w}=-(\sigma+1)w+\delta v+r_2(R+u), \quad \dot{z}=-bz+v.
\end{aligned} \tag{11}$$

Accordingly, the image of system (6) under the map Π: $\mathcal{H}'(x',y',z')\to\mathcal{P}'$, where $\mathcal{P}'$ is a projective space equivalent to $\mathcal{P}$, is the system of equations

$$\begin{aligned}
&\dot{u}'=v'+\mu(R'-u')-\nu w'-v'(1-z'-\rho(R'+u')),\\
&\dot{v}'=-\mu v'+R'-u'+(R'+u')(1-z'-\rho(R'+u')),\\
&\dot{w}'=-\mu w'+\nu(R'+u'), \quad \dot{z}'=-\beta z'+(R'+u').
\end{aligned} \tag{12}$$

We note that in phase spaces $\mathcal{H}$ and $\mathcal{H}'$ the sets $\mathcal{Z}$ and $\mathcal{Z}'$ of points on the z and z' axes are invariant under the flows specified by the systems of equations (1) and (6), respectively. The same is true of the corresponding point sets in spaces $\mathcal{P}$ and $\mathcal{P}'$. It follows, therefore, that the point sets $\mathcal{H}/Z$, $\mathcal{H}'/Z'$, $\mathcal{P}/Z$, and $\mathcal{P}'/Z'$ are also invariant under the flows specified by (1) and (6).

Before using Eqs. (11) and (12) in place of (1) and (6), we need to find a way to obtain information about the motion in $\mathcal{H}$ from the solutions of Eqs. (11). We consider the relation between such characteristics of the dynamical state of the system in $\mathcal{H}$ and in $\mathcal{P}$ as the Lyapunov characteristic exponents and fractal dimensionality of an attractor. For a given trajectory $X_0(t)$ of the dynamical system $dX/dt=F(X)$ the spectrum of Lyapunov exponents Λ_i is defined as

$$\Lambda_i=\lim_{t\to\infty} t^{-1}\ln\left(\frac{|e_i(t)|}{|e_i(0)|}\right), \tag{13}$$

where $e_i(t)$ denotes the fundamental solutions of the linear system of equations

$$\frac{dY}{dt}=\left.\frac{\partial F}{\partial X}\right|_{X=X_0(t)} Y. \tag{14}$$

We introduce local coordinates ξ_i in a neighborhood of a point $X_0\in\mathcal{H}/Z$:

$$\begin{aligned}
&\xi_1=u(x,y), \quad \xi_2=v(x,y), \quad \xi_3=w(x,y),\\
&\xi_4=z, \quad \xi_5=\frac{\mathrm{Im}(\langle X_0,X\rangle)}{\langle X_0,X_0\rangle}.
\end{aligned} \tag{15}$$

From now on we use angle brackets to denote the Hermitian scalar product defined on C^2. The Jacobian of the transformation at the point $X=X_0$ is equal to $|x_0|^2+|y_0|^2>0$, so that (15) is a diffeomorphism in a certain neighborhood of X_0.

Writing Eqs. (14) for the system (1) in the local coordinates $\xi(\xi_1),\ldots,\xi_5)$ defined by (15), we obtain

$$\frac{d\xi}{dt}=\begin{pmatrix} \begin{pmatrix} & & \\ & \hat{A} & \\ & & \end{pmatrix} & \begin{matrix}0\\0\\0\\0\end{matrix} \\ 0,0,0,0 & 0 \end{pmatrix}\xi, \tag{16}$$

where $\hat{A}$ is the Jacobian matrix of the system of equations (11), evaluated at the point $\Pi(X_0(t))$

It is evident from Eq. (16) that the matrix $\hat{A}$ specifies the evolution of perturbations orthogonal to ξ_5 in $X_0(t)$, whereas perturbations along ξ_5 remain neutral. Since (15) is a diffeomorphism, it follows from Eqs. (13) and (16) that the spectrum of Lyapunov exponents for a trajectory in $\mathcal{H}/Z$ differs from the spectrum for its projection in $\mathcal{P}$ only by the presence of a single additional null exponent. In particular, this implies that if a given set in $\mathcal{H}/Z$ is an attractor with Lyapunov dimensionality D_L, its image in $\mathcal{P}$ is an attractor with Lyapunov dimensionality D_L-1. This relation is also valid for fractal dimensionalities of a limit set in $\mathcal{H}/Z$ and its projection in $\mathcal{P}/Z$. The latter result follows from the fact that every limit set in $\mathcal{H}/Z$ can be represented locally (in a neighborhood of the given ray) by the direct product of a set in $\mathcal{P}/Z$ and the ray, i.e., the set $\mathcal{R}^1$.

Another piece of physical information associated with the trajectories in $\mathcal{H}$ and, in our opinion, missing from Eqs. (11) is the relative phase of two states. A rule for comparing the phases of two states of a physical system described by a complex state vector has been introduced by Pancharatnam for the states of classical polarized light[24] and has subsequently been generalized to the case of quantum systems.[25,26] It can be stated as follows from the complex Lorenz model: Two states X_1 and X_2 are said to be in phase if the norm $\|X_1+X_2\|$ is the maximum of all possible values of the total phases of X_1 and X_2. We note that this rule can be used to compare the phases of states associated with different rays. The value of the norm for two given rays is defined as the phase of the complex number $\langle X_1,X_2\rangle$. This phase is called

Pancharatnam's phase. Drawing on the analogy between the complex Lorenz model and the Schrödinger equations, Toronov *et al.*[15] have shown that Pancharatnam's phase occurs naturally in the laser dynamics problem. We now discuss this problem briefly from the standpoint of differential geometry.

The triplet $(\mathcal{H}/Z,P/Z,\Pi)$ forms a fiber bundle ''2'' (Refs. 23 and 27), for which $\mathcal{H}/Z$ is the fiber bundle space, $\mathcal{P}/Z$ is the base, a fiber is a ray, and the structural group is $U(1)$. We note that this fiber bundle is nontrivial, i.e., the entire space $\mathcal{H}/Z$ cannot be represented as a direct product of the base and a ray. The equation

$$\xi_5=0 \tag{17}$$

defines the complexity on the fiber bundle. According to the conventional terminology of differential geometry, a curve in $\mathcal{H}$ is said to be horizontal (relative to a given complexity) if its velocity vector at every point is directed along the tangent to the surface $\xi_5=0$. The complexity defined by Eq. (17) ensures the uniqueness of a horizontal curve in $\mathcal{H}$ that is projected onto a given curve in $\mathcal{P}$ and passes through a given point. For the given type of fiber bundles and for the evolution of the state vector along a horizontal trajectory (relative to the given type of complexity) it has been shown[25,26] that Pancharatnam's phase for two states on a given trajectory $X(t)$ can be expressed in the form

$$\gamma=-\oint_{\Gamma T} A_s ds, \tag{18}$$

where

$$A_s=\mathrm{Im}(\langle X(s)|d/ds|X(s)\rangle)/\langle X(s)|X(s)\rangle, \tag{19}$$

ΓT is the closed contour in $\mathcal{H}$ formed by the segment T of the trajectory between two states and a curve Γ whose projection is geodesic in $\mathcal{P}$.

The integral in (12) has a nonzero value if the state vector of the system in $\mathcal{H}$ does not return to the initial point after transition around the closed contour ΓT in $\mathcal{P}$. This possibility reflects the nontriviality of the fiber bundle.[25,26]

We note that a transformation of the type (2) but with time-dependent ψ

$$\psi(t)=\int_0^t h(\tau)d\tau,$$

where $h(t)$ is a time function, takes (1) into the system of equations

$$\begin{aligned}
\dot{x}&=-(\sigma+ih(t))x+\sigma y,\\
\dot{y}&=-[1+i(h(t)-\delta)]y+(r-z)x,\\
\dot{z}&=-bz+\frac{1}{2}(x^*y+xy^*),
\end{aligned} \tag{20}$$

which is homeomorphic to (1) if $h(t)$ is continuous [in a laser experiment $h(t)$ is the phase difference of the reference signal used in heterodyne measurements[7] from a monochromatic signal having the frequency of the empty-cavity mode].

The horizontal curve whose image is the given trajectory in $\mathcal{P}$ is the trajectory of a dynamical system (20) with $h(t)=\mathrm{Im}(\langle X,F(X)\rangle/\langle X,X\rangle$, where $\mathbf{F}(\mathbf{X})$ is the phase velocity vector of the system of equations (1) [see definition (17)]. It follows from the connection of systems (1) and (20) by the transformation (2) that for arbitrary $h(t)$ the total lead of the common phase of variables x and y in (20) can be represented by the two-term sum

$$\gamma=\gamma_d+\gamma_g, \tag{21}$$

where

$$\gamma_d=\int_0^t[h(\tau)+\mathrm{Im}(\langle X,F(X)\rangle)/\langle X,X\rangle]d\tau \tag{22}$$

is the dynamical phase, and the geometric phase γ_g is given by Eq. (18).

We now show that when $h(t)$ is a given time function or a constant, the phase lead can be determined by solving the system (11) without reference to Eqs. (1). In fact, the dynamical phase is given by the time integral of the function

$$\frac{\mathrm{Im}(\langle X,F(X)\rangle)}{\langle X,X\rangle}=\frac{[\delta(R-u)-(\sigma+r_1-z)w]}{R},$$

expressed in terms of the coordinates of a point in $\mathcal{P}$.

To prove this statement for the geometrical part, in $\mathcal{P}$ we introduce spherical coordinates

$$u=\rho\cos\theta,\quad v=\rho\sin\theta\cos\phi,\quad w=\rho\sin\theta\sin\phi.$$

According to (8), $x=\rho^{1/2}\cos(\theta/2)\exp[i(\Theta)]$, and $y=\rho^{1/2}\sin(\theta/2)\exp[i(\Theta+\phi)]$, where Θ is the common phase. Expressing A_s in (18) in terms of the spherical coordinates, we obtain

$$\gamma_g=\oint_{\Gamma T}\sin^2(\theta/2)d\phi, \tag{23}$$

where the integral is evaluated in $\mathcal{P}$ around the contour formed by the trajectory and a geodesic.

It is evident that the right side of (23) is just equal to one-half the solid angle subtended by the contour.[27]

This discussion of the characteristics of phase evolution makes it clear that when a trajectory of the system in $\mathcal{H}$ represents the image of a limit cycle in $\mathcal{P}$, it is closed only for a special choice of coordinate system (carrier signal) defined by the function $h(t)$. This result is consistent with the following assertion deduced from the relations between the dimensionalities of the limit sets in $\mathcal{H}$ and $\mathcal{P}$: A periodic attractor in $\mathcal{P}$ must correspond to a torus in $\mathcal{H}$. The projection of the torus onto a limit cycle is but one example of how the analysis of system dynamics is simplified by the introduction of the projective space $\mathcal{P}$. In the next section we employ the representations of the complex Lorenz model in $\mathcal{P}$ (11) and in $\mathcal{P}'$ (12) to reveal some general properties of its solutions.

2. BOUNDEDNESS OF LIMIT SETS IN A HOMOCLINIC BIFURCATION

For the Lorentz model it is a well-known fact[2] that all the limit sets of trajectories in phase space are bounded by the sphere

$$x^2+y^2+(z-\sigma-r)^2-K^2(\sigma+r)^2=0,$$

where

$$K^2\geqslant\frac{1}{4}+\frac{b}{4}\max(\sigma^{-1},1).$$

We now show that this property is preserved in the complex Lorenz model, i.e., the limit sets in $\mathcal{H}$ are bounded by the hypersphere

$$|x|^2+|y|^2+(z-\sigma-r_1)^2-K^2(\sigma+r_1)^2=0 \tag{24}$$

with the same K as for the original Lorenz model.

Let us consider the one-parameter family of spheres

$$V_M\equiv|x|^2+|y|^2+(z-\sigma-r_1)^2-M^2=0, \tag{25}$$

where M is the parameter, and the time derivatives are

$$\dot{V}_M=-2\sigma|x|^2-2|y|^2-2b\left(z-\frac{\sigma+r_1}{2}\right)^2+b\frac{(\sigma+r_1)^2}{2}, \tag{26}$$

since

$$\dot{V}_M|_{V_M=0}=(F,\nabla V_M), \tag{27}$$

where F is the phase velocity vector; the trajectories on the sphere $V_M=0$ are directed into or out of the sphere if the right side of Eq. (26) has positive or negative values, respectively, on the sphere. It is evident from (26) that this function does not depend on the parameters δ and r_2. We can therefore use a result obtained[2] for the Lorenz model: The given function is positive on any sphere $V_M=0$ having a radius greater than the radius of the sphere (24).

An equation describing the bounding surface for limit sets in $\mathcal{P}$, corresponding to the hypersphere (24), can be obtained by making the substitution $|x|^2+|y|^2=2R$ in (24), which gives the equation of a spheroid

$$S:2R+(z-\sigma-r_1)^2-K^2(r_1+\sigma)^2=0. \tag{28}$$

We now consider the hypersurface in $\mathcal{P}$ specified by the equation

$$Q(u,v,w,z)\equiv u-\frac{2\sigma}{\delta}w+R=0. \tag{29}$$

In the subspace $\mathcal{R}^3$ of variables u,v,w Eq. (29) specifies a two-dimensional half cone with vertex at the origin, its symmetry axis directed along the unit vector $(\alpha,0,\beta)$, where $\alpha=[1+(2\sigma/\delta)^2]^{-1/2}$ and $\beta=(2\sigma/\delta)\alpha$. The cosine of the angle between the axis and the generatrix of the cone is equal to α. For $\delta>0$ the cone degenerates into the plane $w=0$. For $\delta<0$ the surface is situated in the region of negative w. In space $\mathcal{P}'$ the surface $Q=0$ corresponds to the hyperplane

$$w'=0.$$

Here $Q>(<)\,0$ corresponds to $w'>(<)\,0$. It is evident from Eq. (12) that for $w'=0$ we have $\dot{w}'=\nu(R'+u')=\nu|x'|^2$. The derivative $\dot{w}'$ is therefore nonnegative for $\nu>0$ and is nonpositive for negative ν. Consequently, for $\nu>(<)\,0$ trajectories on the surface are tangent to it or are directed toward the region $\mathcal{P}$, where $Q>(<)\,0$ $[w'>(<)\,0$ in $\mathcal{P}']$.

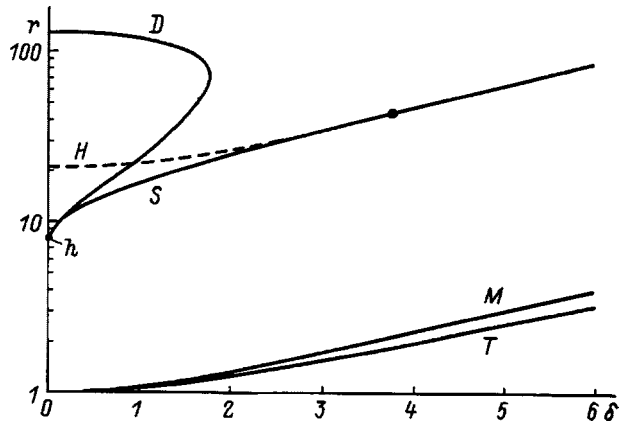

FIG. 1. Bifurcation diagram for the system of equations (11) on the plane of the parameter (δ,r); $\sigma=3$, $r_2=0$, $b=1/9$. The values of the parameter r are plotted in logarithmic scale. *T*) stability threshold of the trivial equilibrium state; *H*) Hopf bifurcation curve; *S*) saddle–node bifurcation curve; *D*) period-doubling bifurcation curve; *M*) curve $\sigma(r-1)-\delta^2/4=0$.

It follows from (12) that the surface (29) is globally stable and invariant against flow for $\nu=0$. This condition can be rewritten in the form

$$r_2=r_{2c}=\delta\frac{1-\sigma}{2\sigma}. \tag{30}$$

We now show that for $r_2>(<)r_{2c}$ every trajectory beginning in the region $Q<(>)\,0$ $[w'<(>)\,0]$ tends toward the region where $Q\geqslant(\leqslant)\,0$ $[w'\geqslant(\leqslant)\,0]$. Let $r_2>r_{2c}$ $(\nu>0)$. We consider the family of hyperplanes in $\mathcal{P}'$: $w'=C<0$. It is evident from (12) that for $\nu>0$ we have $\dot{w}'>0$ on these surfaces (since μ is always positive). Consequently, every trajectory emanating from the region $w'<0$ $(Q<0)$ intersects each of these surfaces in succession and eventually ends up on the surface $w'=0$ $(Q=0)$ or in the region $w'>0$ $(Q>0)$. More precisely, there exists a set of trajectories, having measure zero, which tends to the origin as $t\to\infty$ (see below). All other trajectories enter the region $w'>0$ $(Q>0)$ earlier or later.

To show that for $r_2<r_{2c}$ every trajectory tens to the set of points in $\mathcal{P}$ for which $Q\leqslant 0$ $(w'\leqslant 0)$, we need to analyze the family of surfaces $w'=C>0$. We omit this proof, which is easily done by the same approach as for $r_2>r_{2c}$.

We now look at some important consequences of the existence of the bounding surface $Q=0$. First of all, we note that for a laser $(r_2=0)$ all attractors are located in the region of $\mathcal{P}$ where $Q\geqslant 0$ or $\delta>0$ and in a symmetric region for $\delta<0$ (see Fig. 1). If the discussion is confined to the subspace of variables (u,v,w), the region in question is the region of the solid angle Ω subtended by the half cone $Q=0$. Consequently, for a trajectory associated with a certain attractor the solid angle subtended by the contour ΓT [see (18)] is not greater than Ω. In the limit $\delta\to\pm 0$ the cone is transformed into the flow-invariant plane $w=0$, which is globally stable in this case, so that the solid angle corresponding to ΓT tends to the limiting value $\pm 2\pi$. This result explains the resonance jump of the average slope of the phase of a laser field by an amount equal to the characteristic average frequency of the intensity fluctuations $2\pi/\tau$ (τ is the

average period of the intensity fluctuations, which coincides with the average time for the representative point to go around the origin on the plane $w=0$). It is interesting to note that this jump was first discovered in numerical calculations[13] and was interpreted as a manifestation of the geometric phase in laser dynamics on the basis of a numerical analysis of the behavior of trajectories in ray space.[16]

Another consequence of the existence of the bounding surface $Q=0$ ($w'=0$) is an additional constraint on the values of the system parameters in homoclinic bifurcation. Sensitive (nonrobust) homoclinic loops of the separatrix are known to form a very important structure responsible for the formation of a chaotic set of trajectories in the original Lorenz model.[3] Since the complex Lorenz model subsumes the original Lorenz model as a special case for $\delta=0$ and $r_2=0$, corresponding homoclinics also occur in the complex model. A necessary condition for their existence is the intersection of stable and unstable invariant manifolds of the saddle point located at the origin.[28,29]

The local structure of the invariant manifolds in the vicinity of the saddle point can be determined from a linear analysis of Eqs. (1) or (6) in the vicinity of the solution $x=y=z=0$ ($x'=y'=z'=0$).

The trivial solution of Eqs. (6) $x'=y'=z'=0$ is unstable and is a saddle point when

$$r_1>1+\frac{\delta^2-r_2\sigma+\delta r_2(1-\sigma)}{(1+\sigma)^2}.$$

Its eigenvalues are

$$\lambda_1=-\frac{\mu}{2}+\sqrt{1+\frac{\mu^2}{4}+i\nu},$$

$$\lambda_2=-\frac{\mu}{2}-\sqrt{1+\frac{\mu^2}{4}+i\nu},$$

$$\lambda_3=-\beta, \tag{31}$$

where μ, ν, and β are defined in Eq. (7).

They correspond to the eigenvectors

$$V_1=N_1^{-1}\begin{pmatrix}1\\ \lambda_1\\ 0\end{pmatrix},\quad V_2=N_2^{-1}\begin{pmatrix}1\\ \lambda_2\\ 0\end{pmatrix},\quad V_3=\begin{pmatrix}0\\ 0\\ 1\end{pmatrix},$$

where $N_{1,2}=\sqrt{1+|\lambda_{1,2}|^2}$.

It follows from Eqs. (31) that for sufficiently small ν we have $\mathrm{Re}\lambda_1>0$, $\mathrm{Re}\lambda_2<0$, and $\lambda_3<0$. Moreover, for $\nu>0$

$$\mathrm{Im}\ \lambda_1=-\mathrm{Im}\ \lambda_2>0. \tag{32}$$

We shall consider only the case $|\lambda_3|<|\mathrm{Re}\lambda_2|$, because it corresponds to the possibility of the occurrence of a Lorenz attractor.

The coordinates of points of space $\mathcal{P}'$ belonging to the unstable linear subspace $E_u=\mathrm{span}\{V_1\}$ and to the stable linear subspace $E^s=\mathrm{span}\{V_2,V_3\}$ satisfy the equations

$$E^u:\quad y'=x'\lambda_1,\quad z'=0,$$

$$E^s:\quad y'=x'\lambda_2.$$

The unstable (stable) manifold W^u (W^s) of the origin is tangent to E^u (E^s) at $x'=y'=z'=0$. Inasmuch as the z' axis is flow-invariant and belongs to W^s, points of these manifolds situated in a sphere of small radius ε satisfy the equations

$$W^u_{\mathrm{loc}}:\quad y'=x'\{\lambda_1+O(\varepsilon)\},\quad z'=O(\varepsilon^2) \tag{33}$$

and

$$W^s_{\mathrm{loc}}:\quad y'=x'\{\lambda_2+O(\varepsilon)\}. \tag{34}$$

We now consider the projections of the invariant manifolds W^u and W^s onto $\mathcal{P}'$. Since they are flow-invariant, they must map into themselves under the action of the group $U(1)$. Consequently, Π maps W^s and W^u onto (respectively) a two-dimensional manifold and a one-dimensional manifold in $\mathcal{P}'$. From Eq. (8), replacing x,y,z,w by x',y',z',w', and from Eqs. (33) and (34) we obtain

$$w'|_{\Pi(W^u_{\mathrm{loc}})}=|x'|^2\{\mathrm{Im}\ \lambda_1+O(\varepsilon)\},$$

$$w'|_{\Pi(W^s_{\mathrm{loc}})}=|x'|^2\{\mathrm{Im}\ \lambda_2+O(\varepsilon)\}. \tag{35}$$

Taking Eq. (32) into account, we conclude that for $\nu>0$ all points of $\Pi(W^u_{\mathrm{loc}})$ lie in the half space $w'\geqslant 0$. Next, we infer from the formal solution of the third equation of the system (12)

$$w'(t')=w'(0)+\nu e^{-\mu t'}\int_0^{t'}(R'(s)+u'(s))ds \tag{36}$$

that all points of $\Pi(W^u)$ lie in this half space. Indeed, since the integrand of Eq. (36) is nonnegative, $w'(t')$ is always nonnegative if $w'(0)>0$. In this regard, it follows from (35) that all points of $\Pi(W^s_{\mathrm{loc}})$ except those on the z' axis lie in the half space $w'<0$. Consequently, $\Pi(W^u)$ and $\Pi(W^s_{\mathrm{loc}})$ can intersect only on the z' axis. But this is impossible, because the z' axis does not belong to W^u. Therefore, for $\nu>0$ Eqs. (1) and (6) do not admit trajectories doubly asymptotic to the origin. Making use of the property of invariance of the system (12) under the substitution

$$\nu\to-\nu,\quad w'\to-w' \tag{37}$$

and proceeding in the same way, one can easily show that such trajectories are also nonexistent for $\nu<0$. Thus, $\nu=0$ or

$$r_2=\frac{\delta(1-\sigma)}{2\sigma}$$

is a necessary condition for the existence of a homoclinic Lorenz butterfly.

3. ONE-DIMENSIONAL MAP

Let the following relation hold for $\nu=0$:

$$k=-\frac{\lambda_3}{\lambda_1}<1,$$

where λ_1 and λ_3 are defined in Eq. (31).

This inequality corresponds to the case where the disintegration of a butterfly for $\delta=r_2=0$ is followed by the emer-

gence of a strange invariant set. In the original Lorenz model the corresponding bifurcation is described by the one-dimensional map[30]

$$\xi \to \text{sign } \xi(-\varepsilon_1 + \text{sign } A|\xi|^k),$$
$$0 < |\xi| \ll 1, \quad 0 \leqslant \varepsilon_1 \ll 1, \tag{38}$$

where ξ is a real variable, A is a separatrix variable, and ε_1 describes a small deviation from the point of homoclinic bifurcation in parameter space (we shall assume from now on that A is positive).

We wish to construct a similar map for the complex Lorenz model. Let x' and y' be complex and $\nu=0$ in Eqs. (6). Then, as shown above, a limit set of trajectories of the system belongs to the globally stable hypersurface $x'^*y' - x'y'^* = w = 0$. The solution corresponding to each trajectory on this hypersurface has the form $(x'(t')e^{i\psi}, y'(t')e^{i\psi}, z'(t'))$, where $(x'(t'), y'(t'), z'(t'))$ is the solution of the system (6) for real-valued x' and y', and ψ is a constant that depends on the initial conditions. This result enables us to write the map for complex-valued x' and y' (for $\nu=0$):

$$\xi \to e^{i \arg \xi}(-\varepsilon_1 + |\xi|^k). \tag{39}$$

Here, in contrast with (38), ξ is complex, and A is set equal to unity, which is made possible by the renormalization of ξ. The change of variables (8) transforms the homoclinic butterfly into a single homoclinic loop in the projective space $\mathcal{P}$. A one-dimensional map describing the dynamics of the system in the vicinity of this loop in the space $\mathcal{P}$ can be obtained from Eq. (39):

$$\Xi \to (-\varepsilon_1 + \Xi^{k/2})^2, \quad 0 < \Xi \ll 1, \quad 0 \leqslant \varepsilon_1 \ll 1. \tag{40}$$

Here $\Xi = |\xi|^2$. As in (39), the map (40) is valid only for $\nu=0$. For $\nu \neq 0$ we have

$$\Xi \to G(\Xi, \varepsilon_1, \nu), \quad 0 < \Xi \ll 1, \quad 0 \leqslant \varepsilon_1 \ll 1, \tag{41}$$

where $G(\Xi, \varepsilon_1, 0) = (-\varepsilon_1 + \Xi^{k/2})^2$.

Assuming that the derivative

$$G_{\nu\nu}(\Xi, \varepsilon_1, 0) = \left(\frac{\partial^2 G(\Xi, \varepsilon_1, \nu)}{\partial \nu^2} \right)_{\nu=0}$$

exists for small Ξ and small ε_1, we obtain the following from Eq. (41) for small ν:

$$\Xi \to G(\Xi, \varepsilon_1, 0) + \frac{\nu^2}{2} G_{\nu\nu}(\Xi, \varepsilon_1, 0) + O(\nu^4), \tag{42}$$

where $G(\Xi, \varepsilon_1, 0)$ is defined in Eq. (40).

By virtue of the symmetry property (37) of the system (12), Eq. (42) does not contain any terms linear or cubic in ν. Inasmuch as $\Xi, \varepsilon_1 \ll 1$, the dependence of $G_{\nu\nu}(\Xi, \varepsilon, 0)$ on Ξ and ε_1 can be disregarded. Then, omitting small terms $O(\nu^4)$ in Eq. (42), we obtain the map

$$\Xi \to (-\varepsilon_1 + \Xi^{k/2})^2 + \varepsilon_2^2,$$
$$0 < \Xi \ll 1, \quad 0 \leqslant \varepsilon_1, \varepsilon_2 \ll 1, \tag{43}$$

where $\varepsilon_2^2 = (\nu^2/2) G_{\nu\nu}(0,0,0)$.

Since the variable Ξ is nonnegative, we have $G(\Xi, \varepsilon_1, \nu) \geqslant 0$ and $G(0,0,0)=0$, so that $G_{\nu\nu}(0,0,0) \geqslant 0$. We shall assume that $G_{\nu\nu}(0,0,0) > 0$. We note that the point $\varepsilon_1 = \varepsilon_2 = 0$ corresponds to a homoclinic bifurcation of codimensionality 2, and the parameter ε_2 is proportional to the small quantity ν. The substitution of

$$\Xi = \varepsilon_1^{2/k} \left\{ 1 + \frac{2\lambda}{k^2}(1-2\zeta)\varepsilon_1^{2(1-k)/k} \right\}$$

and

$$\varepsilon_2 = \varepsilon_1^{1/k} \left\{ 1 + \frac{\lambda(2-\lambda)}{2k^2} \varepsilon_1^{2(1-k)/k} \right\} \tag{44}$$

into (43) produces the logistic map

$$\zeta \to \lambda\zeta(1-\zeta) + O(\varepsilon_1^{2(1-k)/k}).$$

Consequently, the bifurcations in the map (43) are similar to the bifurcations of the logistic map. Moreover, in a small neighborhood of a bifurcation point of codimensionality 2 ($\varepsilon_1=0$, $\varepsilon_2=0$) asymptotic expressions can be obtained for the bifurcation sets of the map (43) by substituting the bifurcation values of the parameter λ of the logistic map into Eq. (47). In particular, the first two bifurcations of the logistic map are saddle–node ($\lambda=1$) and period-doubling ($\lambda=3$) types. The asymptotic expressions for the bifurcation curves corresponding to these bifurcations on the plane of the parameters $(\varepsilon_1, \varepsilon_2)$ are

$$\varepsilon_2 = \varepsilon_1^{1/k} \left\{ 1 + \frac{1}{2k^2} \varepsilon_1^{2(1-k)/k} + O(\varepsilon_1^{4(1-k)/k}) \right\} \tag{45}$$

and

$$\varepsilon_2 = \varepsilon_1^{1/k} \left\{ 1 - \frac{3}{2k^2} \varepsilon_1^{2(1-k)/k} + O(\varepsilon_1^{4(1-k)/k}) \right\}, \tag{46}$$

respectively.

To verify the conclusions drawn using the one-dimensional map (43), we have numerically plotted several bifurcation curves for the complex Lorenz model with real r, consistent with the model of a single-mode laser. Figure 1 shows these curves on the (δ, r) plane. Inequality (4), which is the condition for replacing the system (1) by the system (6), is satisfied above curve M [$\sigma(r_1-1) - \delta^2/4 = 0$]. Curve T represents the stability threshold of the trivial steady-state solution $x'=y'=z'=0$, which becomes unstable above this curve. Point h corresponds to a homoclinic bifurcation of codimensionality 2. Because the resulting one-dimensional map is identical to the logistic map, this point must be a limit point for an infinite number of bifurcation curves. Some of these curves shown in the figure do in fact go to point h. Curve S corresponds to the saddle–node bifurcation (45). The intersection of this curve with the parametric vector from the right in projective space $\mathcal{P}$ is accompanied by the onset of two limit cycles, one stable and the other unstable. As the parameter δ is further decreased, the stable limit cycle undergoes a series of period-doubling bifurcations, making the transition to chaos. A numerical analysis shows that the corresponding curves (curve D in Fig. 1 corresponds to the

first doubling bifurcation) converge to point *h*. This result is in complete agreement with results based on an analysis of the above-derived map (43). Curve *H* in the figure corresponds to a Hopf bifurcation of a ''nonzero'' steady-state solution of the equations of motion in projective space $\mathcal{P}$. The dashed and solid segments of the curve represent its subcritical and supercritical parts, respectively. We note that the intersection of curve *H* with the *r* axis corresponds to a subcritical Hopf bifurcation, which is a well-known occurrence in the original Lorenz model.

CONCLUSION

We have shown that all the dynamical properties of a system, including the salient characteristics of its phase dynamics, can be obtained directly from the representation of the complex Lorenz model in ray space. We note that this representation does not contain singularities for certain values of the parameters, contrary to the analogous representation used in Ref. 10, and it provides an effective, simple method for studying the properties of the complex Lorenz model. We have established a correspondence between the properties of the limit sets in the initial phase space and in ray space, and we have elucidated the boundedness properties of the limit sets in these spaces. We have shown that these properties are responsible for the singular behavior — previously observed in a numerical simulation[13] — of the curve representing the mean slope of the phase of a laser field as a function of the detuning. We have proved that the homoclinic bifurcation of the separatrix of a saddle point in the complex Lorenz model has codimensionality 2. For values of the parameters close to the homoclinic bifurcation point we have constructed a one-dimensional map for points of phase space near the separatrix. For $r_2 \neq r_{2c}$ [see (30)] the resulting map (unlike the Lorenz map) is smooth and equivalent to the logistic map. This result lends support to the assertion that a ''true'' Lorenz attractor, which contains only saddle limit cycles, can exist in the complex Lorenz model only for $r_2 = r_{2c}$. We have demonstrated numerically the correspondence between the hierarchy of bifurcations in the single-mode laser model on the plane of the pump-detuning parameters and the sequence of bifurcations of the logistic map.

V. Yu. Toronov and V. L. Derbov have received support from Goskomvuz RF (State Committee on Higher Education of the Russian Federation) Grant No. 94-2.7-1097.

[1] J. D. Gibbon and M. J. McGuinness, Physica D **5**, 108 (1982).
[2] E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[3] C. Sparrow, in *The Lorenz Equations: Bifurcations, Chaos and Strange Attractors* (Springer-Verlag, Heidelberg/Berlin/New York, 1982), p. 269.
[4] N. A. Phillips, Rev. Geophys. **1**, 123 (1963).
[5] J. Pedlosky, J. Atmos. Sci. **36**, 1908 (1979); J. Pedlosky and C. Frenzen, *ibid.* **37**, 1177 (1980); J. Pedlosky, *ibid.* **40**, 1863 (1983); P. Klein and J. Pedlosky, *ibid.* **43**, 1243 (1986).
[6] A. N. Oraevskiĭ, Radiotekh. Elektron. **4**, 718 (1959).
[7] D. Y. Tang and C. O. Weiss, Phys. Rev. A **49**, 1296 (1994); D. Y. Tang, M. Y. Li, and C. O. Weiss, *ibid.* **44**, 7597 (1991); C. O. Weiss, N. B. Abraham, and J. Hubner, Phys. Rev. Lett. **61**, 1587 (1988).
[8] A. C. Fowler, J. D. Gibbon, and M. J. McGuinness, Physica D **4**, 139 (1982); J. D. Gibbon and M. J. McGuinness, *ibid.* **5**, 108 (1982); A. C. Fowler and M. J. McGuinness, SIAM J. Appl. Math. **44**, 681 (1984).
[9] H. Zeghlache and P. Mandel, J. Opt. Soc. Am. B **2**, 18 (1985).
[10] C. Ning and H. Haken, Phys. Rev. A **41**, 3826 (1990).
[11] C. Ning and H. Haken, Z. Phys. B **81**, 457 (1990).
[12] R. Vilaseca, G. J. de Valcarcel, and E. Roldan, Phys. Rev. A **41**, 5269 (1990).
[13] E. Roldan *et al.*, Phys. Rev. A **48**, 591 (1993).
[14] C. Z. Ning and H. Haken, Phys. Rev. Lett. **68**, 2109 (1992).
[15] V. Toronov and V. Derbov, Phys. Rev. A **49**, 1392 (1994).
[16] V. Toronov and V. Derbov, Phys. Rev. A **50**, 878 (1994).
[17] A. N. Oraevskiĭ and V. Yu. Toronov, Kvantovaya Élektron. **16**, 2063 (1989) [Sov. J. Quantum Electron. **19**, 1327 (1989)].
[18] V. I. Yudovich, manuscript deposited at the All-Union Institute of Scientific and Technical Information [in Russian], VINITI Deposit No. 2611-78 (Rostov-on-Don, 1978), 49 pp.
[19] T. Shimizu and N. Morioka, Phys. Lett. A **76**, 201 (1980).
[20] A. L. Shil'nikov, Selecta Math. Sov. **10**, 105 (1991); Physica D **62**, 332 (1993).
[21] A. G. Vladimirov and D. Yu. Volkov, Opt. Commun. **100**, 351 (1993).
[22] A. M. Rucklidge, Physica D **62**, 323 (1993).
[23] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1969).
[24] S. Pancharatnam, Proc.-Indian Acad. Sci., Sect. A **44**, 247 (1956).
[25] J. Samuel and R. Bhandari, Phys. Rev. Lett. **60**, 2339 (1988).
[26] J. Anandan and Y. Aharonov, Phys. Rev. Lett. **65**, 1697 (1990).
[27] B. A. Dubrovin, S. P. Novikov, and L. T. Fomenko, *Modern Geometry* [in Russian], Nauka, Moscow, 1984.
[28] L. P. Shil'nikov, Mat. Sb. **35**, 240 (1968).
[29] L. P. Shil'nikov, Usp. Mat. Nauk **36**, 240 (1981).
[30] L. P. Shil'nikov, Differ. Equations, No. 11, 180 (1981).

Translated by James S. Wood

Influence of the Barnett–London effect on the motion of a superconducting rotor in a nonuniform magnetic field

Yu. M. Urman

(Submitted April 14, 1997)
Zh. Tekh. Fiz. **68**, 10–14 (August 1998)
[S1063-7842(98)00208-6]

In a number of instruments utilizing a contactless suspension system a rotor spins at high speed in a suspension magnetic field. The design of such instruments is feasible only if the properties of bodies rotating in a magnetic field are known. This understanding is especially significant for superconducting bodies in that the rotation of a superconductor in a magnetic field produces several phenomena that can substantially undermine the performance characteristics of the instrument. Of major importance in this regard is the assessment of unwanted effects attending the rotation of a superconductor in magnetic fields. One such effect is the onset of a magnetic field in the interior of the body; this field has nothing to do with the presence of the external field and depends linearly on the angular velocity of the body.

The onset of a magnetic field in a superconductor when it rotates (analog of the Barnett effect[1]) was predicted by Becker et al.[2] and given a theoretical foundation by F. London in 1960. The magnetic field in a rotating superconductor has come to be known as the London moment. The first successful experiment to measure it was reported in Ref. 3 and has been repeated several times.[4–6] In 1940 Kikoin and Gubar' conducted a superconductor experiment analogous to the celebrated Einstein–de Haas experiment, showing that the superconducting current is an electron current, and the magnetization of the superconductor is a consequence of the screening current.

The London moment has been treated in several theoretical papers based on classical electrodynamics,[7] general relativity,[8] and quantum mechanics.[9]

In instruments utilizing a body suspended in a magnetic field the magnetization induced by rotation of the body interacts with the suspension field to generate torque. The investigation of the motion of the body under the influence of this torque is of fundamental importance, particularly in establishing the ultimate error limits of navigation systems that utilize a contactless suspension. Indeed, even if all the factors causing a gyroscope to drift were eliminated, it would still be virtually impossible to avoid the torque induced by the Barnett–London effect. Moreover, the London moment, which is directed strictly along the axis of rotation of a symmetric rotor, can be exploited as an angle sensor, for example, in a cryogenic gyroscope utilizing an electrostatic suspension.

The influence of the Barnett–London effect on the angular motions of a rigid body has been investigated from the general theoretical point of view[10–13] and in application to a superconducting gyroscope.[14,15] All these papers (except Ref. 14) have directed their attention mainly to the influence of the Barnett–London effect on the angular motions of a superconducting body in a uniform magnetic field. However, the behavior of a body in a nonuniform field is of major concern, because a mechanical force also appears in such a field together with the mechanical torque, so that the translational and rotational motions become mutually coupled, inducing a series of phenomena that affect the operation of the cryogenic gyroscope; the process is analogous to an unbalanced rotor in a contactless suspension.[16]

1. TORQUE ACTING ON A ROTATING SUPERCONDUCTING SPHERE IN A NONUNIFORM MAGNETIC FIELD

The magnetic torque on a superconducting solid sphere rotating with angular velocity $\mathbf{\Omega}$ is given by the expression[15]

$$\mathbf{G}=\varkappa a^3\mathbf{\Omega}, \quad \varkappa=\frac{m_e C}{|e|}, \tag{1.1}$$

where m_e is the electron mass, C is the speed of light in vacuum, $|e|$ is the magnitude of the electron charge, and a is the radius of the sphere.

Next we determine the surface current corresponding to the London magnetic moment.

The vector potential at an arbitrary point $\mathbf{r}$ is expressed in terms of the magnetic torque per unit volume according to the equation

$$\mathbf{A}(\mathbf{r})=\int_v \mathbf{G}\times\nabla'\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)dV'. \tag{1.2}$$

Invoking the identity $\operatorname{curl}(\varphi\mathbf{G})=\nabla\varphi\times\mathbf{G}]+\varphi\operatorname{curl}\mathbf{G}$, we transform Eq. (1.2) as follows:

$$\mathbf{A}(\mathbf{r})=\int_v \frac{\operatorname{curl}'\mathbf{G}}{|\mathbf{r}-\mathbf{r}'|}dV'-\int_{v'}\operatorname{curl}\left(\frac{\mathbf{G}}{|\mathbf{r}-\mathbf{r}'|}\right)dV'. \tag{1.3}$$

The prime signifies integration over points in the interior of the body. Since $\mathbf{G}=\text{const}$ inside the sphere, $\operatorname{curl}\mathbf{G}$ is equal to zero everywhere, so that the first integral in Eq. (1.3) vanishes, and the second integral can be converted into a surface integral. We then have

$$\mathbf{A}(\mathbf{r})=\int_{s'}\frac{[\mathbf{G}\times\mathbf{n}]}{|\mathbf{r}-\mathbf{r}'|}ds',$$

where $\mathbf{n}$ is the unit vector normal to the surface of the sphere.

The quantity $C[\mathbf{G}\times\mathbf{n}]$ can be regarded as the surface density of the current. In our case it is equal to

$$\mathbf{j}_L=\frac{3C}{4\pi}\varkappa[\mathbf{\Omega}\times\mathbf{n}], \tag{1.4}$$

where from now on the subscript L attached to the vector $\mathbf{j}$ signifies the London current.

We consider a spherical superconducting rotor spinning in a magnetic suspension formed by N superconducting coils. The vector potential of the system is

$$\mathbf{A}=\frac{1}{C}\sum_{k=1}^{N}\int_{v_k}\frac{\boldsymbol{\delta}_k}{|\mathbf{r}-\mathbf{r}_k|}dV_k+\frac{1}{C}\times\int_{s'}\frac{\mathbf{j}_s}{|\mathbf{r}-\mathbf{r}'|}ds'+\frac{3}{4\pi}\varkappa\int_{s'}\frac{\mathbf{\Omega}\times\mathbf{n}}{|\mathbf{r}-\mathbf{r}'|}ds'. \tag{1.5}$$

Here the first term describes the vector potential of the specified current sources δ_k, the second term represents the potential of the supercurrent $\mathbf{j}_s$ induced by the external field of the sources, and, finally, the third term gives the potential of the London currents. Calculating the potential of the London currents, we obtain

$$\mathbf{A}(\mathbf{r})=\varkappa\frac{a^3}{r^3}[\mathbf{\Omega}\times\mathbf{r}]. \tag{1.6}$$

Using the vector potential (1.5), we find, first, the flux linkage with the kth coil

$$\psi_k=C\sum_{j=1}^{N}I_jL_{jk}+\frac{W_k}{S_k}\varkappa a^3\mathbf{\Omega}\cdot\int_{v_k}\frac{[\mathbf{r}\times\mathbf{t}_k]}{r_k^3}dV_k, \tag{1.7}$$

where L_{jk} is the mutual inductance coefficient of the th and jth coils, I_j is the total current in the jth coil, and W_k and S_k are the number of ampere-turns and the cross-sectional area of the wire in the kth coil; second, we find the magnetic energy

$$U=\frac{1}{2}\sum_{kj}I_kI_jI_{kj}+\frac{1}{C}\varkappa a^3\mathbf{\Omega}\cdot\sum_k\frac{W_k}{S_k}\times\int_{v_k}\frac{[\mathbf{r}\times\mathbf{t}_k]}{r_k^3}dV_k+\frac{3}{8}\varkappa^2a^{-1}\int_s[\mathbf{\Omega}\times\mathbf{r}]^2ds. \tag{1.8}$$

The first term in Eq. (1.8) describes the magnetic energy of interaction of the suspension field with the nonrotating sphere, the second term gives the interaction of the London currents with the suspension field, and the third term represents the self-energy of the London currents. Making use of the fact that

$$\frac{1}{C}\sum_k I_k\frac{W_k}{S_k}\int_{v_k}\frac{[\mathbf{r}\times\mathbf{t}_k]}{r_k^3}dV_k=-\mathbf{H} \tag{1.9}$$

represents the magnetic field generated by the coils at the point in space occupied by the center of the sphere, we can write the magnetic energy of interaction of the London currents with the external field of the suspension in the form

$$U_L=-\varkappa a^3(\mathbf{\Omega}\cdot\mathbf{H}). \tag{1.10}$$

We use Eq. (1.10) to calculate the torques acting on a rotating superconducting sphere in various coil-suspension configurations.

1) Rotating superconducting sphere in the magnetic field of a single coil carrying a circular current

a) $I=\text{const}$

Once the integral (1.8) has been evaluated, the magnetic energy has the form

$$U=\frac{1}{2}I^2L_{11}+\frac{2\pi IW}{C}\varkappa a^3\frac{\sin^2\vartheta}{r}(\mathbf{\Omega}\cdot\mathbf{e}_z), \tag{1.11}$$

where r is the radius of the spherical surface defined by the path of the circular current (current loop), ϑ is half the angle at which the current loop is viewed from the center of the superconducting sphere, $\mathbf{e}_z$ is the unit vector along the axis of the coil, and L_{11} is the inductance.

The torque on a rotating superconducting sphere is given by the equation

$$\mathbf{M}=\frac{2\pi IW}{Cr}\varkappa a^3\sin^2\vartheta[\mathbf{\Omega}\cdot\mathbf{e}_z]. \tag{1.12}$$

It follows from this equation that the torque is perpendicular to the plane formed by the vectors $\mathbf{\Omega}$ and $\mathbf{e}_z$ and depends on the magnitude of the angular velocity, the radius of the sphere, the position of the coil, and the number of ampere-turns W. The maximum torque is attained when the current loop lies in the equatorial plane, $\vartheta=\pi/2_+$, and $\mathbf{\Omega}\perp\mathbf{e}_z$.

b) $\psi=\text{const}$

Using Eq. (1.7), we find the flux linkage with the current loop

$$\psi_0=IL_{11}+\frac{2\pi W}{Cr}\varkappa a^3\sin^2\vartheta(\mathbf{\Omega}\cdot\mathbf{e}_z),$$

so that

$$I=(\psi_0-\psi_L)L_{11}^{-1}. \tag{1.13}$$

Substituting Eq. (1.13) into (1.12), we have

$$\mathbf{M}=\frac{2\pi W}{CrL_{11}}(\psi_0-\psi_L)\varkappa a^3\sin^2\vartheta[\mathbf{\Omega}\times\mathbf{e}_z]. \tag{1.14}$$

In general, the torque at constant flux through the coil differs from the torque at constant current in the coil, because the current in the coil depends on the position of $\mathbf{\Omega}$ relative to the axis of the coil.

2) Rotating superconducting sphere in the field of two coaxial coils with opposing circular currents

The magnetic energy interaction of the rotating sphere with the field of two current-carrying coils has the form

$$U=\frac{1}{2}I_1^2L_{11}+\frac{1}{2}I_2^2L_{22}+I_1I_2L_{12}+\frac{1}{C}\varkappa a^3\mathbf{\Omega}\cdot\sum_{k=1}^{2}I_kW_k\int\frac{[\mathbf{e}_r\times d\mathbf{l}]}{r_k^2}. \tag{1.15}$$

If the center of the sphere coincides with the center of symmetry of two coaxial coils having equal geometries and carrying equal but opposing currents, a torque is not induced by the London currents, because the last term in Eq. (1.15) vanishes. We therefore assume at the outset that the center of

the sphere does not coincide with the center of symmetry of the suspension, but is located at a point O_1 characterized by a radius vector **b** drawn from the symmetry center of the suspension O; we adopt this point as the origin. The z axis of this system coincides with the symmetry axis of the suspension, and the x and y axes lie in the plane perpendicular to the z axis.

Calculating the energy to within first-order terms in the displacement of the center of mass of the rotor, we have

$$U=\frac{1}{2}I^2(L_{11}+L_{22}-2L_{12})+\frac{2\pi IW}{Cr^2}\varkappa a^3 3 \times\sin^2\vartheta\cos\vartheta[2z(\mathbf{\Omega}\cdot\hat{e}_3)-x(\mathbf{\Omega}\cdot\hat{e}_1)-y(\mathbf{\Omega}\cdot\hat{e}_2)], \tag{1.16}$$

where x,y,z are the components of the displacement vector **b**, and $\hat{e}_i$ $(i=1,2,3)$ are the corresponding unit vectors.

The torque has the form

$$\mathbf{M}=\frac{2\pi IW}{Cr^2}\varkappa a^3 3\sin^2\vartheta\cos\theta\{2z[\mathbf{\Omega}\times\mathbf{e}_3]-x[\mathbf{\Omega}\times\mathbf{e}_1]-y[\mathbf{\Omega}\times\mathbf{e}_2]\}. \tag{1.17}$$

It is readily apparent that the rotating sphere is acted upon not only by the torque, but also by a force expressed in the form

$$\mathbf{F}=\frac{2\pi IW}{Cr}\varkappa a^3 3\sin^2\theta\cos\vartheta\{2\Omega_3\hat{e}_3-\Omega_1\hat{e}_1-\Omega_2\hat{e}_2\}. \tag{1.18}$$

It follows from Eq. (1.18) that the force vector is not in the same direction as the angular velocity and does not depend on the displacement of the rotor.

A general analysis shows that the torque induced by the Barnett–London effect occurs when the expansion of the potential of the suspension field contains the first harmonic (a uniform component of the field exists). Consequently, it looks as though to avoid drift errors induced by the London currents, one would have to endow the suspension magnetic field with a configuration such that a uniform component did not exist in the field in the vicinity of the stable equilibrium state of the body. For example, the body could be placed at the center of the system formed by coaxial coils carrying equal but opposing currents. Equations (1.17) and (1.18) indicate that this is not the case. In fact, if the rotation of the sphere without gravitational forces did not alter the equilibrium position, the torque acting on the spinning rotor would be equal to zero. However, the rotation of the sphere generates the force (1.18) [which in general can be written in the form $F_i=\varkappa a^3\Omega_j(\partial H_i/\partial X_j]$, which displaces the rotor to a new equilibrium position dictated by the condition that the forces created by the suspension field is equal to the force induced by the London effect. This displacement induces the torque (1.17). If the rotor nutates, coupling between the translational and rotational motions can set in. In a contactless, adjustable suspension such coupling can produce a number of effects that influence the operation of the instrument in much the same way as rotor disbalance in a contactless suspension.

2. INFLUENCE OF THE BARNETT–LONDON EFFECT ON THE MOTION OF A SUPERCONDUCTING ROTOR IN THE NONUNIFORM MAGNETIC FIELD OF A CONTACTLESS, ADJUSTABLE CRYOGENIC SUSPENSION

Separating out the dimensionless parameter ε, we write the rotor equations of motion in a dimensionless form including the momentum equation and the angular momentum equation:

$$D^2\mathbf{r}+\hat{\omega}(D)\mathbf{r}=(\mathbf{\Omega}\cdot\nabla)\mathbf{h},\quad D\mathbf{k}=\varepsilon[\mathbf{\Omega}\times(\mathbf{r}\cdot\nabla)\mathbf{h}]. \tag{2.1}$$

Here **r**, **k**, **Ω**, and **h** are dimensionless vectors representing the displacement of the center of mass relative to the center of the suspension, the angular momentum, the angular velocity, and the magnetic field; $\omega(D)$ is the tensor operator of dynamic stiffness; D is the operator of differentiation with respect to dimensionl ess time, the scale of which is chosen equal to the ''nutational'' transit time of the rotor $T_0=I_1K_0^{-1}$ (I_1 is the equatorial moment of inertia, and K_0 is a characteristic angular momentum); and $\varepsilon=R_0^2MT^{-2}\sim10^{-8}-10^{-3}$ (R_0 is a characteristic scale of the displacement of the center of mass, and M is the mass of the rotor).

Inasmuch as $\operatorname{curl}\mathbf{H}=0$ and $\operatorname{div}\mathbf{H}=0$, the tensor $\partial H_i/\partial x_j$ is symmetric with zero trace; it can therefore be written in the form of the irreducible second-rank tensor

$$\mathfrak{L}_{2\mu}\left\{\mathfrak{L}_{20}=\partial H_3/\partial x_3,\mathfrak{L}_{22}=\mathfrak{L}_{2-2}=\frac{1}{\sqrt{6}}\left(\frac{\partial H_1}{\partial x_1}-\frac{\partial H_2}{\partial x_2}\right)\right\}.$$

If the field is axisymmetric, the tensor $\mathfrak{L}_{2\mu}$ has only the one component $\mathfrak{L}_{20}$. We write the London force in terms of an irreducible first-rank tensor[17]:

$$F_1=-\sqrt{\frac{5}{2}}\varkappa\{\mathfrak{L}_2\oplus\Omega_1\}_1. \tag{2.2}$$

For subsequent calculations we write the second equation (2.1) in the form of the irreducible tensor

$$k_1=i\sqrt{5}\varepsilon\{\Omega_1\oplus\{\mathfrak{L}_2\oplus r_1\}_1\}_1=M_1. \tag{2.3}$$

We determine **r** from the first equation (2.1), assuming that the eigenfunctions of the problem are damped:

$$r_1=-\sqrt{\frac{5}{2}}\{\mathfrak{L}_2\oplus E\Omega_1\}_1+\frac{5}{2}\{T_2\oplus\{\mathfrak{L}_2\oplus\Omega_1\}\}_1, \tag{2.4}$$

where

$$E=\frac{1}{3}\operatorname{Tr}\hat{\omega}(D),\quad T_{20}=\hat{\omega}_3-E,$$

$$T_{22}=T_{2-2}=\sqrt{\frac{1}{6}}(\hat{\omega}_1-\hat{\omega}_2).$$

Substituting Eq. (2.4) into the right side of (2.3), we obtain

$$M_{1\nu}=-i\frac{5}{\sqrt{2}}\sum_{\lambda npSq'\mu}\mathfrak{L}_{2q'}\mathfrak{L}_{2q}\Omega_{1\lambda}E(D)\Omega_{1S}C^{1\nu}_{1\lambda 1n}\times C^{1n}_{2q'1\mu}C^{1\mu}_{2p1S}+i\frac{5\sqrt{5}}{2}\sum_{\lambda npSqq'\mu}\mathfrak{L}_{2q'}\mathfrak{L}_{2q}\times\Omega_{1\lambda}T_{2p}\Omega_{1t}C^{1\mu}_{2p1S}C^{1\nu}_{2q'1\mu}C^{1S}_{2q1t}. \quad (2.5)$$

Here $C^{1\nu}_{1\lambda 1n}$ is the Clebsch–Gordan coefficient.[17]

Below we consider the axisymmetric case, i.e., we assume that $\mathfrak{L}_{20}=0$, $\mathfrak{L}_{22}=\mathfrak{L}_{2-2}=0$, $T_{20}=0$, and $T_{22}=T_{2-2}=0$. Equation (2.5) is now simplified considerably, acquiring the form

$$M_{1\nu}=\frac{i}{2\sqrt{2}}\sum_{\lambda n}(3n^2-2)\Omega_{1\lambda}\times\left[E(D)-\frac{1}{2}(3n^2-2)T_{20}(D)\right]\Omega_{1n}C^{1\nu}_{1\lambda 1n}. \quad (2.6)$$

We average (2.6) over the free Euler–Poinsot motion and write the torque in the form

$$\mathbf{M}=[\hat{e}\nabla]V_1+\nabla_\Omega V_2+\hat{e}V_3, \quad (2.7)$$

where

$$V_1=\frac{\varepsilon}{8}\mathfrak{L}^2_{20}\left(1-\frac{\chi}{1+\chi}\cos^2\beta\right)^2(\mathbf{k}\cdot\mathbf{h})^2\times[4\hat{\omega}_3(0)-\hat{\omega}_1(0)]+\frac{\varepsilon}{16}\mathfrak{L}^2_{20}\left(\frac{\chi}{1+\chi}\right)^2\times\sin^2\beta\cos^2\beta(\mathbf{k}\cdot\mathbf{h})^2\mathrm{Re}[\hat{\omega}_1(ik)-4\hat{\omega}_3(ik)],$$

$$V_2=\frac{\varepsilon}{16}\mathfrak{L}^2_{20}\left(\frac{\chi}{1+\chi}\right)^2\sin^2\beta\cos^2\beta(\mathbf{k}\cdot\mathbf{h})^2\times\mathrm{Im}[\hat{\omega}_1(ik)-4\hat{\omega}_3(ik)],$$

$$V_3=\frac{\varepsilon}{4}k^2\mathfrak{L}^2_{20}\left(\frac{\chi}{1+\chi}\right)^2\sin^2\beta\cos^2\beta\times\left\{2[\mathbf{k}\times\mathbf{h}]\mathrm{Im}\,\hat{\omega}_3(ik)+\frac{1}{2}(1+(\mathbf{k}\cdot\mathbf{h})^2)\mathrm{Im}\,\hat{\omega}_1(ik)\right\}.$$

Here **e** is the unit angular momentum vector, ∇_Ω denotes the transverse part of the gradient, $\chi=(I_3-I_1)/I_1$, and β is the nutation angle (the angle between the vector **k** and the dynamical symmetry axis of the rotor). Also determining the torque in projection onto the symmetry axis of the body and averaging, we write the equations of evolutionary motions of the angular momentum vector and the nutation angle:

$$\dot{k}=\frac{\varepsilon}{4}k^2\mathfrak{L}^2_{20}\left(\frac{\chi}{1+\chi}\right)^2\sin^2\beta\cos^2\beta\times\left[2\sin^2\rho\,\mathrm{Im}\,\hat{\omega}_3(ik)+\frac{1}{2}(1+\cos^2\rho)\mathrm{Im}\,\hat{\omega}_1(ik)\right],$$

$$\dot{\rho}=-\frac{\varepsilon}{8}k\mathfrak{L}^2_{20}\left(\frac{\chi}{1+\chi}\right)^2\sin^2\beta\cos^2\beta\sin\rho\times\cos\rho\,\mathrm{Im}[\hat{\omega}_1(ik)-4\omega_3(ik)],$$

$$\sin\rho\dot{\sigma}=-\frac{\varepsilon}{4}\mathfrak{L}^2_{20}k\left(1-\frac{\chi}{1+\chi}\cos^2\beta\right)\sin\rho\cos\rho[4\omega_3(0)-\omega_1(0)]-\frac{\varepsilon}{8}k\mathfrak{L}^2_{20}\left(\frac{\chi}{1+\chi}\right)^2\sin^2\beta\cos^2\beta\sin\rho\times\cos\rho\,\mathrm{Re}[\omega_1(ik)-4\omega_3(ik)],$$

$$\dot{\beta}=-\frac{\varepsilon}{4}k\mathfrak{L}^2_{20}\frac{\chi}{1+\chi}\left(1-\frac{\chi}{1+\chi}\cos^2\beta\right)\sin\beta\cos\beta\times\left[2\sin^2\rho\,\mathrm{Im}\,\hat{\omega}_3(ik)+\frac{1}{2}(1+\cos^2\rho)\mathrm{Im}\,\omega_1(ik)\right]. \quad (2.8)$$

Here σ is the angle of precession of the angular momentum vector about the field axis, and ρ is the angle between the angular momentum vector and the field axis. Using the first and last equations of the system (2.8), we find the first integral

$$k^2\left(1-\frac{\chi}{1+\chi}\cos^2\beta\right)=\mathrm{const}, \quad (2.9)$$

which corresponds in the physical sense to the conservation of kinetic energy. If the suspension is isotropic ($\omega_3=\omega_1$), we have one more integral

$$k\cos\rho^3\sqrt{Ctg^2\rho}=\mathrm{const}. \quad (2.10)$$

From the last equation of the system (2.8) we deduce a sufficient condition for asymptotic stability of the nutational oscillations of the body:

$$\chi\left[2\sin^2\rho\,\mathrm{Im}\,\omega_3(ik)+\frac{1}{2}(1+\cos^2\rho)\mathrm{Im}\,\omega_1(ik)\right]>0. \quad (2.11)$$

For an isotropic suspension, condition (2.11) goes over to the inequality

$$\chi\,\mathrm{Im}\,\omega(ik)>0. \quad (2.12)$$

We now consider the case in which a superconductor in a magnetic suspension is acted upon by an elastic force proportional to the displacement of the geometrical center of the sphere $q_0\mathbf{r}$, and the viscous friction force is proportional to $\lambda\mathbf{r}$. The expression for the transfer function assumes the form

$$\omega(0)=q_0(1+\tau D),$$

whereupon

$$\omega(ik)=\frac{1}{(q_0-k^2)+iq_0\tau k}=\frac{(q_0-k^2)^2-i\tau q_0 k}{(q_0-k^2)^2+q_0^2\tau^2k^2},$$

and, accordingly,

$$\mathrm{Im}\,\omega(ik) = -\frac{\tau q_0 k}{(q_0 - k^2)^2 + q_0^2 \tau^2 k^2} < 0.$$

To satisfy inequality (2.12), it is necessary that $\chi = (I_3 - I_1)/I_1 < 0$, i.e., $I_3 < I_1$. In the given situation, therefore, nutations decay for a prolate body and grow for an oblate body.

The existence of the integrals enables us to analyze the motion of the angular momentum vector. The steady-state motion is precession of the angular momentum vector about the axis of the suspension at the precession velocity

$$\dot{\sigma} = -\frac{\varepsilon}{4} k \mathfrak{Q}_{20}^2 \frac{\cos\rho_0}{(1+\chi)^2}[4\omega_3(0) - \omega_1(0)]. \tag{2.13}$$

It follows from Eq. (2.13) that the London currents can cause the gyroscope to drift, even in an equal-stiffness suspension.

We now investigate the motion of the body in an isotropic suspension on the basis of the first integrals (2.9) and (2.10). Let the initial steady state correspond to the values of the variables

$$\beta = 0, \quad k = k_0, \quad \rho = \rho_0,$$

$$\dot{\sigma} = -\tfrac{3}{4}\varepsilon k_0 \mathfrak{Q}_{20}^2 (1+\chi)^{-2} \cos\rho_0 \omega(0).$$

Let us assume that nutations set in. It follows from the presence of the integrals that the decay of the nutations is followed by a new steady state characterized by the variables $\beta = 0$, $k_H = \sqrt{I_3/I_1}k_0$, and $\rho = \rho_H$. Since $I_3 < I_1$, in the new steady state we have $k_H < k_0$, i.e., the rotor slows down, and the angle between the field axis and the angular momentum decreases. Consequently, not only the precession velocity, but also the angle of cone of precession varies.

This work has received financial support from the Russian Fund for Fundamental Research (Project No. 95-01-00612).

[1] S. I. Barnett, Rev. Mod. Phys. **7**, 129 (1935).
[2] R. Becker, F. Sauter, and G. Holler, Z. Phys. **85**, 77 (1935).
[3] A. F. Hildebrandt, Phys. Rev. Lett. **12**, 190 (1964).
[4] M. Bol and W. M. Fairbank, Low Temp. Phys. **9**, 471 (1964).
[5] N. F. Brinkman, Phys. Rev. **184**, 460 (1969).
[6] I. B. Hendrichs, C. A. King, and H. E. Roschah, Low Temp. Phys. **4**, 202 (1971).
[7] V. V. Borisov, Zh. Tekh. Fiz. **45**, 2257 (1975) [Sov. Phys. Tech. Phys. **20**, 1413 (1975)].
[8] G. Parini, Nuovo Cimento **14**, 66 (1966).
[9] B. I. Verkin and N. O. Kulik, Zh. Éksp. Teor. Fiz. **61**, 2067 (1971) [Sov. Phys. JETP **34**, 1103 (1971)].
[10] A. A. Burov and G. I. Subkhankurov, Prikl. Mat. Mekh. **50**, 960 (1986).
[11] N. E. Egarmin, in *Aerophysics and Geocosmic Research* [in Russian], Izd. MOTI, Moscow (1983), pp. 95–96.
[12] V. V. Kozlov, Izv. Akad. Nauk SSSR Mekh. Tverd. Tela, No. 6, 95 (1985).
[13] V. A. Samsonov, Izv. Akad. Nauk SSSR Mekh. Tverd. Tela, No. 4, 32 (1984).
[14] Yu. M. Urman, Dokl. Akad. Nauk SSSR **276**, 1402 (1984) [Sov. Phys. Dokl. **29**, 512 (1984)].
[15] Yu. M. Urman, Zh. Tekh. Fiz. **56**, 2081 (1986) [Sov. Phys. Tech. Phys. **31**, 1251 (1986)].
[16] Yu. T. Martynenko, *Rigid Body Motion Body in Electric and Magnetic fields* [in Russian], Nauka, Moscow (1988), 386 pp.
[17] Yu. M. Urman, in *All-Republic Joint Collection*, No. 15 (Naukova Dumka, Kiev, 1983), pp. 75–87.

Translated by James S. Wood

Kinetics of processes in the middle atmosphere during the laser excitation of O_2 molecules

A. M. Starik and O. V. Taranov

(Submitted July 8, 1996; resubmitted October 2, 1997)
Zh. Tekh. Fiz. **68**, 15–23 (August 1998)

Kinetic processes taking place in the atomic–molecular system $O–O_2–O_3$ in the middle atmosphere with the participation of oxygen molecules in the excited electronic states $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ are analyzed in detail. The possibility of increased ozone production under the influence of solar radiation during the laser excitation of O_2 molecules in the $a^1\Delta_g$ state is demonstrated on the basis of numerical modeling. Upper and lower bounds are determined for the densities of $O_2(a^1\Delta_g)$ molecules at which the ozone concentration increases in the irradiated zone. © *1998 American Institute of Physics.* [S1063-7842(98)00308-0]

INTRODUCTION

The possibility of an increase in the ozone production rate in the stratosphere during the photodissociation of excited O_2 molecules has been discussed at some length in recent times.[1–4] The increased rate is caused by the accelerated formation of atomic oxygen in the photodissociation of O_2 molecules either from excited electronic states $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ (Ref. 1) or from highly excited vibrational levels of the electronic ground state.[2,3] In neither case, however, has the ozone production kinetics been analyzed in detail. The problem is that the presence of excited O_2 molecules also tends to intensify collisional O_3 fragmentation processes. Consequently, an increase in the rate of ozone production during excitation of, for example, the $O_2(a^1\Delta_g)$ state is attainable only when certain conditions are observed. It should be noted that excited $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules play an important role in the upper and lower stratosphere. They determine the intensity of atmospheric emissions at wavelengths of 1.27 μm and 762 nm and can contribute significantly to O_3 production at altitudes above 45 km.[5,6] Interest in research on kinetic processes in the atmosphere involving $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ is also prompted by the need to analyze the propagation of laser radiation subjected to resonance absorption in electronic–vibration transitions of the O_2 molecule.[3,4]

The specific objective of the present study is to analyze the kinetics of ozone production and destruction in the atmosphere during the laser excitation of O_2 molecules in the $a^1\Delta_g$ state.

KINETIC PROCESSES INVOLVING $O_2(a^1\Delta_g)$ AND $O_2(b^1\Sigma_g^+)$

The main source of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ molecules in the stratosphere is the photodissociation of O_3 (Ref. 6). It is customary to single out the short-wavelength part of the spectrum with $\lambda<198$ nm, the Hartley bands with $\lambda=198-310$ nm, the Huggins bands with $\lambda=310-360$ nm, and the Chappuis band with $\lambda=410-730$ nm (Refs. 6 and 7). An analysis[8,9] has shown that the following processes take place in these spectral intervals:

$$O_2+h\nu(\lambda<198\text{ nm})\rightarrow\text{products} \quad J_{\mathrm{I}}, \quad \text{(I)}$$
$$\rightarrow O_2(b^1\Sigma_g^+)+O(^1D) \quad J_{\mathrm{I}'}, \quad \text{(I}')$$
$$\rightarrow O_2(a^1\Delta_g)+O(^1S) \quad J_{\mathrm{I}''}, \quad \text{(I}'')$$
$$O_3+h\nu$$
$$(198\leqslant\lambda\leqslant310\text{ nm})\rightarrow\text{products} \quad J_{\mathrm{II}}, \quad \text{(II)}$$
$$\rightarrow O_2(b^1\Sigma_g^+)+O(^3P) \quad J_{\mathrm{II}'}, \quad \text{(II}')$$
$$\rightarrow O_2(a^1\Delta_g)+O(^1D) \quad J_{\mathrm{II}''}, \quad \text{(II}'')$$
$$O_3+h\nu$$
$$(310<\lambda\leqslant360\text{ nm})\rightarrow\text{products} \quad J_{\mathrm{III}}, \quad \text{(III)}$$
$$\rightarrow O_2(b^1\Sigma_g^+)+O(^3P) \quad J_{\mathrm{III}'}, \quad \text{(III}')$$
$$\rightarrow O_2(a^1\Delta_g)+O(^3P) \quad J_{\mathrm{III}''}, \quad \text{(III}'')$$
$$\rightarrow O_2(X^3\Sigma_g^-)+O(^1D) \quad J_{\mathrm{III}'''}, \quad \text{(III}''')$$
$$O_3+h\nu$$
$$(410<\lambda\leqslant610\text{ nm})\rightarrow O_2(a^1\Delta_g)+O(^3P) \quad J_{\mathrm{IV}}, \quad \text{(IV)}$$
$$O_3+h\nu$$
$$(610<\lambda<750\text{ nm})\rightarrow O_2(X^3\Sigma_g^-)+O(^3P) \quad J_{\mathrm{V}}. \quad \text{(V)}$$

Here J_i is the rate of the ith photodissociation process, and $h\nu$ is the absorbed radiation quantum. We call attention to the fact that process (II) is dominant up to altitudes $H=30$ km in the upper atmosphere, and processes (IV) and (V) are dominant at lower altitudes. This condition is illustrated in Fig. 1, which shows altitude profiles of J_i [$i=1-5$ in Roman numerals, corresponding to processes (I)–(V)], calculated for a solar zenith angle $\chi=0$ from the equation

$$J(z,\chi)=\int_{\lambda_0}^{\lambda_0+\Delta\lambda} I_0(\lambda)\sigma(\lambda)\varphi(\lambda)\tau(\lambda,z,\chi)d\lambda, \quad (1)$$

where $I_0(\lambda)$ is the extra-atmospheric solar irradiance, $\sigma(\lambda)$ is the radiation absorption cross section, $\varphi(\lambda)$ is the quantum

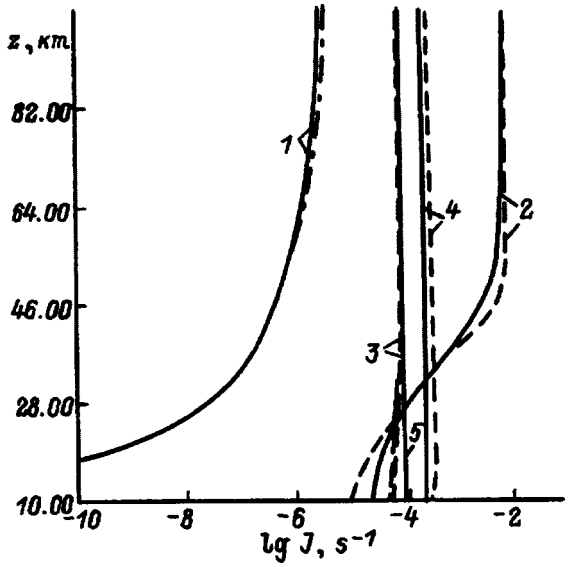

FIG. 1. Altitude profiles of the O_3 photodissociation rates in different spectral intervals. *1)* J_{I}; *2)* J_{II}; *3)* J_{III}; *4)* J_{IV}; *5)* J_{V}.

yield of the reaction, $\tau(\lambda,z,\chi)$ is the transmission function of the atmosphere at the altitude $H=z$ for the solar zenith angle χ. The same procedure as in Ref. 9 is used to calculate this function with allowance for Rayleigh scattering and the altitude variation of the principal absorbing components (O_2 and O_3). The values of $I_0(\lambda)$, $\sigma(\lambda)$, and $\varphi(\lambda)$ are taken from Refs. 6, 10, and 11. Also shown are the results of calculations in Refs. 7 and 12, represented by the dot-dashed and dashed lines, respectively.

The most important source of atomic oxygen needed for ozone production in the atmosphere is the dissociation of O_2 molecules under the influence of ultraviolet radiation from the sun.[6] It is currently assumed that the photodissociation of molecular oxygen existing in the electronic ground state $O_2(X^3\Sigma_g^-)$ takes place during the absorption of radiation with $\lambda=115-242$ nm. Four spectral intervals are usually distinguished in this case[6,7]: 1) the Lyman bands (L_α with $\lambda=115-121.6$ nm; 2) the Schumann–Runge continuum with $\lambda=130-175$ nm; 3) the Schumann–Runge bands with $\lambda=175-198$ nm; 4) the Herzberg continuum with $\lambda=198-242$ nm. In this case the photodissociation reactions are written in the form

$$O_2(X^3\Sigma_g^-)+h\nu(\lambda\leqslant 121.6\ \text{nm}) \rightarrow O_2(e^1\Delta_u)\rightarrow 2O(^1D) \quad J_{\mathrm{VI}}, \tag{VI}$$

$$O_2(X^3\Sigma_g^-)+h\nu(130<\lambda\leqslant 175\ \text{nm}) \rightarrow O_2(1^3\Pi_u)\rightarrow 2O(^3P) \quad J_{\mathrm{VII}}, \tag{VII}$$

$$O_2(X^3\Sigma_g^-)+h\nu(175<\lambda\leqslant 198\ \text{nm}) \rightarrow O_2(B^3\Sigma_u^-)\rightarrow 2O(^3P) \quad J_{\mathrm{VIII}}, \tag{VIII}$$

$$O_2(X^3\Sigma_g^-)+h\nu(198<\lambda\leqslant 242\ \text{nm}) \rightarrow \left\{\begin{array}{l} O_2(C^3\Delta_u) \\ O_2(A^3\Sigma_u^+) \\ O_2(c^1\Sigma_u^-) \end{array}\right\} \rightarrow 2O(^3P) \quad J_{\mathrm{IX}}. \tag{IX}$$

Recently, the photodissociation of O_2 from the excited states $a^1\Delta_g$ and $b^1\Sigma_g^+$ is being considered as a secondary source of free oxygen in the stratosphere.[1–3] According to Ref. 1, the most probable transitions in the Herzberg continuum are

$$O_2(a^1\Delta_g)+h\nu(200\leqslant\lambda<320\ \text{nm}) \rightarrow \left\{\begin{array}{l} O_2(C^3\Delta_u) \\ O_2(A^3\Sigma_u^+) \end{array}\right\} \rightarrow 2O(^3P) \quad J_{\mathrm{X}}, \tag{X}$$

$$O_2(b^1\Sigma_g^+)+h\nu(250\leqslant\lambda<370\ \text{nm}) \rightarrow \left\{\begin{array}{l} O_2(C^3\Delta_u) \\ O_2(A^3\Sigma_u^+) \end{array}\right\} \rightarrow 2O(^3P) \quad J_{\mathrm{XI}}. \tag{XI}$$

We note that the change in spin multiplicity in transition from the singlet state $a^1\Delta_g$ ($b^1\Sigma_g^+$) to the triplet state $C^3\Delta_u$ presents a relatively mild hindrance to the indicated transitions. At the present time, unfortunately, there are no data from measurements of the cross sections of processes (X) and (XI). Previously calculated[13] values of σ do not exceed half the corresponding values of σ for $O_2(X^3\Sigma_g^-)$. Possible channels of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ photodissociation in the shorter-wavelength spectral range corresponding to the Schumann–Runge continuum and bands are transitions to the antibonding terms $1^3\Pi_u$ and $1^1\Pi_u$ (Refs. 2 and 9)

$$O_2(a^1\Delta_g)+h\nu(141\leqslant\lambda<212\ \text{nm}) \rightarrow 1^1\Pi_u\rightarrow 2O(^3P) \quad J_{\mathrm{XII}}, \tag{XII}$$

$$O_2(a^1\Delta_g)+h\nu(141\leqslant\lambda<239\ \text{nm}) \rightarrow 1^3\Pi_u\rightarrow 2O(^3P) \quad J_{\mathrm{XIII}}, \tag{XIII}$$

$$O_2(b^1\Sigma_g^+)+h\nu(158.3\leqslant\lambda<231\ \text{nm}) \rightarrow 1^1\Pi_u\rightarrow 2O(^3P) \quad J_{\mathrm{XIV}}, \tag{XIV}$$

$$O_2(b^1\Sigma_g^+)+h\nu(158.3\leqslant\lambda<275\ \text{nm}) \rightarrow 1^3\Pi_u\rightarrow 2O(^3P) \quad J_{\mathrm{XV}}. \tag{XV}$$

These transitions are allowed optically, so that their cross sections σ must be several orders of magnitude greater than the cross sections of transitions (X) and (XI). For absorption in the Lyman bands the $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ photolysis products are $O(^1D)$ and $O(^1S)$ (Ref. 14):

$$O_2(a^1\Delta_g)+h\nu(124.3\leqslant\lambda<141\ \text{nm}) \rightarrow {}^1\Delta_u\rightarrow 2O(^1D) \quad J_{\mathrm{XVI}}, \tag{XVI}$$

$$O_2(a^1\Delta_g)+h\nu(\lambda\leqslant 124.3\ \text{nm}) \rightarrow k^1\Delta_u\rightarrow 2O(^1D)+O(^1S) \quad J_{\mathrm{XVII}}, \tag{XVII}$$

$$O_2(b^1\Sigma_g^+)+h\nu(\lambda\leqslant 158.3\ \text{nm}) \rightarrow \alpha^1\Sigma_u^+ \begin{cases} \nearrow\ O(^3P)+O(^1S) & J_{XVIII} \\ \searrow\ O(^1D)+O(^1D) & J_{XIX}. \end{cases} \quad \text{(XIX)}$$

In general, a functional dependence $\sigma_i(\lambda)$ analogous to the function $\sigma(\lambda)$ for transitions from the $X^3\Sigma_g^-$ state must exist for $i=a^1\Delta_g, b^1\Sigma_g^+$. Figure 2 shows graphs of $\sigma(\lambda)$ for the $a^1\Delta_g$ and $b^1\Sigma_g^+$ states (curves *1* and *2*, respectively) in the interval $\lambda=160-360$ nm; the graphs have been obtained analytically on the assumption that the behavior of $\sigma(\lambda)$ is identical for transitions from these states and from the $X^3\Sigma_g^-$ state both in the Herzberg continuum and in the Schumann–Runge bands and continuum.[9] Also shown is a graph of $\sigma(\lambda)$ for transitions from the $X^3\Sigma_g^-$ (curve *3*, Ref. 6); the dashed curves represent graphs of $\sigma_i(\lambda)$ obtained in Ref. 13 for the Herzberg continuum. We call attention to the fact that the shift of the frequency interval toward higher values of λ in the photodissociation of O_2 from the $a^1\Delta_g$ and $b^1\Sigma_g^+$ states leads not only to an increase in the solar irradiance, but also to an increase in $\tau(\lambda,z,\chi)$. Indeed, as an example, an O_3 absorption maximum in the Hartley band ($\lambda=255$ nm) occurs within the interval of the Herzberg continuum for $O_2(X^3\Sigma_g^-)$, so that the solar radiation intensity in this spectral interval decreases rapidly with decreasing z. For the intervals $\lambda=240-300$ nm and $\lambda=275-356$ nm corresponding to the Herzberg continuum for $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$, respectively, the absorption by O_3 is substantially (50 times) weaker, and $\tau(\lambda,z,\chi)$ is much greater here. Consequently, J_X and J_{XI} must be higher than J_{IX}, even though the values of σ for $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ in the corresponding spectral intervals are lower than for $O_2(X^3\Sigma_g^-)$. This assertion is clearly evident from Table I, which shows the values of $I_0(\lambda)$ for $\lambda=200-360$ nm, the values of $\sigma_i(\lambda)$, $i=X^3\Sigma_g^-$, $a^1\Delta_g$, and $b^1\Sigma_g^+$ for transitions (IX), (X), and (XI), and the values of J_{IX}, J_X, and J_{XI} at various altitudes, calculated from Eq. (13). At $z\leqslant 20$ km there is a decisive contribution to the photodissociation of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ from absorption in the Herzberg continuum, and at $z>20$ km the same is true for absorption in the Schumann–Runge bands and continuum.[9] The concentration

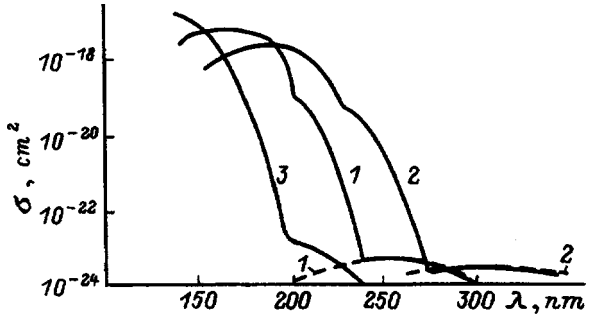

FIG. 2. Absorption cross sections versus wavelength.

of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ in the stratosphere is insignificant, so that despite the large values of $J(z,\chi)$, the photodissociation of $O_2(a^1\Delta_g)$ and $O_2(b^1\Sigma_g^+)$ as a source of atomic oxygen can be disregarded up to $z=45$ km.

Excited O_2 molecules are also formed in collisional recombination, in E–E exchange processes, and in chemical reactions involving the interaction of O_3 with $O(^3P)$ and $O(^1D)$ atoms:[7,8,15]

$$O(^3P)+O(^3P)+M \rightarrow \begin{cases} \nearrow\ O_2(X^3\Sigma_g^-)+M, \\ O_2(a^1\Delta_g)+M, \\ \searrow\ O_2(b^1\Sigma_g^+)+M, \end{cases} \quad \text{(XIXa)}$$

$$O_2(X^3\Sigma_g^-)+O(^1D) \begin{cases} \nearrow\ O_2(a^1\Delta_g)+O(^3P), \\ \searrow\ O_2(b^1\Sigma_g^+)+O(^3P), \end{cases} \quad \text{(XX)}$$

$$O_2(a^1\Delta_g)+O_2(a^1\Delta_g)\rightarrow O_2(b^1\Sigma_g^+)+O_2(X^3\Sigma_g^-), \quad \text{(XXI)}$$

$$O_3+O(^3P)\rightarrow \begin{cases} \nearrow\ O_2(b^1\Sigma_g^+)+O_2(X^3\Sigma_g^-), \\ O_2(a^1\Delta_g)+O_2(X^3\Sigma_g^-), \\ \searrow\ 2O_2(X^3\Sigma_g^-), \end{cases} \quad \text{(XXII)}$$

TABLE I. Absorption cross sections and photodissociation rates of $O_2(X^3\Sigma_g^-)$, $O_2(a^1\Delta_g)$, $O_2(b^1\Sigma_g^+)$ in the Herzberg continuum.

Wavelength λ, nm	Extra-atmospheric solar radiation flux $I_0\times10^{-4}$, $lx\cdot cm^{-2}$, s^{-1} nm	Absorption cross section of O_2 mol. in Herzberg continuum σ, cm^2			Altitude z, km	Photodissociation rate of molecular oxygen, s^{-1}		
		$X^3\Sigma_g^-$	$a^1\Delta_g$	$b^1\Sigma_g^+$		J_{IX}	J_X	J_{XI}
200	7.2(11)	1.29(−23)	1.1(−24)	0	10	2.22(−19)	1.61(−9)	1.86(−8)
220	5.53(12)	6.0(−24)	2.9(−24)	1.5(−25)	15	1.69(−15)	2.09(−9)	2.14(−8)
240	5.16(12)	1.0(−24)	4.3(−24)	8.0(−25)	20	2.62(−13)	2.81(−9)	2.32(−8)
260	1.25(13)	0	4.2(−24)	1.65(−24)	25	5.55(−12)	3.75(−9)	2.48(−8)
280	1.41(13)	0	2.9(−24)	2.6(−24)	30	4.16(−11)	5.05(−9)	2.64(−8)
290	9.21(13)	0	2.0(−24)	2.85(−24)	35	1.66(−10)	6.50(−9)	2.82(−8)
300	6.10(13)	0	1.1(−24)	2.95(−24)	40	3.66(−10)	7.78(−9)	2.98(−8)
310	9.51(13)	0	4.5(−25)	2.85(−24)	45	5.53(−10)	9.11(−9)	3.10(−8)
320	1.59(14)	0	1.0(−25)	2.6(−24)	50	6.80(−10)	1.04(−8)	3.18(−8)
330	2.11(14)	0	0	2.3(−24)	55	7.57(−10)	1.15(−8)	3.21(−8)
340	2.41(14)	0	0	2.05(−24)	60	8.04(−10)	1.20(−8)	3.22(−8)
350	2.73(14)	0	0	1.65(−24)				
360	3.10(14)	0	0	1.3(−24)				

TABLE II. Kinetic model and process rate constants in an $O-O_2-O_3$ mixture.

Process No.	Reaction	$K_i(J_k)$, $(cm^3)^{n-1}/s(s^{-1})$	τ_p, s $H=20$ km	References
1	$O_3+h\nu(198<\lambda<610\ nm)\rightarrow O_2(^1\Delta g)+O(^3P)$	$J_1=J_{II}+J_{III}+J_{IV}$	2.92×10^3	[8,9]
2	$O_3+h\nu(610\leqslant\lambda\leqslant730\ nm)\rightarrow O_2(^3\Sigma_g^-)+O(^3P)$	$J_2=J_V$	1.13×10^4	[9]
3	$O_2(^3\Sigma_g^-+h\nu(198,<\lambda<242\ nm)\rightarrow O(^3P)+O(^3P)$	$J_3=J_{IX}$	1.76×10^{12}	[6,9]
4	$O_2(1^\Delta_g)+h\nu(220\leqslant\lambda\leqslant320\ nm)\rightarrow O(^3P)+O(^3P)$	$J_4=J_X$	3.28×10^8	[9,13]
5	$O_2(^3\Sigma_g^-+h\nu_I(\lambda=1.27\ \mu m)\rightarrow O_2(^1\Delta_g)$	$W_I\left(I=100\ \frac{W}{cm^2}\right)$	6.78	
6′	$O_3+O(^3P)\rightarrow O_2(^1\Delta_g)+O_2(^3\Sigma_g^-)$	$K_{6'}=0.9\times10^{-11}\exp\left(-\frac{2300}{T}\right)$	2×10^{-3}	[15,18]
6″	$O_3+O(^3P)\rightarrow 2O_2(^3\Sigma_g^-)$	$K_{6''}=0.9\times10^{-11}\exp\left(-\frac{2300}{T}\right)$	2.5×10^{-3}	[15,18]
7	$O_3+O_2(^1\Delta_g)\rightarrow 2O_2(^1\Sigma_g^-)+O(^3P)$	$K_7=5.22\times10^{-11}\exp\left(-\frac{2840}{T}\right)$	5.22×10−3	[18]
8	$O_2(^1\Delta_g)+M\rightarrow O_2(^3\Sigma_g^-)+M$	$K_8^M=\begin{cases}M=O_2 & 1.6\times10^{-18}\\ M=O_3 & 4\times10^{-15}\\ M=O & 1.6\times10^{-16}\end{cases}$	0.31 1.35×10^{-4} 3.4×10^{-5}	[18] [18] [8]
9′	$O^3P)+O(^3P)+M\rightarrow O_2(^3\Sigma_g^-)+M$	$K_{9'}=2.4\times10^{-33}\cdot(T/300)^{-3}$	7.32×10^{-5}	[15,18]
9″	$O(^3P)+O(^3P)+M\rightarrow O_2(^1\Delta_g)+M$	$K_{9''}=2.4\times10^{-33}\cdot(T/300)^{-2}$	7.32×10^{-5}	[15,18]
10	$O(^3P)+O_2+M\rightarrow O_3+M$	$K_{10}=6.9\times10^{-34}\cdot\left(\frac{T}{300}\right)^{-1.25}$	3.14×10^{-4}	[18]
11	$O_2(^1\Delta_g)\rightarrow O_2(^3\Sigma_g^-)+h\nu(\lambda=1.27\ \mu m)$	$A_{mn}=2.58\times10^{-4}$	3.88×10^3	[6]

$$O_3+O(^1D)\begin{cases}\nearrow & 2O_2(X^3\Sigma_g^-),\\ \nearrow & O_2(a^1\Delta_g)+O(X^3\Sigma_g^-),\\ \searrow & O_2(b^1\Sigma_g^+)+O(X^3\Sigma_g^-),\\ \searrow & O_2(X^3\Sigma_g^-)+2O(^3P)\end{cases}\qquad\text{(XXIII)}$$

Comparatively rapid fragmentation of singlet oxygen occurs together with its production:[6–8]

$$O_2(a^1\Delta_g)+M\rightarrow O_2(X^3\Sigma_g^-)+M,\qquad\text{(XXIV)}$$

$$O_2(b^1\Sigma_g^+)+M\rightarrow O_2(a^1\Delta_g)+M,\qquad\text{(XXV)}$$

$$O_2(a^1\Delta_g)\rightarrow O_2(X^3\Sigma_g^-)+h\nu,\qquad\text{(XXVI)}$$

$$O_2(b^1\Sigma_g^+)\rightarrow O_2(X^3\Sigma_g^-)+h\nu,\qquad\text{(XXVII)}$$

$$O_3+O_2(a^1\Delta_g,b^1\Sigma_g^+)\rightarrow 2O_2(X^3\Sigma_g^-)+O(^3P).\qquad\text{(XXVIII)}$$

The recombination of O and O_2 results in the production of ozone:

$$O(^3P)+O_2(X^3\Sigma_g^-)+M\rightarrow O_3+M.\qquad\text{(XXIX)}$$

For all practical purposes reactions (I)–(XXIX) determine the kinetics of processes in the oxygen atmosphere.

KINETICS OF OZONE PRODUCTION IN THE STRATOSPHERE UNDER THE INFLUENCE OF RADIATION WITH $\lambda_I\approx1.27\ \mu m$

We consider the simple model mixture $O-O_2-O_3$ exposed to radiation with a frequency ν_I equal to the resonance frequency of the electronic–vibration transition of the O_2 molecule $m(X^3\Sigma_g^-,V',J',K')\rightarrow n(a^1\Delta_g,V'',J'',K'')$, where V' and V'' are the vibrational quantum numbers, and J',K' and J'',K'' are the rotational quantum numbers in the electronic states $X^3\Sigma_g^-$ and $a^1\Delta_g$. We recall that optical transitions between the $X^3\Sigma_g^-$ and $a^1\Delta_g$ states are forbidden by the selection rules. However, spin-orbit interaction produces absorption lines that are allowed in the magnetic-dipole approximation. Rotational levels appear in the $a^1\Delta_g$ state only for $J''=K''\geqslant2$, whereas in the $X^3\Sigma_g^-$ state each rotational level with quantum number K' consists of three spin components with $J'=K'+1$, $J'=K'$, and $J'=K'-1$ (Refs. 16 and 17).

It is evident from the foregoing discussion that the processes exhibit rather complicated kinetics even in such a simple atomic–molecular system. However, our analysis has shown that for $H<25$ km a simplified scheme based exclusively on reactions involving $O(^3P)$, $O_2(X^3\Sigma_g^-)$, $O_2(a^1\Delta_g)$, and O_3 can be used in the case of laser-excited $O_2(a^1\Delta_g)$. Table II lists the reactions included in this scheme. Also shown are the rate constants K_i of collisional processes, the photodissociation process rates J_i determined for $\chi=0$ on the basis of earlier recommendations,[6–9,13,18], the induced transition rates W_I, and characteristic time constants of these processes at $z=20$ km. The quantity τ_I is determined from the relation

$$(\tau_I)^{-1}=W_I=\sigma_{mn}I/h\nu_I,\quad \sigma_{mn}=\frac{\lambda_{mn}^2}{4\pi b_D}A_{mn}\sqrt{\frac{\ln 2}{\pi}}H(x,a),$$

where I is the laser intensity, λ_{mn} is the wavelength at the center of the absorption transition line, A_{mn} is the Einstein coefficient for this process, b_D is the Doppler linewidth (FWHM), and $H(x,a)$ is the Voigt function, which is calculated from standard relations with the influence of both Doppler and collisional spectral line-broadening mechanisms

taken into account; here the line-broadening cross sections for the collision of O_2 molecules with various partners $M=O_2, O_3, O$ are assumed to be equal to the gas-kinetic values.

We shall assume that thermodynamic equilibrium exists between the vibrational, rotational, and translational degrees of freedom in the investigated spectral intervals. The system of equations describing processes *1–11* (Table II) can be written in the form

$$\frac{dN_1}{dt}=-J_3N_1+J_2N_3-W_I\left(\frac{g_n}{g_m}N_m-N_n\right)+N_3N_4(K_6'+2K_6'')+N(K_9'N_4^2-K_{10}N_1N_4)+N_2\left(\sum_M K_8^M N_M+W_S+2K_7N_3\right), \quad (2)$$

$$\frac{dN_2}{dt}=J_1N_3+W_I\left(\frac{g_n}{g_m}N_m-N_n\right)-N_2\left(\sum_M K_8^M N_M+W_S+K_7N_3+J_4\right)+N_3N_4K_6'+K_9''N_4^2N, \quad (3)$$

$$\frac{dN_3}{dt}=-(J_1+J_2)N_3-K_6N_3N_4-K_7N_3N_2+K_{10}N_1N_4N, \quad (4)$$

$$\frac{dN_4}{dt}=2(J_3N_1+J_4N_2)+N_3(J_1+J_2-K_6N_4+K_7N_2)-2K_9N_4^2N-K_{10}N_1N_4N. \quad (5)$$

Here N_i is the density of particles of the ith species: $i=1$ for $O_2(X^3\Sigma_g^-)$; $i=2$ for $O_2(a^1\Delta_g)$; $i=3$ for O_3, and $i=4$ for $O(^3P)$; N_m and N_n are the densities of O_2 molecules in states m and n, g_m and g_n are the multiplicities of their degeneracy; $N=\Sigma_{i=1}^4 N_i$; $K_6=(K_6'+K_6'')$, $K_9=(K_9'+K_9'')$; and $W_S=A_{mn}$. Under the stated assumptions

$$N_m=N_1\varphi_m,$$

$$\varphi_m=\frac{g_mB_{v'}}{kT}\frac{\exp[-\theta_1V'/T]}{1-\exp(-\theta_1/T)}\exp\left(-\frac{E_{j'}}{kT}\right),$$

$$N_n=N_2\varphi_n,$$

$$\varphi_n=\frac{g_nB_{v''}}{kT}\frac{\exp[-\theta_2V''/T]}{1-\exp(-\theta_2/T)}\exp\left(-\frac{E_{j''}}{kT}\right),$$

where θ_1 and θ_2 are characteristic vibrational temperatures, and $E_{j'}$ and $E_{j''}$ are the rotational energies of the O_2 molecules in the $X^3\Sigma_g^-$ and $a^1\Delta_g$ states, $B_{v'}$ and $B_{v''}$ are the rotational constants in vibrational states V' and V'', and k is Boltzmann's constant.

The quantities $E_{j'}$ and $E_{j''}$ are calculated with allowance for splitting of the j' level in the $X^3\Sigma_g^-$ state into three components with $j'=K'+1$, $J'=K'$, and $J'=K'-1$ (Ref. 16).

We analyze the dynamics of $N_i(t)$, assuming that the density of the mixture is constant. We consider the conditions when $\tau_{pul}\gg\tau_D, \tau_T, \tau_K$, where τ_{pul} is the duration of irradiation, and τ_D, τ_T, and τ_K are characteristic time constants of diffusion, heat conduction, and convection without regard for the temperature variation in the irradiation zone. The initial conditions are consistent with atmospheric conditions at the given altitude at time t_0, and the variation of the solar irradiance is modeled by varying the optical thickness of the atmosphere ($I=0$ at $t<t_0$, and $I=I_0$ at $t_0\leqslant t<t_0+\tau_{pul}$).

We consider the basic processes responsible for the production and destruction of ozone in the atmosphere. In the absence of an artificial source ($W_I=0$) the $O_2(a^1\Delta_g)$ concentration at $H<45$ km is low as a result of rapid extinction in collisions. The photodissociation of $O_2(a^1\Delta_g)$ as a source of $O(^3P)$ can be disregarded. The principal source of $O(^3P)$ in this case is the photodissociation of O_3 and $O_2(X^3\Sigma_g^-)$, although reaction *7* involving $O_2(a^1\Delta_g)$ also plays an appreciable role. Ozone production occurs in the recombination of $O(^3P)$ and $O_2(X^3\Sigma_g^-)$ (reaction *10*).

Exposure to radiation whose frequency coincides with the frequency of the electronic–vibration transition $m(X^3\Sigma_g^-,V')\rightarrow n(a^1\Delta_g,V'')$ significantly alters the kinetics of O_3 production. In fact, the $O_2(a^1\Delta_g)$ photodissociation rate J_4 at an altitude of 20 km is 10^4 times the rate J_3 for $O_2(X^3\Sigma_g^-)$ (Table I). Consequently, on the one hand, the advent of a large quantity of $O_2(a^1\Delta_g)$ leads to the production of additional atomic oxygen and the acceleration of reaction *10*, but then, on the other hand, it raises the O_3 fragmentation rate in reaction *7*. This circumstance is illustrated in Fig. 3, which shows the rates S_i of production and fragmentation of $O_2(a^1\Delta_g)$ (a), $O(^3P)$ (b), and O_3 (c). The values of the rates S_i, where $S_1=(J_1+J_2)N_3$; $S_3=J_3N_1$; $S_4=J_4N_2$; $S_5=W_I(g_n/g_mN_m-N_n)$; $S_6=K_6N_3N_4$; $S_7=K_7N_3N_2$; $S_8=N_2\Sigma_MK_8^MN_M$; $S_9=K_9N_4^2N$; $S_{10}=K_{10}N_1N_4N$; $S_{11}=W_SN_2$, were obtained by numerical solution of the system (2)–(5) for the case of irradiation at a frequency equal to the resonance frequency of the center of the $^0P(3)$ line of the $m\rightarrow n$ transition with $V'=V''=0$, $J'=4$, $J''=3$, and $K'=K''=3$ for $I=50$ W/cm^2, $\tau_{pul}=5$ h and for $I=0$ (dashed lines) at 3 h after sunrise on the equator ($t_0=8$ h 41 min) at $H=20$ km (the numbers *1–11* in Fig. 3 correspond to the rates of processes 1–11 in Table II, and $S_9=S_{9'}=S_{9''}$).

The rates of O_3 production and fragmentation under the influence of laser radiation depend on W_I, τ_{pul}, and t_0. The induced transition rate W_I (which depends on I_0) determines the rate of replenishment of excited $O_2(a^1\Delta_g)$ molecules, the laser pulse duration τ_{pul} determines the number of these molecules during irradiation, while the time t_0 at which irradiation begins dictates the zenith angle and, hence, the solar radiation intensity, i.e., J_4. Figure 4 shows the results of numerical calculations of the variation of the density of O_3 molecules N_3 at an altitude of 20 km under the influence of radiation with various values of $I_0(W_I)$, τ_{pul}, and t_0. It is evident that when t_0 differs only slightly from the time of sunrise (t_*), laser irradiation lowers the O_3 concentration, owing to the dominant influence of process *7*, since the $O_2(a^1\Delta_g)$ photodissociation rate is negligible during this time (see Fig. 3). Later, however, as J_4 increases, the ozone concentration is observed to increase as well. For sufficiently large τ_{pul} the final value of N_3 can be much greater than the

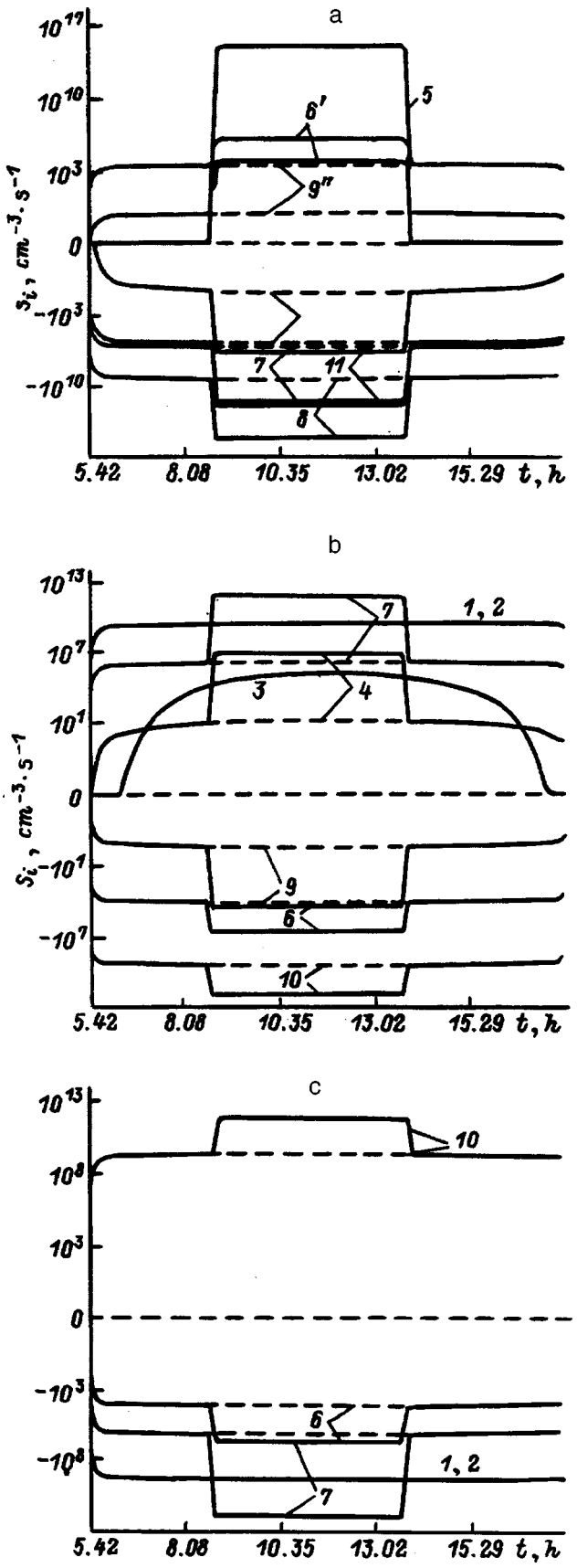

FIG. 3. Time variation of the rates of production and destruction of $O_2(a^1\Delta_g)$ (a), $O(^3P)$ (b), and O_3 (c) during laser irradiation of the atmosphere at $\lambda \approx 1.25\ \mu$m for $W_I > 0$ (solid curves) and $W_I = 0$ (dashed curves).

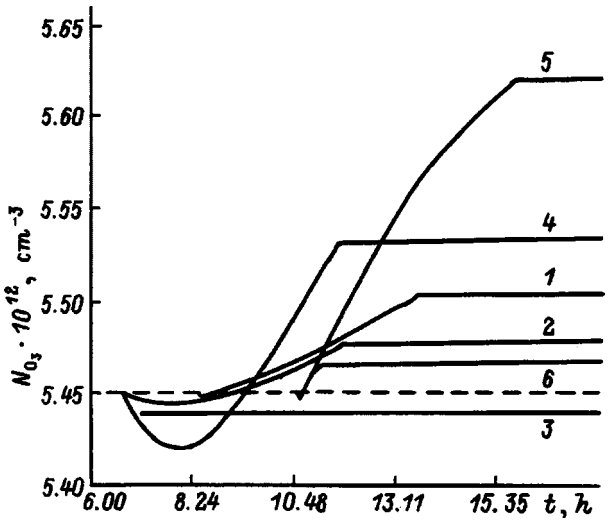

FIG. 4. Time variation of the density of O_3 molecules in the irradiated zone. *1,2,5*) $I_0 = 50$ W/cm^2; *3,4,6*) $I_0 = 200$ W/cm^2; *1,2,4,5*) $\tau_0 = 5$ h; *3,6*) $\tau_0 = 20$ min.

density of O_3 molecules in the undisturbed atmosphere (dashed line). For $I = 50$ W/cm^2, $\tau_{\text{pul}} = 5$ h, and $t_0 = 8$ h 41 min this increase attains 1.5%, and for $t_0 = 10$ h 40 min it even be as much as 3.3%.

These results imply the existence of a domain of parameters $(t_0, \tau_{\text{pul}}, W_I)$ in which the O_3 production rate in the atmosphere can increase under the influence of radiation with $\lambda \approx 1.27\ \mu$m. These parameters determine the density of replenished $O_2(a^1\Delta_g)$ molecules. It is therefore pertinent to consider the following problem: What must $N_2(t)$ be in order for dN_3/dt at $W_I > 0$ and $t_0 < t \leqslant t_0 + \tau_{\text{pul}}$ to be greater than dN_3^0/dt at $W_I = 0$. We seek a solution for $N_3(t)$ in the form $N_3(t) = \varepsilon(t) N_3^0(t)$, where $\varepsilon(t) > 1$ is a bounded continuous function; here (and from now on) $N_i^c(t) = N_i$ ($W_I = 0$) and $N_i^e = N_i^c(t_0)$. We make use of the fact that the contributions of reactions 6 and 9 to the dynamics of N_i under the influence of radiation can be disregarded under the given conditions, and $\Sigma_i K_{10}^i N_i = K_{10} N = \text{const}$. It is now readily shown that

$$\frac{dN}{dt}[2(N_1 + N_2) + 3N_3 + N_4] = 0, \tag{6}$$

$$\frac{dN_4}{dt} + \frac{dN_3}{dt} = 2(J_3 N_1 + J_4 N_2). \tag{7}$$

Taking Eq. (7) into account, we can write the inequality $dN_3/dt > dN_3^0/dt$ in the form

$$-(J_1 + J_2) N_3^0 (\varepsilon - 1) - K_7 N_3^0 (\varepsilon N_2 - N_2^0) + K_{10} N N_1 \times \left[N_3^e + N_4^e - \varepsilon N_3^0 + 2 \int_{t_0}^{t} (J_3 N_1 + J_4 N_2) dt \right] - K_{10} N N_1^0 \times \left[N_3^e + N_4^e - N_3^0 + 2 \int_{t_0}^{t} (J_3 N_1^0 + J_4 N_3^0) dt \right] > 0. \tag{8}$$

Let $J_4(t)$ and $J_3(t)$ be monotonic, nondecreasing functions in the interval $[t_0 \leqslant t \leqslant t_0 + \tau_{\text{pul}}]$ (these conditions pre-

vail in the atmosphere from sunrise to noon). Using these conditions along with relations (6) and (7), we readily obtain the upper bounds

$$\sum_{i=1}^{3} N_i^e-(N_2+N_1+\varepsilon N_3^0)\leqslant N_1\delta_1+N_2\delta_2, \tag{9}$$

$$\sum_{i=1}^{3} N_i^e-(N_1+N_2+\varepsilon N_3^0)\geqslant N_1^e\delta_1^0+N_2^e\delta_2^0. \tag{10}$$

Here $\delta_1=\tau_{\rm pul}/\tau_{F3}$, $\delta_2=\tau_{\rm pul}/\tau_{F4}$, $\tau_{Fi}(i=3,4)=[J_i(t)]^{-1}$, and $\delta_i^0=\delta_i(t_0)$. Making use of (9), (10), and the following relations, which hold at $z=20$ km, $\tau_{F4}/\tau_{F3}\approx 1.2\times 10^{-4}$, $N_1^e\approx N_1^0(t)$, and $N_1^0(t)\gg\Sigma_{i=2}^4 N_i^c(t)$, we write inequality (8) in the form

$$N_2^2+bN_2+c<0,$$

$$b=-N_1^0\left[1-\varepsilon\frac{\tau_F}{2\tau_{\rm pul}}\gamma_3^0\right],$$

$$c=N_1^0\left[\frac{N_3^0(\varepsilon-1)\tau_{F4}}{2\tau_{\rm pul}}+N_2^0\left(1-\frac{\tau_F}{2\tau_{\rm pul}}\gamma_3^0\right)\right],$$

$$\tau_{10}=[N^2K_{10}]^{-1},\quad \tau_7=[NK_7]^{-1},$$

$$\tau_F=\tau_{F4}\tau_{10}/\tau_7,\quad \gamma_i^c=\frac{N_i^0}{N_1^0}. \tag{11}$$

Inequality (11) has two real positive roots if $b^2-4c>0$ and $b<0$. In this case there exists a set of densities N_2 such that $N_2^{\rm inf}\leqslant N_2\leqslant N_2^{\rm sup}$. For $b>0$ there is one positive root, and $0<N_2\leqslant N_2^{\rm sup}$. We now estimate $N_2^{\rm sup}$ and $N_2^{\rm inf}$ for the indicated cases. The inequality $b<0$ holds for $\tau_{\rm pul}>\varepsilon\gamma_3^0\tau_F/2$, which for $z=20$ km at noon and $\varepsilon\approx 1$ corresponds to $\tau_{\rm pul}>120.8$ s (here the condition $b^2-4c>0$ holds for any $\tau_{\rm pul}>0$). We note that for any τ_F it is possible to choose $\tau_{\rm pul}$ such that $b<0$. This indicates that even when $\tau_{F4}\gg\tau_7$ [the photodissociation of $O_2(a^1\Delta_g)$ is a slower process than the chemical fragmentation of O_3], $\tau_{\rm pul}$ can be chosen so that $dN_3/dt>dN_3^0/dt$. Allowing for the fact that $\tau_{10}\gamma_3^0/\tau_7\ll 1$, for $N_2^{\rm sup}$ and $N_2^{\rm inf}$ we obtain

$$N_2^{\rm sup}\approx N_1^0\left[1-\gamma_3^0\frac{(\varepsilon-1)\tau_{F4}}{2\tau_{\rm pul}}-\gamma_2^0\right],$$

$$N_2^{\rm inf}\approx N_3^0\left[\frac{\gamma_2^0}{\gamma_3^0}+\frac{(\varepsilon-1)\tau_{F4}}{2\tau_{\rm pul}}\right]. \tag{12}$$

It follows from Eqs. (12) that $N_2^{\rm sup}$ increases and, conversely, $N_2^{\rm inf}$ decreases as τ_{F4} decreases and as $\tau_{\rm pul}$ increases. We have thus shown that if the quantity of $O_2(a^1\Delta_g)$ replenished by irradiation with $\lambda=1.27$ μm satisfies the indicated set of solutions, the rate of O_3 production in the atmosphere for $W_I>0$ is greater than for $W_I=0$, and $N_3(t)>N_3^0(t)$.

In the case $b>0$ a necessary condition for the existence of $N_2^{\rm sup}>0$ is the relation $c<0$, which is valid for $\varepsilon<1$ and $\tau_{\rm pul}<\varepsilon\gamma_3^0\tau_F/2$, i.e., the density of O_3 molecules in the interval $\tau_{\rm pul}$ decreases in this case $[N_3(t)<N_3^0(t)]$. If $\varepsilon\ll(1+\gamma_2^0\tau_{10}/\tau_7)/(1+\tau_{10}/4\tau_7)$, then $N_2^{\rm sup}$ is given by the expression

$$N_2^{\rm sup}\approx\frac{N_3^0}{2}\sqrt{\frac{2}{\gamma_3^0\tau_{\rm pul}}(\tau_{F4}(1-\varepsilon)+\gamma_2^0\tau_F)}. \tag{13}$$

In contrast with the preceding case, $N_2^{\rm sup}$ now decreases as $\tau_{\rm pul}$ increases for all $\tau_{\rm pul}<\varepsilon\gamma_3^0\tau_F/2$.

Inasmuch as N_2 depends on W_I, it is clear that conditions also exist for the quantity I. However, even in the elementary case when τ_I is much shorter than the characteristic times of collisional processes, no one has succeeded in obtaining expressions in explicit form for the bounds on W_I and I. The value of t_0 also occurs implicitly in $N_2^{\rm sup}$ and $N_2^{\rm inf}$, because τ_{F4} and τ_{F3} depend on t. As t_0 to noontime, the value of τ_{F4} decrease more rapidly than τ_{F3} (Ref. 9). A natural lower bound on N_2 is N_2^0. The density N_2 also has an upper bound, which depends on the irradiation technique. In the event of saturation of the absorbing transition we have $N_2^{\max}=G_{mn}N_1^0/(1+G_{mn})$, where $G_{mn}=g_nN_2^eN_m^e/g_mN_1^eN_n^e$. Figure 5 shows the bounds on N_2 as functions of the dimensionless parameter $\omega=(\varepsilon-1)\tau_{F4}/\tau_{\rm pul}$, which characterizes the relative variation of the O_3 concentration under the influence of radiation during the time $\tau_{\rm pul}$ starting with the time $t=t_0$. Here curves *1* and *2* correspond to $N_2^{\rm inf}/N_3^0=f(\omega)$ and $N_2^{\rm sup}/N_3^0=\varphi(\omega)$ for the case $b<0$, when $\varepsilon>1$, curve *3* corresponds to $N_2^{\rm sup}/N_3^0=\psi(\omega)$, when $b>0$ and $\varepsilon<1$ for $\tau_{\rm pul}<\varepsilon\gamma_3^0\tau_F/2$ $(\omega<0)$, and curves *4* and *5* correspond to the natural bounds N_2^0/N_3^0 and $N_2^{\max}/N_3^0$.

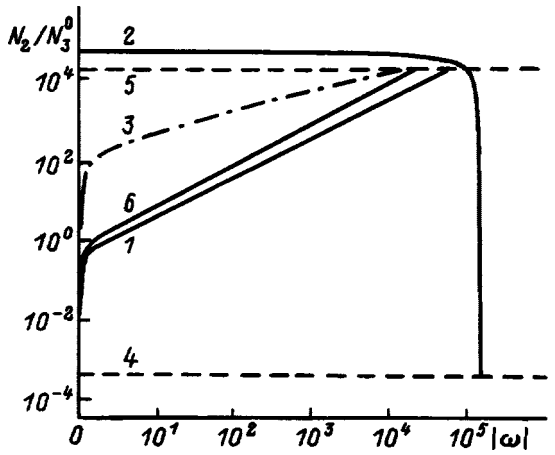

FIG. 5. Bounds on the $O_2(a^1\Delta_g)$ concentration versus the parameter ω.

We now examine the values of N_2 such that for a given ω the maximum values of ε or, equivalently, the maximum values of N_3 are attained at $t=\tau_{\rm pul}$. Let $\tau_8\ll\min\{\tau_7,\tau_{F4},W_S^{-1}\}$, $\tau_8=(\Sigma_M K_8^M N_M)^{-1}$ (this condition is satisfied at $H\leqslant 30$ km). The variations of N_1 and N_2 in the interval $[t_0,t_0+\tau_{\rm pul}]$ can now be written in the approximation of a two-level system:

$$N_1=B/A+(N_1^e-B/A)\exp(-A(t-t_0)/\tau_\Delta),$$

$$N_2=N_2^e+(N_1^e-B/A)[1-\exp(-A(t-t_0)/\tau_\Delta)],$$

$$A=\left(\frac{g_n}{g_m}\varphi_m+\varphi_n\right)\frac{\tau_\Delta}{\tau_I}+1,$$

$$B=\left(\varphi_n\frac{\tau_\Delta}{\tau_I}+1\right)(N_1^e+N_2^e),\quad \tau_\Delta=\tau_8. \tag{14}$$

Taking Eq. (7) into account, for $N_3(t)$ we have

$$N_3=\exp(-a_1(t))\left[N_3^e+\int_{t_0}^{t}b_1(t)\exp(a_1(t))dt\right],$$

$$a_1(t)=\int_{t_0}^{t}[J_1+J_2+K_7N_2+K_{10}NN_1]dt,$$

$$b_1(t)=\left[N_3^e+N_4^e+2\int_{t_0}^{t}[J_3N_1+J_4N_2]dt\right]K_{10}N_1N. \tag{15}$$

We now introduce simplifying assumptions. Let $\tau_{\text{pul}}>A^{-1}\tau_\Delta$ and $t_0\gg\tau_*$. We can then write the solutions (14) in the form

$$N_1=(N_1^e+N_2^e)y,\quad N_2=(N_1^e+N_2^e)(1-y),$$

$$y=\left(1+\varphi_n\frac{\tau_\Delta}{\tau_I}\right)\left[1+\varphi_n\left(1+\frac{g_n}{g_m}\frac{\varphi_m}{\varphi_n}\right)\frac{\tau_\Delta}{\tau_I}\right]^{-1}. \tag{16}$$

Next we make use of the fact that $\tau_{10}\ll\tau_{F1},\tau_{F2}$ for the given conditions. If $J_4N_2\gg J_3N_1$ in this case (the latter condition holds for $N_2>N_2^{\text{inf}}$), then at $t_0>t_*$

$$\int_{t_0}^{t}(J_4N_2+J_3N_1)dt=J_4\int_{t_0}^{t}N_2dt,$$

and for $a_1(t)$ and $b_1(t)$ we readily obtain

$$a_1(t)=a_2(t-t_0)\tau_{10},$$

$$b_1(t)=ya_1^e(N_3^e+N_4^e)[1+b_1^e(1-y)(t-t_0)],$$

$$a_1^e=(\gamma_1^e+\gamma_2^e)/\tau_{10},\quad b_1^e=2/\tau_{F4}(\gamma_1^e+\gamma_2^e)(\gamma_3^e+\gamma_4^e)^{-1},$$

$$a_2=a_1^e[y+(1-y)\tau_{10}/\tau_7]. \tag{17}$$

Substituting Eqs. (17) into (15) and making use of the fact that $N_3^e\gg N_4^e$, we obtain

$$N_3=N_3^e\{c_1(1-c_2)[1-\exp(-(t-t_0)a_2)]+c_1c_2a_2(t-t_0)+\exp(-(t-t_0)a_2)\},$$

$$c_1=y[y(1-\tau_{10}/\tau_7)+\tau_{10}/\tau_7]^{-1},\quad c_2=(1-y)\frac{b_1^e}{a_2}. \tag{18}$$

The function $N_3(t)$ determined from (18) has a minimum for

$$t_m=t_0+\frac{\ln c_3}{a_2},\quad c_3=1+\frac{1-c_1}{c_1c_2},$$

$$N_3^{\min}=N_3^e\left\{c_1+\frac{(1-(1-c_2)c_1)+c_1c_2(\ln c_3-1)}{c_3}\right\}.$$

At $\tau_{\text{pul}}<t_m$, therefore, irradiation causes N_3 to decrease, and at $\tau_{\text{pul}}>t_m$ the ozone concentration can increase in the irradiated zone. At $t_m\ll t\leqslant t_0+(b_1^e)^{-1}$ the density N_3 essentially increases linearly with time [for $z=20$ km, $I_0=10$ W/cm^2, $t_0=8$ h 41 min, $t_m=t_0+1.8\times10^{-2}$ s, and $(b_1^e)^{-1}=3.15\times10^3$ s]: $N_3=N_3^ec_1\{1-c_2+c_2a_2(t-t_0)\}$. In this time interval the function $N_3(y)$ has a maximum at $y=y_0$:

$$y_0=\left(\frac{\tau_7}{\tau_{10}}-1\right)^{-1}\times\left(\sqrt{1+\left(\frac{\tau_7}{\tau_{10}}-1\right)\left(1+\frac{\tau_{\text{pul}}\gamma_3^e(t-t_0)}{2\tau_{F4}}\right)}-1\right),$$

if

$$\frac{\gamma_3^e\tau_{10}}{2\tau_7}\leqslant(t-t_0)/\tau_{F4}\leqslant\frac{\gamma_3^e}{2}\frac{\tau_{10}}{\tau_7}\times\left[(1+G_{mn})^{-1}-\frac{\tau_{10}}{\tau_7}(g-1)^2\right].$$

The maximum value of ε in this case is given by the expression $\varepsilon_{\max}=N_3^{\max}/N_3^e=c_1+(1-y_0)b_1^e(t-t_0)$. Since y depends on the ratio τ_Δ/τ_I, it is clear that a decrease in τ_I causes ε to decrease for $y<y_0$ and, conversely, to increase for $y>y_0$. Figure 5 shows the dependence of N_2/N_3^e on the parameter ω corresponding to $\varepsilon_{\max}$ (curve *6*). We note that the range of possible values of N_2 for which $\varepsilon_{\max}$ is attained for a given τ_{pul} increases as ω increases.

From these considerations we infer the existence of a large set of parameters characterizing irradiation $(t_0,\tau_{\text{pul}},W_I)$, for which the excitation of molecular oxygen in the $O_2(a^1\Delta_g)$ state causes the rate of production of O_3 in the stratosphere to increase and for which the O_3 concentration in the irradiated zone increases significantly.

This work has received financial support form the Russian Fund for Fundamental Research (Grant No. 96-01-01617).

[1] J. E. Frederic and R. J. Cicerone, J. Geophys. Res. **90**, 10.733 (1985).
[2] T. G. Slanger, L. E. Jusinski, G. Black, and G. E. Gadd, Science **241**, 945 (1988).
[3] G. E. Gadd, and T. G. Slanger, Science **250**, 1433 (1990).
[4] A. M. Starik, O. N. Favorskiĭ, O. S. Khabarov, and B. N. Amelin, Vestn. Ross. Akad. Nauk **63**, 1082 (1993).
[5] R. P. Wayne, J. Photochem. **25**, 345 (1984).
[6] M. J. McEwan and L. F. Phillips, *The Chemistry of the Atmosphere* [Halsted Press, New York, 1975; Mir, Moscow, 1978].
[7] R. Rodrigo, J. J. Lopez-Moreno *et al.*, Planet. Space Sci. **34**, 723 (1986).
[8] Yu. A. Kulagin, L. A. Shelepin, and V. I. Yarygina, Tr. Fiz. Inst. Ross. Akad. Nauk **212**, 166 (1994).
[9] A. M. Starik, O. V. Taranov, and O. E. Pashkova, ''Photodissociation of molecular oxygen and ozone in the atmosphere,'' TsIAM Preprint No. 24 [in Russian] (Central Scientific-Research Institute of Aircraft Engines, 1995).
[10] I. T. Molina and M. J. Molina, J. Geophys. Res. **91**, 14.501 (1986).
[11] G. H. Mount and J. Rottman, J. Geophys. Res. **88**, 5403 (1983).
[12] T. Shimazaki, J. Atmos. Terr. Phys. **46**, 173 (1984).
[13] R. P. Saxon and T. G. Slanger, J. Geophys. Res. **91**, 9877 (1986).
[14] P. Krupenie, J. Phys. Chem. Ref. Data **1**, 423 (1972).
[15] A. I. Zakharov, K. S. Klopovskiĭ, A. P. Osipov *et al.*, Fiz. Plazmy **14**, 327 (1988) [Sov. J. Plasma Phys. **14**, 191 (1988)].
[16] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Vol. 3 [Pergamon Press, New York–Oxford, 1974; Nauka, Moscow, 1974].
[17] C. Amiot and J. Verges, Can. J. Phys. **59**, 1391 (1981).
[18] D. C. Baulch, R. A. Cox, F. F. Hampson *et al.*, J. Phys. Chem. Ref. Data **13**, 1259 (1984).

Translated by James S. Wood

Increase in the hydraulic pressure in various zones of a two-fluid hydrophilic capillary due to nonuniformity of the external electric field

D. V. Tikhomolov and O. N. Slyadneva

St. Petersburg State University, 199124 St. Petersburg, Russia
(Submitted October 16, 1996; resubmitted May 6, 1997)
Zh. Tekh. Fiz. **68**, 24–30 (August 1998)

The phenomenon whereby the thickness of the water film next to the inner surface of a hydrophilic capillary filled with two fluids increases in the presence of an external static electric field is investigated. A hypothesis of the essential nature of the phenomenon is submitted, along with a corresponding equation for calculating numerically the thickness variations as a function of several parameters of the working system, including the strength of the external field. Experimental results are given. The results of theoretical calculations and the experimental data are shown to be in good agreement. The orientation of stationary dipoles of the molecules constituting highly polar liquids in an electric field is estimated. © *1998 American Institute of Physics.* [S1063-7842(98)00408-5]

The subject of this paper is the experimentally established phenomenon that the thickness of a polar electrolyte (water) film contained between a nonpolar fluid (whether a gas or a liquid is immaterial) and a solid surface increases in systems investigated by us with a special intricate geometrical structure (the details of which are described below) under the influence of a static or slowly alternating (guaranteed up to 50 Hz) electric field. In our opinion, the specific geometrical structure of the system (the measurement cell) described below is not really significant at the conceptual, qualitative level, but was chosen by us because, first, it enabled us to carry out a specific numerical calculation of the physical constants from the experimental results and, second, this cell was the one used in the actual experiments. Consequently, the results of the study can be applied to systems having other geometrical configurations, e.g., to classical direct emulsions. The sum-total of the results available to us at the present time, characterizing various aspects of the phenomenon (optical, electrical, etc.), has led us to hypothesize that this phenomenological macrophysical property of the given system is attributable to an anomaly of the molecular structure of water as a special case of a liquid consisting of highly polar molecules.

One of the present authors has published the results of an experimental study[1–3] of the dc electrical conductivity of a model emulsion in a hydrophilic cylindrical capillary. The system consisted of a quartz glass capillary filled with a current-conducting aqueous solution of an electrolyte (the dispersion medium) and containing in its interior a nonpolar fluid (gas or liquid — the disperse phase) in the form of a column extending along the axis of the capillary. The length of the column (l_c) was greater than the diameter of the capillary ($2r_k$), but no more than tenfold. When the capillary was filled with fluids, owing to the natural hydrophilic quality of glass, a thin film of water remained between the lateral surface of the column and the inner surface of the capillary; the thickness of the film in the equilibrium state was $r_f = 10-100\ \mu$m.

It followed from the experiments that the film thickness increased by an order of magnitude or more when the capillary was exposed to an external static electric field characterized by an average strength along the capillary $E_k = U_k/l_k = 0.1-4$ V/cm (U_k is the potential difference between the ends of the capillary, and l_k is the length of the capillary). Optical observations confirmed that the new equilibrium state acquired by the film under the influence of the external field, as a rule[1)], had a constant thickness along the axis of the capillary; synchronous optical observations correlate qualitatively with electrical measurements.[2,4] At this point, however, we note that a change of size and shape can be detected but not measured by optical means.

These observations have provided the basis of an electrokinetic logging method which has been tested in on-site geophysical industrial-exploration operations[5,6] but whose development has been halted because of (among other reasons) a lack of well-developed parametric relations.

In 1993 one of the present authors published a paper[7] advancing a hypothesis to account for the increase in the film thicknesses under the influence of an external field. However, the vast inventory of experimental data accumulated previously over a wide range of possible variables has underscored the advisability of elaborating the advanced hypothesis and subjecting it to more rigorous theoretical analysis. Underlying the theoretical solution of the problem is a model working system used in virtually all scientific studies concerned with modeling the behavior of capillary-porous bodies filled with two fluids.[8–13] Schematically the given model system (Fig. 1) consists of a hydrophilic cylindrical capillary having a constant circular cross section (of radius r_k). A nonpolar fluid column in the shape of a cylinder (of radius r_c), terminated at the ends in hemispheres of the same radius as the cylindrical part of the column[2)], is placed in the middle part of the capillary. The remaining interior space of

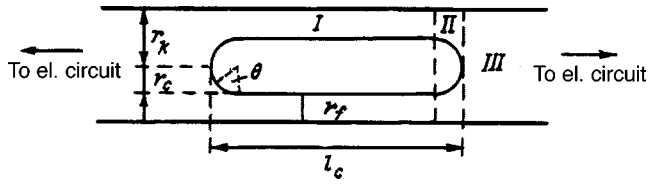

FIG. 1. Diagram of the model capillary system (shown schematically — actually $r_k \gg r_f$).

the capillary is filled with an aqueous electrolyte solution, which forms a thin electrolytic buffer layer (wetting fluid) of thickness $r_f = r_k - r_c$ between the outer cylindrical surface of the column and the inner surface of the capillary.

In the absence of external force (including electric) fields the equilibrium value of r_f depends on the balance of forces: the capillary pressure from the cylindrical part of the column and all possible components of the disjoining pressure present in the film.[14]

When the ends of the capillary (with the indicated equilibrium already established in it) are connected to the electric circuit, the magnitudes of the external field E are constant in the axial direction in the region occupied by the aqueous solution alone (zone *III* in Fig. 1) and in the water film adjoining the cylindrical surface of the column (zone *I* in the figure); in other words, the field is uniform in both zones *I* and *III*. The water in these zones is polarized in the external field and, like any highly polar substance, primarily because of the orientation component[3)], i.e., orientation of the polar water molecules as constant dipoles in the external field.[16] Here the poles of the H_2O dipole[4)], being of opposite sign and equal magnitude, are acted upon by electric forces equal in absolute value but oppositely directed.

A different situation is encountered in the case of water dipoles in the electrolyte-occupied zone between the surfaces of the hemispheres and the inner surface of the capillary (zone *II* in Fig. 1). The dipoles are oriented in this zone as well, but here the lines of force of the electric field bend (i.e., the field is not uniform) and, accordingly, the electric forces acting on the pole of the dipole do not cancel out. We assume that no other factors influence the dipole in the system. According to classical mechanics principles, the molecule then acquires angular momentum in the direction of increasing external field. By virtue of the symmetry of the force fields the motion of the dipoles in both directions of the column is directed from zone *III* to zone *I*, creating in zone *I* an elevated hydraulic pressure against the cylindrical surface of the nonpolar fluid and thereby pushing it in the direction normal to the axis of the capillary. The motion stops if a countereffect occurs in the working system during application of the external field. Since the water phase is continuous and the column has a finite volume, when this motion of the molecules is summed over all dipoles in zone *II*, it should cause the thickness of the water film to increase and, accordingly, the radius of the column to decrease. The latter effect implies an increase in the capillary pressure in the direction from zone *I* to zone *III*.[15]

Our theoretical derivation of the force driving the motion of the dipoles rests primarily on a series of simple constructions (Fig. 1), which for now are based on the conventional classical physics description of the significant parameters of molecular media in terms of a statistical mean, which remains invariant in an infinitely small unit volume of the medium.

We choose the direction of flow past the hemispheres as the coordinate along which the variations of the parameters in the investigated process are to be estimated:

$$x = r_c \sin\theta = (r_k - r_f)\sin\theta. \tag{1}$$

In accordance with the foregoing discussion, we write the relation between the x-component of the external electric field and the geometrical parameters of zone II in the form

$$E_x = \frac{d(IR)}{dx} = I\frac{1}{\chi\pi(a^2+x^2)}, \tag{2}$$

where I is the current intensity, which as a charge flow in the absence of concentration (charge) polarization[5)] does not depend on position in the aqueous solution, i.e., is not a function of x; R is the electrical resistance, χ is the electrical conductivity of the solution, and $\pi a^2 = \pi(r_k^2 - r_c^2)$ is the transverse cross section of the annular film in zone *I*.

In the postulated physical scenario of increasing film thickness we are concerned not with the absolute strength of the field as with its increment dE:

$$dE = \frac{I}{\chi\pi} d\left(\frac{1}{a^2+x^2}\right). \tag{3}$$

We now describe the medium with its molecular structure taken into account. For this purpose we arbitrarily delimit a vacuous region in zone *II*. We characterize the size of this void by the radius of a sphere whose volume equals the statistical mean volume of the medium associated with the fraction of each molecule, taking into account existing notions as to the structural configuration of specific liquids (water in this case). At the center of the sphere we place a water molecular dipole characterized by the distance between its poles d_{dip} and by the charge e_{dip}. Inasmuch as the force characteristics of the external electric field acting on a molecule in the systems treated here depend mainly on the Ohmic losses of the medium (and not on the dielectric losses as in Onsager's paper[16]), we place the delimited void in a continuum characterized by the conductivity χ of the solution.

It is a well-known fact that in an external static electric field the dipole electric vector of a molecule of a polar substance (water) tends to align itself in a direction tangential to the direction of the external field vector. We assume by virtue of the entropy factor that the dipole vector and the external field vector are in different directions separated by an angle γ. It is obvious that the effectiveness of the external field acting on the molecular dipole is determined quantitatively in the geometrical sense by the average projection of the constant dipole moment of the molecule onto the direction of the external field, i.e., in accordance with the relation $E_{\text{eff}} = E\cos\gamma$. In the statistical sense the efficiency of the external field acting on the dipole is proportional to the cor-

relation w of the rotation of the dipole moment in the direction of the external field.[16] We can therefore write

$$w=\cos\ \gamma. \tag{4}$$

The electric force df_1 acting in zone *II* on the dipoles in the volume of a water ring of thickness equal to the characteristic diameter d_{H_2O} of the void for the water molecule:

$$df_1=e_{\text{dip}}\frac{\rho_M N_A}{M}\pi(a^2+x^2)d_{H_2O}\frac{I}{\chi\pi}d\left(\frac{1}{a^2+x^2}\right)\frac{d_{\text{dip}}}{d_{H_2O}}\cos\ \gamma, \tag{5}$$

where $e_{\text{dip}}\omega_M N_A/M$ is the statistical mean density for one of the poles of the dipole per unit volume of water, E_{dip} is the pole charge of the H_2O dipole (equal in absolute value to the electron charge), ρ_M is the bulk density of H_2O, M is the molecular mass of H_2O, $\pi(a^2+x^2)d_{H_2O}$ is the volume of the indicated ring,

$$\frac{I}{\chi\pi}d\left(\frac{1}{a^2+x^2}\right)$$

is the difference between the field strengths on the surfaces bounding the indicated ring, and d_{dip} is the distance between poles of the dipole in the H_2O molecule (we shall assume from now on that d_{dip} is equal to 0.43×10^{-10} m, in accordance with Ref. 18); we use the last term d_{dip}/d_{H_2O} to adjust for the fact that the probability density of the dipoles in the ring is lower in the statistical mean than for H_2O (the distance between poles of the dipole is smaller than the diameter of the water molecule and so is even smaller than the diameter of the void).

The electric force acting on all the dipoles in each zone *II* is

$$\begin{aligned}F_{\text{el}}&=\int_{a^2}^{r_k^2}\left[\frac{I}{\chi}d_{\text{dip}}e_{\text{dip}}\frac{\rho_M N_A}{M}\right](a^2+x^2)d\left(\frac{1}{a^2+x^2}\right)\cos\ \gamma\\&=\frac{I}{\chi}d_{\text{dip}}e_{\text{dip}}\frac{\rho_M N_A}{M}\ln\frac{r_k^2}{r_k^2-r_f^2}w\\&=\frac{I}{\chi}d_{\text{dip}}e_{\text{dip}}\frac{\rho_M N_A}{M}\ln\frac{r_k^2}{r_k^2-(r_k-r_f)^2}w.\end{aligned} \tag{6}$$

The force created by the evolution of the capillary pressure in the case of a cylindrical surface is

$$F_k=\left(\frac{\sigma}{r_c}\right)\pi(r_k^2-r_c^2)=\frac{\sigma}{r_k-r_f}\pi[r_k^2-(r_k-r_f)^2], \tag{7}$$

where σ is the two-fluid interfacial surface tension.

Thus, proceeding from the basic concepts of forces acting in the system and governing the equilibrium film thickness in the capillary connected into the circuit producing the external field, we obtain

$$w\ \frac{e_{\text{dip}}\rho_M N_A I d_{\text{dip}}}{M\chi}\ln\frac{r_k^2}{r_k^2-(r_k-r_f)^2}=\frac{\sigma\pi}{r_k-r_f}\{r_k^2-(r_k-r_f)^2\}. \tag{8}$$

However, since the electric field parameter appearing in the equation, i.e., the current intensity, depends on the specific structure of the conducting channel, it is incorrect to substitute the experimentally recorded current intensity into the theoretical equation.[6)] The parameters of the electric field must be represented in such a way as to give explicit form to the geometrical structure of the current-carrying channel of the aqueous solution. This requirement is easily met if the current intensity is represented as the quotient of the potential difference between the ends of the total cylinder $U_k=\text{const}$ divided by the resistance of (again) the total capillary $R_k=f(r_k)$. In the postulated system model the capillary resistance is the sum of the resistances of three parts occurring in series: the part outside the column (on both sides in summation), the part encompassing both menisci (again in summation), and the part where the water film is bounded by the cylindrical part of the surface of the column. The corresponding equation has been published earlier:[7]

$$\begin{aligned}R_k=&\frac{2}{\chi\pi\sqrt{r_k^2-(r_k-r_f)^2}}\arctan\frac{(r_k-r_f)}{\sqrt{r_k^2-(r_k-r_f)^2}}\\&+\frac{l_c-2(r_k-r_f)}{\chi\pi[r_k^2-(r_k-r_f)^2]}+\frac{l_k-l_c}{\chi\pi r_k^2}.\end{aligned} \tag{9}$$

The first term on the right side reflects the resistance in the vicinity of the menisci, the second term refers to the water ring in the cylindrical part of the column, the third term gives the resistance outside the column, l_c is the length of the column including the menisci, l_k is the length of the capillary, and χ is the conductivity of the solution. Replacing I in Eq. (8) by the voltage U_k as a quantity independent of the structure of the conducting channel in explicit form and by R_k as written on the right side of Eq. (9), we obtain the final form of the equation derived in accordance with the postulated physical notions regarding the cause of the increase in the film thickness and the mechanism characterizing the equilibrium state of the current-carrying film:

$$\begin{aligned}&U_k w\frac{e_{\text{dip}}\rho_M N_A d_{\text{dip}}}{M[r_k^2-(r_k-r_f)^2]}\ln\frac{r_k^2}{r_k^2-(r_k-r_f)^2}\\&\times\left\{\frac{2}{\sqrt{r_k^2-(r_k-r_f)^2}}\arctan\frac{r_k-r_f}{\sqrt{r_k^2-(r_k-r_f)^2}}\right.\\&\left.+\frac{l_c-2(r_k-r_f)}{r_k^2-(r_k-r_f)^2}+\frac{l_k-l_c}{r_k^2}\right\}=\frac{\sigma}{r_k-r_f}.\end{aligned} \tag{10}$$

In concluding the theoretical discussion, we call attention to the fact that the latter equation contains two variables whose values are known *a priori*. One is the unknown resistance r_f itself, and the second is w. The substitution of the above-determined value of w into the equation is ruled out by the extremely debatable issue of the measure of orientation of the H_2O molecule in water (liquid) when the latter is exposed to a static electric field.

The question of w can be answered in principle by comparing the results of calculations of r_f by an independent method unrelated to the conceptualization proposed here.

Such a comparison can be based on an equation expressing Ohm's law for the resistance of the capillary as a whole (R_k) in terms of the resistances of the individual parts of the complex-structured current-carrying channel. The numerical values of R_k can be obtained from experiments to measure the current–voltage curves ($R_k = U_k/I$).

Here we base the Ohm's law calculations on the same model system as in the theoretical derivation. The corresponding computational equation is obviously Eq. (9). To underscore the fact that the resistance r_f calculated in this case is not connected to the theoretical derivation, but reflects the thickening phenomenon in its own right, we attach the superscript e' to the pertinent value of r_f, writing r_f^e to signify that the value corresponds to the experimental result. Like the theoretically calculated value, the ''experimental value'' can contain an error due to inconsistency between the real structure and the model structure, but this fact has no bearing on the physical postulates concerning the mechanism of the current-induced film thickening process.

The parameter w can be evaluated by two methods (according to control calculations for special cases, the values of are very close). The first method is to substitute the value of r_f^e determined from resistance measurements into Eq. (10). In this case the only unknown is w. The second method, which we use below, entails the following. It is initially assumed that $w=1$, i.e., that the axes of all dipoles are strictly aligned with the external electric field. We denote the corresponding calculated thickness by r_f^t. The corresponding special form of (10) is

$$U_k \frac{e_{\text{dip}}\rho_M N_A d_{\text{dip}}}{M[r_k^2-(r_k-r_f^t)^2]} \ln \frac{r_k^2}{r_k^2-(r_k-r_f^t)^2} \times \left\{ \frac{2}{\sqrt{r_k^2-(r_k-r_f^t)^2}} \arctan \frac{r_k-r_f^t}{\sqrt{r_k^2-(r_k-r_f^t)^2}} + \frac{l_c-2(r_k-r_f^t)}{r_k^2-(r_k-r_f^t)^2} + \frac{l_k-l_c}{r_k^2} \right\} = \frac{\sigma}{r_k-r_f^t}. \quad (11)$$

We use Eq. (11) to find the corresponding value of r_f^t. The value of w is obviously determined from the expression

$$w = \frac{r_f^e}{r_f^t}. \quad (12)$$

To illustrate the values of $r_f^t = f(U_k)$ obtained from Eq. (11), curve *1* in Fig. 2 is calculated for the system octane–KCl with the geometrical characteristics $r_k = 126\,\mu$m and $l_c = 1\times10^3\,\mu$m.[7)]

We now describe the experimental part of the study. The following characteristics were varied in the experiment: the radius of the capillary ($40-170\,\mu$m); the length of the column $[(10-20)r_k]$; the constituency of the electrolyte (KCL, NaCl); the nature of the nonpolar fluid (octane, air). The length of the capillary was 25 cm in all the tests. The electrolyte concentration was $0.1M$ ($\chi\sim1$ S/m). It has been remarked previously[10,11] that the system takes a long time to settle into equilibrium at zero current. Since the application of an external electric field establishes a new equilibrium

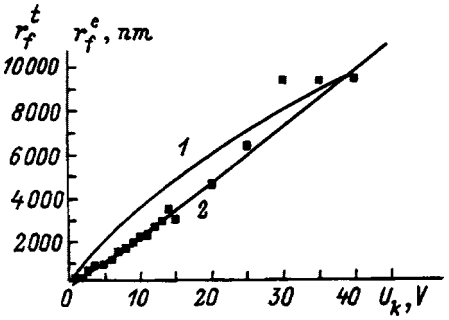

FIG. 2. Graphs of $r_f^t = f(U_k)$ and $r_f^e = f(U_k)$.

theoretically unrelated to the initial zero-current equilibrium, in the experimental work we did not wait for equilibrium in the initial system and instead began to apply the external field 25–30 min after filling the capillary with the fluid, at which time the period of rapid variations of the geometrical parameters of the film had come to an end. Figure 2 shows the experimental data for $r_f^e = f(U_k)$ as points on the graph, along with the approximating curve *2* (the correlation coefficient is 0.95) as representative of typical curves obtained in our experiments. The geometrical characteristics of the system represented by curve *2* are exactly identical with those underlying the model calculations of $r_f^t = f(U_k)$ (curve *1* in Fig. 2).[8)] We have subjected the experimental data to statistical processing[19] with a view toward discerning reliable consistencies in the variation of the parametric relations obtained from the experiment using Eq. (9) and calculated from Eq. (11). We analyzed 26 experimentally determined I–V curves at 15–20 measurement points each. The results of the calculations of r_f^e and r_f^t were compared for each curve by setting the geometrical characteristics of the model cell in the calculation of r_f^t exactly equal to the actual measured values in each experiment for the determination of r_f^e. We grouped these 26 tests into three samples differing in the general indices of the nature of the nonpolar fluid and the constituency of the electrolyte.

Thus, sample 1 comprised systems with octane and KCl solution as the fluids and with the following dimensions r_k and l_c (μm) for the ten experimental systems included in the sample (respectively): 126 and 1000; 108 and 1000; 68 and 950; 44 and 2600; 126 and 1000; 126 and 1500; 64 and 1250; 114 and 550; 157 and 1250; 143 and 1250.

Sample 2 comprised systems with octane and NaCl solution as the fluids and with the following dimensions r_k and l_c (μm) (respectively): 67 and 2350; 59 and 1500; 132 and 2350; 67 and 950; 74 and 1170; 59 and 1660; 59 and 4000; 50 and 1660; 132 and 2320; 46 and 1480.

Sample 3 comprised systems with air and KCl solution as the fluids and with the following dimensions r_k and l_c (μm) (respectively): 123 and 1400; 68 and 1800; 150 and 2350; 119 and 1750; 104 and 1600; 117 and 2550.

Calculations show that the electric field in the vicinity of the menisci (zone *II*) for all the samples was $(2-2.7)\times10^3$ V/m for $U_k = 3$ V and $(5-6)\times10^3$ V/m for $U_k = 25-40$ V. It follows from the data, in particular, that E_M depends scarcely at all on the composition and only very

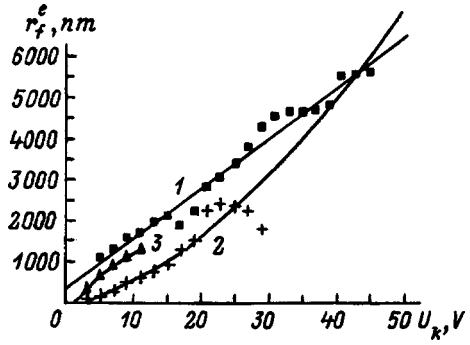

FIG. 3. Graphs of $r_f^e = f(U_k)$. The symbols represent experimental data points, and the curves are approximations to them. *1*, ■) octane–KCl; *2*, +) octane–NaCl solution; *3*, ▲) air–KCl solution.

slightly on U_k. The results of calculations of the average values from Eqs. (9)–(11) for all the samples are shown in Figs. 3 and 4.

Following are, in our opinion, the most fundamental conclusions drawn from an analysis of the actual data indicated above and shown in the figures.

1. To a high degree of approximation the values of w do not depend on U_k.

2. The values of w are essentially independent of the geometrical characteristics of the systems.

3. The sample-average values of r_f^e and r_f^t for a model having geometrical characteristics equal to the real averages and having physical parameters equal to those used in the experimental sample are of the same order, differing at most by a factor of two or three in absolute value.

4. The values of r_f^e are always lower than the corresponding values of r_f^t, indicating that the error of estimation of r_f by means of r_f^t is systematic, the discrepancy being essentially independent of the nature of the nonpolar fluid and depending very slightly on the constituency of the electrolyte in the investigated systems. In the first approximation the ratio of r_f^e to r_f^t (i.e., w or $\cos\gamma$) is constant with an estimated value in the interval 0.3–0.6.

It is obvious that for the given physical structure of the model the indicated systematic error can be attributed to the omission of several probable phenomena associated, for example, with motion (including flow), the onset of viscous friction forces, and a dearth of specific details of the geometrical structure of the conducting channel with the evaluated electrical resistances of the system. However, based on the experimentally deduced logic of the model solution, practical experience in experiments, and the analogy to the data-similarity issue encountered in an analysis of published information on the polarization of dipolar dielectric liquids, we have assumed that the most probable factor responsible for our actually observed discrepancy between the numerical data for r_f^e and r_f^t is the influence of the entropy factor on the orientation of the dipole molecules.

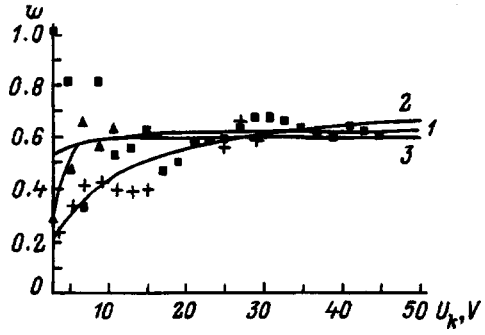

FIG. 4. Graphs of $w = f(U_k)$. The legend is the same as in Fig. 3.

The comparative analysis of the best-known theories interests us only as it bears on the final results, and for this reason we discuss in greater detail certain aspects of consistency between the resulting data and classical canons of polarization theories for dipolar dielectrics. In the physical sense we interpret the values of $w = \cos\gamma = 0.4 - 0.6$ obtained in our work in an external field of $3000 - 5000$ V/m as the average value of the projection of the constant dipole moment of the water molecule onto the direction of the external field vector. The values obtained for w are in good agreement with Kirkwood's data obtained in a calculation of the dielectric constant of water on the basis of an analysis of his model. He estimates the orientation parameter $\cos\gamma$ calculated in the present study to have a value of 0.4–0.5, and he interprets it in the geometrical sense as a projection. On the other hand, it should be noted that our values of w do not conform nearly as well to the theoretical principles of Debye and Onsager. This summarization appears legitimate to us from the practical standpoint. In particular, it is generally known[16] that the application of the Debye and Onsager theories and combinations formulated on the basis thereof to liquids consisting of polar molecules yields substantially inferior results in calculations of the main electrical parameters (e.g., the dielectric constant). In light of the historical evolution of Debye's opinions, strictly speaking, any Debye model will be inferior to Kirkwood's model in consistency with the final result. For the sake of objectivity we note that the physical significance of the entropy factor is transparent and well defined in the Debye theory: the molecular-thermal motion of molecules. In the Kirkwood theory this factor is abstract and is incorporated into the theory independently of the nature of the forces governing it. Nonetheless, both theories contain semiempirical clues. Both the local field in the Debye–Onsager theories and the height of the potential barrier in Kirkwood's theory can only be determined experimentally, in essence representing empirical correction factors to the analytical relations.

With these considerations and the above-described experimental data as guidelines, we can only state that our investigated phenomenological effect cannot be described on the basis of a Debye theory, most likely because of incorrect notions concerning the thermal motion of molecules in liquids in the indicated theoretical studies. This assertion is probably consistent with the latest opinions expressed by Debye himself (1935).

The main conclusions of our work are summarized as follows.

1. The closeness of the film thicknesses calculated on the basis of our proposed hypothesis, even without regard for the influence of the entropy factor on the orientation of the water

dipoles, to the thicknesses determined by an independent method dissociated from our hypothesis indicates the validity of the hypothesis explaining why the thickness of a water film increases under the influence of an external electric field, all the more so in overview of the evolution of the pressure gradient in highly polar liquids exposed to a nonuniform electric field.

2. The value obtained for w is interpreted as the projection of the orientation of the dipole moment vector of H_2O molecules onto the direction of the external electric field as a result of the entropy factor. The indicated value of w is very close to the parameter given by Kirkwood, which essentially characterizes the same property in highly polar liquids — the cause of the orientation effect in static electric fields.

[1)]The local instability of fluids sometimes for high values of E or prolonged exposure to the external electric field has caused optically detected lenticular water layers to evolve, adjacent to the film, so that the calculated average thickness of the film in the initial stages of its formation are too high. The experiment was halted when such a ''dimple'' was observed to develop.

[2)]In the strict theoretical sense the only dubious part of the film in the zero-field equilibrium state occurs at the junction of the hemisphere and the cylinder. Diametrically opposing notions are advanced, but all are based on the authors' own logical deliberations without rigorous proof. However, considering the order of magnitude of the thicknesses obtained experimentally and the length of the disputed zone, its resistance could not contribute significantly to the resistance of the capillary as a whole. Practically speaking, the structure of the film can differ initially from the model in that scarcely any time variation of the conductivity is observed, owing to the long time to settle into equilibrium and the formation of dimples, so that in reality equilibrium never occurs. Moreover, long, thin fluid columns are characterized by the development of a variety of changes due to fluctuation effects.

[3)]Overlooking for now the debate as to what should be the measure of orientation, we note that the viewpoints of most authors known to us converge in this assertion. For example, according to Kirkwood's model, the fraction of the orientation component of the polarization of water amounts to 0.85. All other authors report an even higher fraction.[15]

[4)]In the ensuing discussion we regard water molecules as an integrated physical system endowed with definite electrical and geometrical properties known from other data.

[5)]Concentration polarization should not occur for the thicknesses encountered in the experiments.[17]

[6)]The exclusion of I from the derivatives with respect to x in the derivation does not contradict the stated assumption: The current intensity differs for different structures, but in any case it is constant over the entire conducting channel for any structure, regardless of I.

[7)]Difficulties are met in solving both Eq. (11) and Eq. (9) in explicit form for r_f^t and r_f^e, respectively. We have relied on an iterative numerical procedure implemented on a computer using standard applied algorithms and programs.

[8)]The scatter of the results pertaining to systems of the given sample is extremely small for $U_k > 3$ V; at lower voltages U_k the scatter is greater and irregular, but the average values are close to the high-voltage averages.

[1] D. V. Tikhomolov, N. N. Krasikov, and L. M. Marmorshteĭn, Kolloidn. Zh. **48**, 1034 (1986).
[2] K. P. Tikhomolova, N. S. Arndt, D. V. Tikhomolov, and É. E. Voznaya, Kolloidn. Zh. **52**, 805 (1990).
[3] D. V. Tikhomolov, Author's Abstract of Candidate's Dissertation [in Russian], Leningrad Gos. Univ., Leningrad, 1986.
[4] N. N. Krasikov and D. V. Tikhomolov, Kolloidn. Zh. **48**, 1164 (1986).
[5] D. V. Tikhomolov, A. P. Konshin, and N. N. Krasikov, Geol. Razvedka, No. 2, 136 (1989).
[6] D. V. Tikhomolov, P. A. Aleksandrov, A. P. Konshin *et al.*, Inventor's Certificate No. 1484119 [in Russian], Priority April 9, 1987.
[7] D. V. Tikhomolov, Zh. Prikl. Khim. **66**, 519 (1993).
[8] A. Bretherton, J. Fluid Mech. **10**, 166 (1961).
[9] C. W. Parc and G. M. Homsy, J. Fluid Mech. **139**, 291 (1984).
[10] K. P. Tikhomolova and O. V. Kokorina, Kolloidn. Zh. **55**, 147 (1993).
[11] K. P. Tikhomolova and O. V. Kokorina, Kolloidn. Zh. **55**, 167 (1993).
[12] V. I. Ivanov, V. V. Kalinin, and V. M. Starov, Kolloidn. Zh. **55**, 32 (1993).
[13] N. P. Astrakhantseva and O. G. Us'yarov, in *Proceedings of the Conference on Surface Forces in Thin Films and Stability of Colloids* [in Russian], Nauka, Moscow (1974), pp. 155–163.
[14] B. V. Deryagin, N. V. Churaev, and V. M. Muller, *Surface Forces* [in Russian], Nauka, Moscow, 1987.
[15] Ya. I. Frenkel' and A. I. Gubanov, Usp. Fiz. Nauk **24**, 68 (1940).
[16] O. A. Osipov and V. I. Minkin, *Handbook of Dipole Moments* [in Russian], Vysshaya Shkola, Moscow (1965), 264 pp.
[17] K. P. Tikhomolova, *Electroosmosis* [in Russian], Khimiya, Leningrad (1989), 247 pp.
[18] V. V. Sinyukov, *Structure of Monatomic Liquids, Water, and Aqueous Electrolyte Solutions* [in Russian], Nauka, Moscow (1976), 255 pp.
[19] V. P. D'yakonov, *Handbook of BASIC Algorithms and Programs for Personal Computers* [in Russian], Nauka, Moscow (1989), 240 pp.

Translated by James S. Wood

Breakup of charged drops into droplets of comparable dimensions under strong spheroidal virtual deformations

V. A. Koromyslov, A. I. Grigor'ev, and S. O. Shiryaeva

Yaroslavl State University, 150000 Yaroslavl, Russia
(Submitted March 28, 1997; resubmitted February 16, 1998)
Zh. Tekh. Fiz. **68**, 31–38 (August 1998)

On the basis of the principle of minimization of energy of a closed system in which spontaneous processes are occurring, we investigate the breakup of a highly charged drop into two and three droplets of comparable dimensions under conditions of virtual spheroidal deformations.
 [S1063-7842(98)00508-X]

INTRODUCTION

As was shown in Refs. 1 and 2, a drop carrying a charge greater than or equal to the Rayleigh limit disintegrates by emitting roughly two hundred fine daughter droplets, roughly two orders of magnitude smaller in size, and losing in such an event only 0.5% of its initial mass and 23% of its initial charge. At the same time, a number of experimental studies (see, e.g., Refs. 3–5) have observed the breakup of charged drops into a small number of droplets of comparable dimensions. This type of breakup has been noted by observers for ball lightning.[6] Of course, under different experimental and natural conditions breakup of a drop into two droplets of comparable dimensions can be explained by different causative factors. For example, in the experiments performed in Ref. 5 such a breakup was explained as caused by significant viscosity of the medium damping the instability of the high modes of capillary oscillations of the unstable drop, and the mechanism of this process was analyzed on a qualitative level in Refs. 7 and 8. In the experiments in Ref. 3, breakup of a drop into droplets of comparable dimensions at a subcritical charge of the parent drop was explained as being due to significant mechanical deformations of the initially spherical drop. In Ref. 4 the same phenomenon was explained as caused by braking in a highly nonuniform external electric field.[9] In the experiments in Ref. 10 the breakup of a drop into several droplets of comparable dimensions took place in a periodically varying external electric field.

In the theory of atmospheric electricity in connection with the problem of the microseparation of charges in a thundercloud, and also in a study of the temporal evolution of charged liquid-droplet aerodispersed systems, of greatest interest is the situation in which a drop with subcritical charge (less than the Rayleigh limit) breaks up under conditions of significant spheroidal deformations caused by forces of a non-electrical nature. In connection with the above-said we will solve the problem of the breakup under conditions of virtual spheroidal deformations of a weakly charged drop into two or three drops of comparable dimensions, in analogy with how this was done in Ref. 11 in an effort to explain the phenomenon of radioactivity in the liquid-drop model of the atomic nucleus, without repeating, however, the miscalculations made in this work. The authors of Ref. 11 attempted to investigate the stability relative to strong spheroidal deformations of a charged drop with subcritical charge. However, some inaccuracies in the physical statement of the problem were made in Ref. 11, having no good effect on the ensuing analysis. In particular, they neglected the electrostatic interaction of the droplets in the final state (after breakup), which makes a substantial contribution to the balance of energy.

Thus, we will attempt to derive a condition for the breakup of a charged drop of an inviscid, incompressible liquid into two or three droplets of comparable dimensions, proceeding on the basis that: 1) the total potential energy of the daughter droplets in the final count should be the minimum value in accordance with the principle of minimum potential energy of the final state of a closed system, 2) we will take the kinetic energy of the system before and after breakup to be equal to zero, and 3) the total volume and charge of the liquid are not changed by breakup of the drop.

In the solution of the problem we consider three limiting cases: 1) a surface-charged drop of a liquid insulator, 2) a charged, ideally conducting drop, and 3) a volume-charged drop of a liquid insulator.

1. STUDY OF CONDITIONS OF DECAY OF A CHARGED DROP INTO TWO DROPLETS OF COMPARABLE DIMENSIONS

a) *Surface-charged drop of a liquid insulator.* We consider a drop of radius R with coefficient of surface tension σ and charge Q. We assume that the charge is uniformly distributed over the surface of the drop and is frozen in it. In other words, we assume that the possible breakup of a drop into smaller droplets takes place during a time much shorter than the characteristic time of surface charge redistribution effected by mobility in an electric field and diffusion. After breakup of the parent drop under strong deformation, the total potential energy of the system of two spherical daughter droplets (we ignore distortion of their spherical shape ensuing from their electrostatic interaction as small[12]) is equal to

$$U=\frac{q_1^2}{2r_1}+\frac{q_2^2}{2r_2}+4\pi\sigma r_1^2+4\pi\sigma r_2^2+\frac{q_1q_2}{2r}, \tag{1}$$

where q_1 and q_2 are the charges of the daughter droplets, r_1 and r_2 are their radii, and r is the distance between the centers of the daughter droplets.

Let $V_1/V=n$ (V_1 is the volume of the first daughter droplet, V is the volume of the parent drop). Then, from the condition of constant volume we obtain

$$V_2=(1-n)V \Rightarrow r_1=n^{1/3}R; \quad r_2=(1-n)^{1/3}R. \tag{2}$$

According to what has been said above, the total charge of the parent drop is distributed between the daughter droplets in proportion to their surface areas. Thus

$$\frac{q_1}{q_2}=\frac{S_1}{S_2}=\frac{n^{2/3}}{(1-n)^{2/3}} \quad \text{or} \quad q_1=\frac{n^{2/3}}{(1-n)^{2/3}}q_2.$$

From the condition of constant charge we obtain

$$q_2=\frac{(1-n)^{2/3}}{n^{2/3}+(1-n)^{2/3}}Q, \quad q_1=\frac{n^{2/3}}{n^{2/3}+(1-n)^{2/3}}Q.$$

Thus, the potential energy of the final state is equal to

$$U=f(n)=\left(\frac{1}{n^{2/3}+(1-n)^{2/3}}\right)^2\frac{Q^2}{2R}+4\pi\sigma R^2 \times[n^{2/3}+(1-n)^{2/3}]+\frac{n^{2/3}(1-n)^{2/3}}{[n^{2/3}+(1-n)^{2/3}]^2}\frac{Q^2}{2r}.$$

The above expression, according to the formulation of the problem, should be minimum with respect to n, i.e., the conditions

$$\frac{\partial f(n)}{\partial n}=0 \quad \text{and} \quad \frac{\partial^2 f(n)}{\partial n^2}>0. \tag{3}$$

should be fulfilled.

Requiring that the first of these conditions be fulfilled, we obtain the critical value of the Rayleigh parameter $W=Q^2/(16\pi\sigma R^3)$ of the deformed parent drop for its breakup into two daughter droplets

$$W=\frac{[n^{2/3}+(1-n)^{2/3}]^3}{4+\dfrac{4(n-n^2)^{2/3}}{m}-\dfrac{2(n-n^2)^{-1/3}(1-2n)[n^{2/3}+(1-n)^{2/3}]}{[n^{-1/3}-(1-n)^{-1/3}]m}}, \tag{4}$$

where $m=r/R$.

Note that for $W\geqslant 1$ a spherical drop disintegrates via the Rayleigh channel analyzed in Refs. 1 and 2. In the construction of the dependence $W(n)$ an indeterminacy in the choice of r arises. For the purpose of a qualitative analysis we assume the minimum possible distance, i.e., $r=r_1+r_2$. Then the graph of $W(n)$ has the form shown by curve *1* in Fig. 1a.

From the requirement that the second derivative be positive, analogously we obtain a second relation between W and n, determining the boundary of existence of possible channels

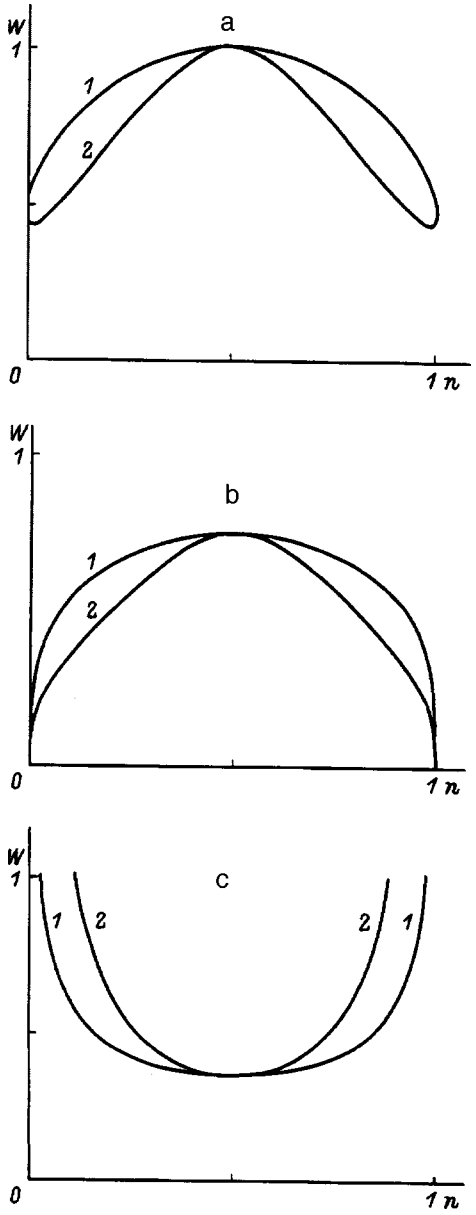

FIG. 1. Dependence of the parameter W of the parent drop on n. Locus of points of curve *1* lying above curve *2* corresponds to minimum variation of the free energy; a — surface-charged insulating drop, b — conducting drop, c — volume-charged insulating drop.

$$W>[n^{-4/3}+(1-n)^{-4/3}] \times\left[\frac{4n^{-4/3}+4(1-n)^{-4/3}}{[n^{2/3}+(1-n)^{2/3}]^3}\left(1+\frac{(n-n^2)^{2/3}}{m}\right)+24\frac{[n^{-1/3}-(1-n)^{-1/3}]^2}{[n^{2/3}+(1-n)^{2/3}]^4}\left(1+\frac{(n-n^2)^{2/3}}{m}\right)-\frac{2(n-n^2)^{-1/3}}{m[n^{2/3}+(1-n)^{2/3}]^2}\left(6+\frac{(1-2n)^2}{n-n^2}\right)+8\frac{(1-2n)[n^{-1/3}-(1-n)^{-1/3}]}{[n^{2/3}+(1-n)^{2/3}]}\right]^{-1}.$$

The graph of this function is plotted in Fig. 1a by curve

2. The requirement that the second derivative be positive is satisfied by the locus of points above curve *2*. Thus it can be seen from Fig. 1a that, depending on the value of W, the drop can break up in an arbitrary fashion, but asymmetric breakup is most likely. Symmetric breakup is possible in principle only for $W=1$, if the zero value of the second derivative at this point is taken as being positive. The possibility of the realization of such a channel is treated in more detail in Ref. 13.

b) *Charged, ideally conducting drop.* In the situation considered here, the charge, as in the preceding case, is distributed uniformly over the surface of the drop, but this idealization means that in the given case the characteristic charge redistribution time is less than the breakup time of the drop. Then the energy of the final state, as in Sec. 1a, is given by expression (1).

In contrast to the model considered earlier, now the volume and charge of the daughter droplets are both free parameters. From conservation of the total volume and charge of the liquid, assuming as before that $V_1/V=n$, we now obtain expressions (2) relating the two radii. Next, setting $q_1/Q=k$, from the condition of constant charge we obtain $q_1=Qk$ and $q_2=(1-k)Q$. Taking the notation introduced above into account, we rewrite the expression for the potential energy of the system in its final state as

$$U=f(k,n)=\frac{k^2}{n^{1/3}}\frac{Q^2}{2R}+\frac{(1-k)^2}{(1-n)^{1/3}}\frac{Q^2}{2R}+4\pi\sigma R^2 \times[n^{2/3}+(1-n)^{2/3}]+(1-k)k\frac{Q^2}{2r}.$$

As in the preceding case, this function should be minimum in the final state, i.e., the conditions

$$\left(\frac{\partial f}{\partial n}\right)_k=0,\quad \left(\frac{\partial f}{\partial k}\right)_n=0,\quad \begin{vmatrix} A & B \\ B & C \end{vmatrix}>0, \tag{5}$$

should be fulfilled, where $A=\partial^2 f/\partial k^2$, $C=\partial^2 f/\partial n^2$, $B=\partial^2 f/\partial k\partial n$.

Finding the first derivatives of $f(k,n)$ and scaling them by $(8/3)\pi\sigma R^2$ (to make them dimensionless), we obtain

$$\frac{k}{n^{1/3}}-\frac{(1-k)}{(1-n)^{1/3}}+\frac{1}{2m}(1-2k)=0,$$

$$-W\frac{k^2}{n^{4/3}}+W\frac{(1-k)^2}{(1-n)^{4/3}}+n^{-1/3}-(1-n)^{-1/3}=0, \tag{6}$$

where W is given by expression (4).

Expressing k in terms of n from the first equation of system (6)

$$k=\frac{(1-n)^{-1/3}+(2m)^{-1}}{n^{-1/3}+(1-n)^{-1/3}-m^{-1}}$$

and substituting this result in the second equation of system (6), we find the desired critical dependence $W=W(n)$

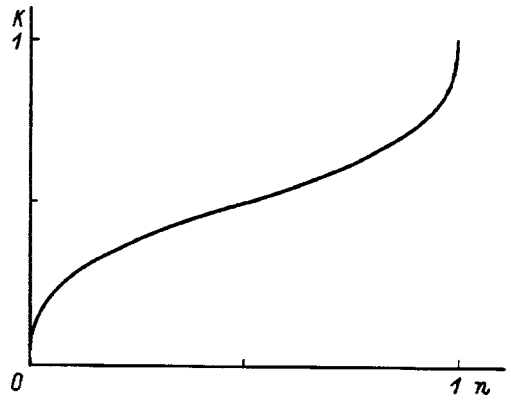

FIG. 2. Dependence of the charge fraction k of a conducting daughter droplet on its dimensionless volume n.

$$W=(n^{-1/3}-(1-n)^{-1/3})\left(\frac{k^2}{n^{4/3}}-\frac{(1-k)^2}{(1-n)^{4/3}}\right)^{-1}.$$

This function is plotted in Fig. 1b by curve *1*.

Determining the second derivatives of the function $f(k,n)$ from the extremality condition (5), we find a restriction on the existence region of possible channels of breakup

$$W>[n^{-4/3}+(1-n)^{-4/3}](n^{-1/3}+(1-n)^{-1/3}-m^{-1}) \times\left(2\frac{k^2}{n^{8/3}}+2\frac{(1-k)^2}{(1-n)^{8/3}}-4m\frac{k^2}{n^{7/3}}-4m\frac{(1-k)^2}{(1-n)^{7/3}} +4\frac{k^2}{n^{7/3}(1-n)^{1/3}}+4\frac{(1-k)^2}{(1-n)^{7/3}n^{1/3}} -4\frac{(1-k)k}{(1-n)^{4/3}n^{4/3}}\right)^{-1}. \tag{7}$$

A graph of the function $W(n)=0$ corresponding to relation (7) is plotted in Fig. 1b by curve *2*. As in Sec. 1a, we find that, depending on the value of W, the drop can break up in an arbitrary fashion, but asymmetric breakup is most likely. Symmetric breakup is possible in principle only for $W=1$, if the zero value of the second derivative at this point is treated as being positive (for more details of the possibility of realizing such a channel, see Ref. 13).

Figure 2 plots the dependence of the charge fraction of the daughter droplet on its volume. Numerical calculations show that when a daughter droplet comprising 10% of the volume of the mother drop splits off, it carries away 27.5% of the charge of the mother drop, and when the volume fraction of the daughter droplet is 0.01, the charge fraction is 0.123, when it is 0.001 the charge fraction is 0.054, and when it is 0.0001 the charge fraction is 0.024. As was already noted above, for the given case the most probable event is breakup with detachment of a daughter droplet having volume much less than that of the mother drop. Hence it is clear that for detachment of droplets having dimensionless volume $n\leqslant 0.001$ our results are in good agreement with those presented in Refs. 1 and 2. It is interesting to note that the values of W for the daughter droplets are above-critical

(above the Rayleigh limit) and, consequently, they can disintegrate anew, as was predicted in Refs. 1 and 2.

c) *Volume-charged insulating droplet.* Here we assume that the charge is uniformly distributed throughout the entire volume (frozen into the material) of the parent drop with volume density ρ. Then the electrostatic energy of this drop is given by

$$E=\int_0^R \frac{8\pi^2\rho r^4}{9\varepsilon}dr+\int_R^\infty \frac{Q^2}{2r}dr=\frac{Q^2}{2R}\left(1+\frac{1}{5\varepsilon}\right),$$

where ε is the dielectric constant of the drop.

After breakup of the parent drop the total energy of the system of two daughter droplets is equal to

$$U=\frac{q_1^2}{2r_1}\left(1+\frac{1}{5\varepsilon}\right)+\frac{q_2^2}{2r_2}\left(1+\frac{1}{5\varepsilon}\right)+4\pi\sigma r_1^2+4\pi\sigma r_2^2+\frac{q_1q_2}{2r}.$$

For a uniform charge distribution throughout the volume of the parent drop the charge divides in the same ratio as the volume, $q_1=nQ$, $q_2=(1-n)Q$. Thus, for the energy of the final state of the system for breakup of the parent drop into two daughter droplets we obtain

$$U=f(n)=\frac{n^2}{n^{1/3}}\frac{Q^2}{2R}\left(1+\frac{1}{5\varepsilon}\right)+\frac{(1-n)^2}{(1-n)^{1/3}}\frac{Q^2}{2R}\left(1+\frac{1}{5\varepsilon}\right)+4\pi\sigma R^2[n^{2/3}+(1-n)^{2/3}]+(1-n)n\frac{Q^2}{2r}.$$

In the given case the condition of minimum potential energy of the final state has the form (3). To determine the position of the minimum, we take the first derivative of the energy of the final state and set it equal to zero. After scaling out dimensions we find

$$5[n^{2/3}-(1-n)^{2/3}]\left(1+\frac{1}{5\varepsilon}\right)W+[n^{-1/3}-(1-n)^{-1/3}]+3W(1-2n)\frac{R}{r}=0.$$

Introducing m as in the previous cases and calculating W for water ($\varepsilon=81$), we obtain the dependence shown in Fig. 1c by curve *1*. From condition (3) for the second derivative we have

$$W>\frac{n^{-4/3}+(1-n)^{-4/3}}{10\left(1+\frac{1}{5\varepsilon}\right)[(n)^{-1/3}+(1-n)^{-1/3}]-18/m}. \qquad (8)$$

A graph of the function $W(n)=0$ corresponding to relation (8) is plotted in Fig. 1c by curve *2*. It is clear from the figure that breakup of the drop is possible in the given model only for $V_1\approx V_2$, as was obtained in Ref. 13.

Figure 3 plots the dependence $W(\varepsilon)$ for $n=0.5$. It can be seen that for the case of breakup of a surface-charged drop, breakup can take place in any ratio depending on W. In this case asymmetric breakup predominates. For a volume-charged drop, breakup occurs only for equal dimensions of the daughter droplets.

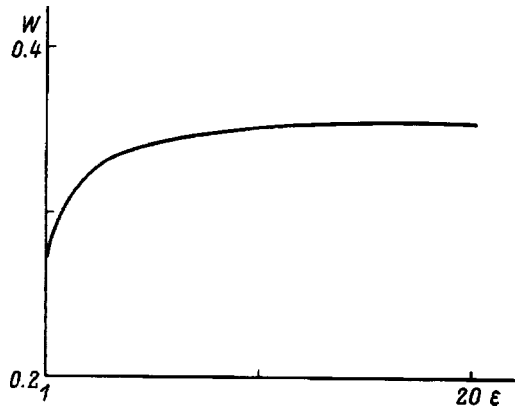

FIG. 3. Value of the parameter W necessary for breakup of a volume-charged drop into two droplets of equal size, plotted as a function of the dielectric constant ε of the liquid.

2. STUDY OF THE CONDITIONS OF BREAKUP OF A CHARGED DROP INTO THREE DROPLETS OF COMPARABLE DIMENSIONS UNDER CONDITIONS OF STRONG VIRTUAL DEFORMATIONS.

a) *Surface-charged insulating drop.* Here we assume that the charge is uniformly distributed over the surface of the parent drop and is frozen in place. After breakup under conditions of virtual deformation of the drop the potential energy of the system of three daughter droplets, where the two outer droplets are assumed to be identical by virtue of the symmetry of the problem, is equal to

$$U=\frac{2q_1^2}{2r_1}+\frac{q_2^2}{2r_2}+8\pi\sigma r_1^2+4\pi\sigma r_2^2+\frac{2q_1q_2}{2r}. \qquad (9)$$

Here q_1 and q_2 are the charges of each of the two small daughter droplets and the large daughter droplet, respectively, r_1 and r_2 are their radii, and r is the distance between the centers of the large daughter droplet and either of the two small daughter droplets.

In expression (9) we left out the term taking into account the electrostatic interaction between the two outer daughter droplets, a correct account of which, as a consequence of screening of this interaction by the central drop, is very problematic.

Let $V_1/V=n$ (V is the volume of the parent drop, V_1 is the volume of one of the small daughter drops). Then from the condition of constant volume we obtain

$$r_1=Rn^{1/3},\quad r_2=R(1-2n)^{1/3}. \qquad (10)$$

It is natural to assume that the charge of the initial drop divides between the daughter droplets in proportion to their surface areas. Then, in analogy with how we proceeded in Sec. 1a, we obtain

$$q_2=\frac{(1-2n)^{2/3}}{2n^{2/3}+(1-2n)^{2/3}}Q,$$

$$q_1=\frac{n^{2/3}}{2n^{2/3}+(1-2n)^{2/3}}Q. \tag{11}$$

Substituting expressions (10) and (11) into expression (9), we obtain an expression for the variation of the potential energy of the final state of the system for breakup of a drop into three daughter droplets

$$U=f(n)=\left[\frac{1}{2n^{2/3}+(1-2n)^{2/3}}\right]^2\frac{Q^2}{2R}+4\pi\sigma R^2[2n^{2/3}+(1-2n)^{2/3}]+\frac{2(n-2n^2)^{2/3}}{[2n^{2/3}+(1-2n)^{2/3}]^2}\frac{Q^2}{2r}.$$

We require, as before, that conditions (3) be satisfied. According to the first of these conditions, we have

$$W=\frac{[2n^{2/3}+(1-2n)^{2/3}]^3}{4+\dfrac{8(n-2n^2)^{2/3}}{m}-\dfrac{2(n-2n^2)^{-1/3}(1-4n)[2n^{2/3}+(1-2n)^{2/3}]}{[n^{-1/3}-(1-2n)^{-1/3}]m}}, \tag{12}$$

where W and m, as before, is given by expression (4).

In the construction of the dependence $W(n)$ an indeterminacy arises in the choice of the distance between the centers of the droplets r at the instant of breakup. For the purpose of a qualitative analysis we set $r=r_1+r_2$. Then the graph of the dependence $W(n)$ will have the form shown in Fig. 4a by curve *1*.

From the requirement that the second derivative be positive, we similarly obtain a second relation between W and n, defining the boundary for the existence of possible channels of breakup,

$$W>[n^{-4/3}+2(1-2n)^{-4/3}]\left\{\left[\frac{4[n^{-4/3}+2(1-2n)^{-4/3}]}{[2n^{2/3}+(1-2n)^{2/3}]^3}+48\frac{[n^{-1/3}-(1-n)^{-1/3}]^2}{[2n^{2/3}+(1-2n)^{2/3}]^4}\right]\left[1+\frac{2(n-2n^2)^{2/3}}{m}\right]-\frac{(n-2n^2)^{-1/3}}{m[2n^{2/3}+(1-2n)^{2/3}]^2}\left[24+2\frac{(1-4n)^2}{(n-2n^2)}+32\frac{(1-4n)[n^{-1/3}-2(1-2n)^{-1/3}]}{[2n^{2/3}+(1-2n)^{2/3}]}\right]\right\}^{-1}. \tag{13}$$

The region of parameter values marked off by this inequality is located above curve *2* in Fig. 4a. It is clear from Fig. 4a that, depending on the magnitude of the parameter W, the drop can break up into smaller droplets in different proportions, but only asymmetrically. It is clear that the most probable such breakup is the one in which two small droplets split off from opposite sides of the parent drop.

If in expressions (12) and (13) we double the distance between the daughter droplets in the final state (double the parameter m), then the corresponding curves will be curves *3* and *4* in Fig. 4a. If we increase the distance between the daughter droplets by tenfold, then the corresponding curves in Fig. 4a, calculating from relations (12) and (13), will be curves *5* and *6*. In summary, it is not hard to see that with growth of the distance between the daughter droplets there is a growth in the tendency toward symmetric breakup.

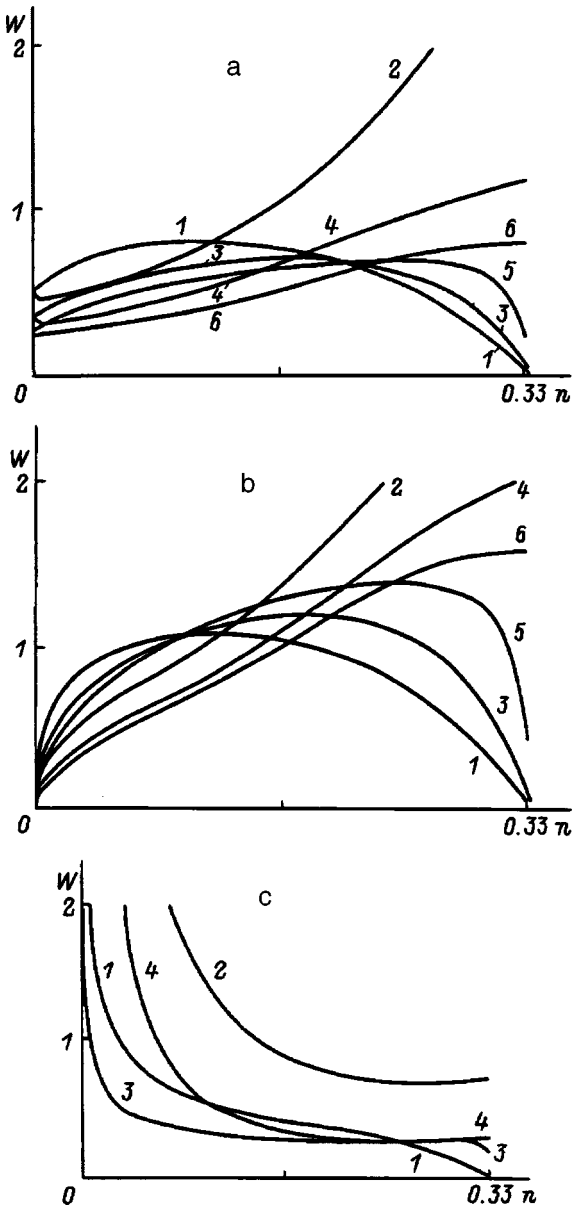

FIG. 4. a — Curves of the parameter W of a surface-charged insulating drop as a function of n, corresponding to extremal variation of the free energy of the system (curves *1*, *3*, and *5*); *1* — minimum possible deformation of the drop r, *3* — deformation $r_*=2r$, *5* — deformation $r_*=10r$; the locus of points of the ith curve lying above the $(i+1)$-th curve corresponds to minimum variation of the free energy for the corresponding initial deformation: $i=1\rightarrow r_*=r$; $i=2\rightarrow r_*=5r$; $i=3\rightarrow r_*=10r$. b — The same dependences as in (a), but for a conducting liquid. c — Dependence of the parameter W on n for a drop of insulating liquid with charge uniformly distributed over its volume, corresponding to extremal variation of the free energy of the system; *1* — minimum possible deformation $r_*=r$, *3* — $r_*=28r$; the locus of points of the ith curve lying above the $(i+1)$-th curve corresponds to minimum variation of the free energy for initial deformation $i=1\rightarrow r_*=r$; $i=2\rightarrow r_*=28r$.

b) *Charged, ideally conducting drop.* As in the preceding case, the charge is uniformly distributed over the surface of the drop, but in contrast to the previously considered case

the characteristic charge redistribution time is now much smaller than the breakup time of the drop and it becomes necessary to seek the charges of the daughter droplets from the requirement of constancy of the electric potential. The final energy state, as in the preceding case, is given by expression (9).

In contrast to the previously considered model, the dimensions and charges of the daughter droplets do not depend on one another. From conservation of the total volume of the liquid, setting $V_1/V=n$ as before, we find expressions for r_1 and r_2 defined by relation (10).

Setting $q_1/Q=k$, from conservation of charge we obtain $q_1=Qk$, $q_2=(1-2k)Q$. Taking the notation introduced above into account, the expression for the variation of the energy of the system now depends on the two parameters k and n

$$f(k,n)=\frac{2k^2}{n^{1/3}}\frac{Q^2}{2R}+\frac{(1-2k)^2}{(1-2n)^{1/3}}\frac{Q^2}{2R}+4\pi\sigma R^2[2n^{2/3}+(1-2n)^{2/3}]+2(1-2k)k.$$

As in Sec. 1b, this function should take its minimum at breakup of the parent drop, i.e., conditions (5) should be satisfied, which give a system of two equations determining the position of the extremum of the function $f(k,n)$. Hence we obtain the critical dependence for W

$$W=(n^{-1/3}-(1-2n)^{-1/3})\left(\frac{k^2}{n^{4/3}}-\frac{(1-2k)^2}{(1-2n)^{4/3}}\right)^{-1}, \quad (14)$$

where

$$k=\frac{2(1-2n)^{-1/3}+m^{-1}}{2n^{-1/3}+2(1-2n)^{-1/3}-4m^{-1}}.$$

The desired critical dependence $W=W(n)$, following from system (14), is represented in Fig. 4b by curve *1*. We introduce the notation

$$A^*=4n^{-1/3}+8(1-2n)^{-1/3}-8/m,$$

$$B^*=\frac{4}{3}\left(\frac{k}{n^{4/3}}+\frac{2(1-2k)}{(1-2n)^{4/3}}\right),$$

$$C^*=\frac{8}{9}\left(\frac{k^2}{n^{7/3}}+\frac{2(1-2k)^2}{(1-2n)^{7/3}}\right),$$

$$D^*=\frac{1}{9}[n^{-4/3}+(1-n)^{-4/3}].$$

With this simplifying notation, the restriction on the region of existence of possible channels of breakup can be written in the form

$$W>D^*A^*[A^*C^*-B^{*2}]^{-1}. \quad (15)$$

The region of values of the parameters W and n in which breakup is possible lies above curve *2* in Fig. 4b.

Doubling the distance between the daughter droplets, we obtain curves *3* and *4* in Fig. 4b. Increasing the distance between the daughter droplets by tenfold, we obtain curves *5* and *6* in Fig. 4b. Comparing Figs. 4a and 4b, it is clear that in both these idealizations the results are similar: with growth of the initial deformation of the parent drop (for large m) symmetric breakup of the drop becomes possible. In this case, in the model of an ideally conducting drop, the value of the Rayleigh parameter W at which breakup occurs is observed to grow. In the limit of very large deformations $W\to 1.5$, so that such breakup (breakup of the drop into droplets of comparable dimensions) becomes problematic, since for $W=1$ the Rayleigh instability is observed for an initially spherical drop for which the charge of the parent drop is dispersed in the form of a series of highly dispersed, highly charged droplets.[1,2] Nevertheless, it can be realized under conditions of extremely rapid deformations of the drop, when its viscosity is high and the formation of emitting bulges is hindered.[7] It is also clear from general physical considerations that strongly charged drops can break up in this way in a highly viscous liquid medium (as was noted in the experiments in Ref. 11). In this case the high viscosity of the drop itself or the medium damps the instability of the higher modes, which hinders the formation of Taylor cones and the development of the instability via the Rayleigh mechanism.[7]

c) *Volume-charged insulating drop.* We assume that the charge is uniformly distributed throughout the entire volume (frozen into the material) of the initial drop with volume density ρ. Then the electrostatic energy of this drop is given by the relation derived in Sec. 1c.

After breakup of the initial drop the potential energy of the system of three daughter droplets is given by

$$U=\frac{2q_1^2}{2r_1}\left(1+\frac{1}{5\varepsilon}\right)+\frac{q_2^2}{2r_2}\left(1+\frac{1}{5\varepsilon}\right)+8\pi\sigma r_1^2+4\pi\sigma r_2^2+\frac{2q_1q_2}{2r}.$$

We assume in accordance with the assumption of frozen-in charge in the material that the charge divides in the same ratio as the volume $q_1=Qn$, $q_2(1-2n)Q$. Then for the energy of the final state of the system upon breakup of the parent drop into three daughter droplets we obtain

$$5[n^{2/3}-(1-2n)^{2/3}]\left(1+\frac{1}{5\varepsilon}\right)W+[n^{-1/3}-(1-2n)^{-1/3}]+3W(1-4n)\frac{R}{r}=0.$$

Setting $\varepsilon=81$ for definiteness (as in water), we find

$$W=\frac{[n^{-1/3}-(1-n)^{-1/3}]}{5[n^{2/3}-(1-2n)^{2/3}]\left(1+\frac{1}{5\varepsilon}\right)+\frac{3(1-4n)}{m}}.$$

A graph of this function is plotted by curve *1* in Fig. 4c. From the condition on the second derivative we obtain

$$W>\frac{n^{-4/3}+2(1-2n)^{-4/3}}{10\left(1+\frac{1}{5\varepsilon}\right)[n^{-1/3}+2(1-2n)^{-1/3}]-\frac{36}{m}}. \quad (14a)$$

The region in Fig. 4c in which breakup is possible is located above curve *2*. From the mutual arrangement of curves *1* and *2* in Fig. 4c it is clear that for $r=r_1+r_2$ the drop will not break up into three daughter droplets.

Let us consider the case in which the initial drop is strongly deformed, i.e., when $r \gg r_1+r_2$. As the calculations show, breakup becomes possible only at sufficiently large deformations, when m is 28 times larger than its minimum possible value (the corresponding curves are curves *3* and *4* in Fig. 4c). It is easy to see that upon breakup the ratio of volumes of the daughter droplets $n \approx 0.27$, i.e., breakup has a substantially symmetric character ($V_1 \approx V_2$). With further increase of the initial deformation, the degree of symmetry of the breakup also increases.

Thus it is clear from the above-said that for breakup of a surface-charged non-deformed or weakly deformed drop, breakup can occur only asymmetrically. In this case the ratio of volumes of the daughter droplets is found to depend on the parameter W. The most probable outcome is breakup for which $V_1 \ll V_2$. Such breakup becomes possible for a weakly deformed drop already at $W_{\min}=0.5$ ($W_{\min}$ is defined for $n \gg 0.1$). For strongly deformed surface-charged insulating drops, with growth of the degree of deformation the probability of symmetric breakup is increased, but in this case asymmetric breakup into three daughter droplets predominates all the same. For an ideally conducting drop, with growth of the degree of deformation the probability of symmetric breakup is decreased. For extremely large values of the initial deformation m, breakup becomes possible already at $W_{\min}=0.25$. For a weakly deformed volume-charged insulating drop, breakup into three daughter droplets is generally not realized. For large deformations, such an event becomes possible. In this case, the volumes of the daughter droplets are roughly equal, i.e., for the give case symmetric breakup is more probable.

3. INFLUENCE OF PRIMARY DEFORMATION ON THE CRITICAL CHARACTERISTICS OF THE INSTABILITY OF STRONGLY CHARGED DROPS

Figure 5 plots the dependence $W(x)$ for the first and second derivatives for the case in which the charge is uniformly distributed over the surface of an insulating drop, where x is a parameter defining the degree of deformation of the initial drop and is equal to the distance between the ends of the drops. Curves *1* and *2* correspond to breakup of the initial drop into two daughter droplets; curves *3* and *4* correspond to breakup into three daughter droplets (curves *1* and *3* correspond to setting the first derivative of the energy of the final state equal to zero, curves *2* and *4* define allowed channels of breakup). In the calculations we varied the magnitude of the primary deformation for fixed n, i.e., in the given case the volumes of each of the small daughter droplets are equal (by the mother droplet here we understand the drop from which one droplet splits off in the first case, and from which two droplets split off in the second). The smallest deformation of the initial drop was found from the calculation where in the final state after breakup into three droplets, the droplets are touching each other (this quantity, scaled by the radius of the initial drop, corresponds to the leftmost point on the abscissa). Figure 5 plots curves for $n=0.0001$ and the magnitude of the smallest primary deformation $x_{\min}=2.19$. From the mutual arrangement of curves *1* and *3* it can be seen that for the case of relatively small deformations (up to $x \approx 6$) breakup into three droplets is more probable, and for the deformation $x>6$ breakup into two droplets predominates. With growth of n the magnitude of the deformation necessary for breakup into two daughter droplets to predominate falls rapidly. Starting at roughly $n=0.175$ (Fig. 6), breakup into two droplets is observed to predominate over the entire range in x.

In the case when the charge is distributed throughout the volume of the drop, as was mentioned earlier, for a weakly deformed drop only breakup into two daughter droplets of roughly equal volumes is possible while breakup into three daughter droplets becomes possible only for very strong virtual deformations.

The third case of an ideally conducting drop is similar to the first case, but in this case only breakup into two droplets is observed already for $n>0.01$.

To summarize, for instability of a weakly deformed surface-charged drop, breakup into three daughter droplets with emission of two daughter droplets much smaller than the mother drop is energetically favored. Disintegration into

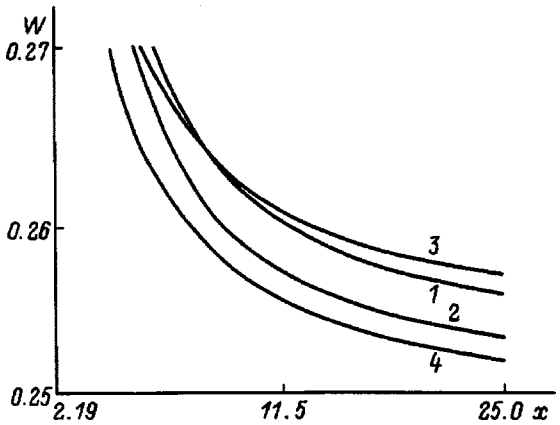

FIG. 5. The parameter W of a surface-charged drop of insulating liquid corresponding to extremal variation of the free energy of the system, plotted as a function of the deformation parameter x for $n=0.0001$. Locus of points of the ith curve lying above the $(i+1)$-th curve corresponds to minimum variation of the free energy.

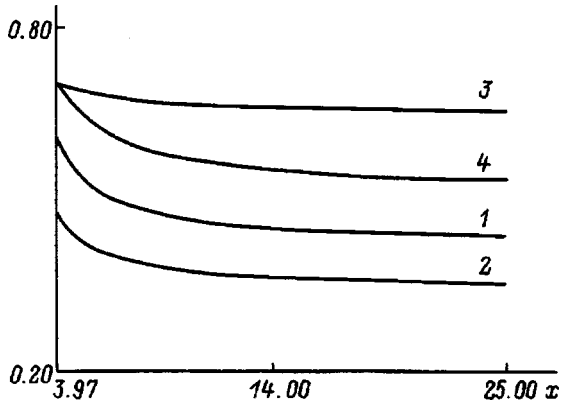

FIG. 6. Same as in Fig. 5, but for $n=0.175$.

two droplets becomes more probable for strong deformations. For emission of daughter droplets with volumes of the order of tenths of the volume of the mother drop, breakup into two droplets is energetically favored.

[1] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61**(3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].
[2] A. I. Grigor'ev and S. O. Shiryaeva, J. Phys. D **23**, 1361 (1991).
[3] J. R. Adam, N. R. Lindblad, and C. D. Hendricks, J. Appl. Phys. **39**, 5173 (1968).
[4] T. G. O. Berg, R. J. Trainor , Jr., and U. Vaughan, J. Atmos. Sci. **27**, 1173 (1970).
[5] S. A. Ryce and D. A. Patriarche, Can. J. Phys. **43**, 2192 (1965).
[6] A. I. Grigor'ev, I. D. Grigor'eva, and S. O. Shiryaeva, Khim. Plazmy, No. 17 [in Russian], Energoatomizdat, Moscow (1991), pp. 218–248.
[7] A. I. Grigor'ev and A. É. Lazaryants, Zh. Vychisl. Mat. Mat. Fiz. **32**, 929 (1992).
[8] A. I. Grigor'ev and A. É. Lazaryants, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza No. 5, 11 (1991).
[9] S. O. Shiryaeva and A. I. Grigor'ev, Zh. Tekh. Fiz. **62**, 35 (1992) [Sov. Phys. Tech. Phys. **37**, 254 (1992)].
[10] V. I. Krasnitskiĭ, A. M. Apasov, and S. M. Kontush, Pis'ma Zh. Tekh. Fiz. **16**(18), 77 (1990) [Tech. Phys. Lett. **16**(9), 717 (1990)].
[11] S. A. Ryce and R. R. Wyman, Can. J. Phys. **42**, 2185 (1964).
[12] A. I. Grigor'ev and E. I. Mukhina, Zh. Tekh. Fiz. **60**, 25 (1990) [Sov. Phys. Tech. Phys. **35**, 1015 (1990)].
[13] A. I. Grigor'ev, S. O. Shiryaeva, I. D. Grigor'eva *et al.*, Zh. Tekh. Fiz. **61**, 25 (1991) [Sov. Phys. Tech. Phys. **36**, 385 (1991)].

Translated by Paul F. Schippnick

Optical study of parameters of the passage of a shock wave from air to water

S. V. Gribin, I. I. Komissarova, G. V. Ostrovskaya, B. I. Spesivtsev, V. N. Filippov, and E. N. Shedova

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted April 23, 1997)
Zh. Tekh. Fiz. **68**, 39–43 (August 1998)

A study is made of the penetration of shock waves from air into water. The shock wave in air is generated as a result of dielectric breakdown induced by pulsed CO_2-laser radiation. A combination of the double-exposure shadow method and holographic interferometry is used to measure the shock-wave parameters. Density and pressure profiles behind the wave front are obtained at different times after onset of breakdown. It is shown experimentally that as the wave passes through the interface from the air to the water, there is a fourfold amplification of the pressure in the shock wave front. Estimates of the width of the shock wave front formed in the water are given in the context of studies of large-scale explosion processes. It is shown that simple empirical dependences, established in the course of studies of large-scale explosions, are also valid with certain corrections for microscopic laboratory experiments.
 [S1063-7842(98)00608-4]

INTRODUCTION

The efficacy of using shock waves in a liquid to solve various medical, technological, and environmental problems depends substantially on the spatiotemporal structure of the shock wave (pressure in its front, steepness of the front, density and pressure profiles behind the front, etc.). One way of forming shock waves in water, with steep shock fronts, is by generation of a primary shock wave in the air above the water surface. In this case, a steep shock front typical of such a wave in air can be expected along with amplification of the pressure in the shock front upon passage of the shock wave from air to water.[1]

In the present work the primary shock wave was generated by dielectric breakdown in air produced by focused radiation of a pulsed CO_2 laser. To study the parameters of the shock wave and its penetration into the water medium, we used a combination of the double-exposure shadow method[2] and holographic interferometry.

SHOCK WAVE STUDY BY THE DOUBLE-EXPOSURE SHADOW METHOD

Pulsed CO_2-laser radiation with an energy per pulse of ~ 4 J was focused with the help of a spherical mirror *1* (Fig. 1a) near the free surface of the water medium. The time dependence of the laser pulse was typical for a CO_2 laser, i.e., it consisted of a main peak with duration around 200 ns and a trailing edge stretching out to 2.5 μs. As a result of dielectric breakdown a laser spark was created at the mirror focus *3*, accompanied by generation of a shock wave *4*. The investigated process was probed in the direction perpendicular to the plane of the figure, in ruby laser light synchronized with the CO_2 laser.

The arrangement for obtaining double-exposure shadowgrams is described in detail in Ref. 3. The main circuit element of the setup is a light delay line providing a time shift (129 ns) between the two probe pulses. Figures 2a and 2b display shadowgrams illustrating the penetration of a shock wave into water at angles of incidence close to the normal (a — 2.3 μs and b — 4.7 μs after onset of breakdown). Figure 2c displays shadowgrams illustrating almost grazing incidence of a shock wave from air into water. Here ''irregular'' reflection of the shock wave and formation of a Mach wave are observed.[1] In this experiment the shock wave in air was initiated at the focus of laser radiation on the surface of a solid target. Shock-front images (Figs. 2a and 2b) corresponding to the first and second exposure are visible in the shadowgrams, and the speed of propagation of the shock wave can be determined from the distance between these images.

Figure 3a plots the time dependence of the small radius R and the velocity v_1 of the shock front in air, obtained by processing the shadowgrams, and Fig. 3b plots the time dependence of the pressure P_1 and density ρ_1 in the shock front, calculated from the formulas[1]

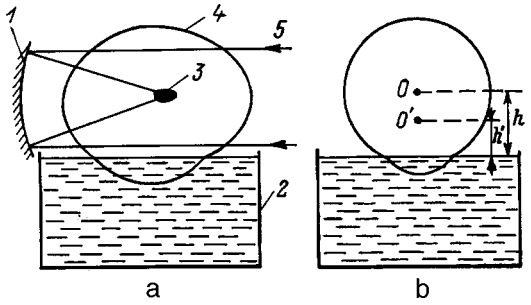

FIG. 1. a: Diagram of excitation of a shock wave: *1* — spherical mirror, *2* — cuvette filled with water, *3* — laser spark, *4* — shock wave, *5* — rays from CO_2-laser. b: Cross section of shock wave perpendicular to symmetry axis.

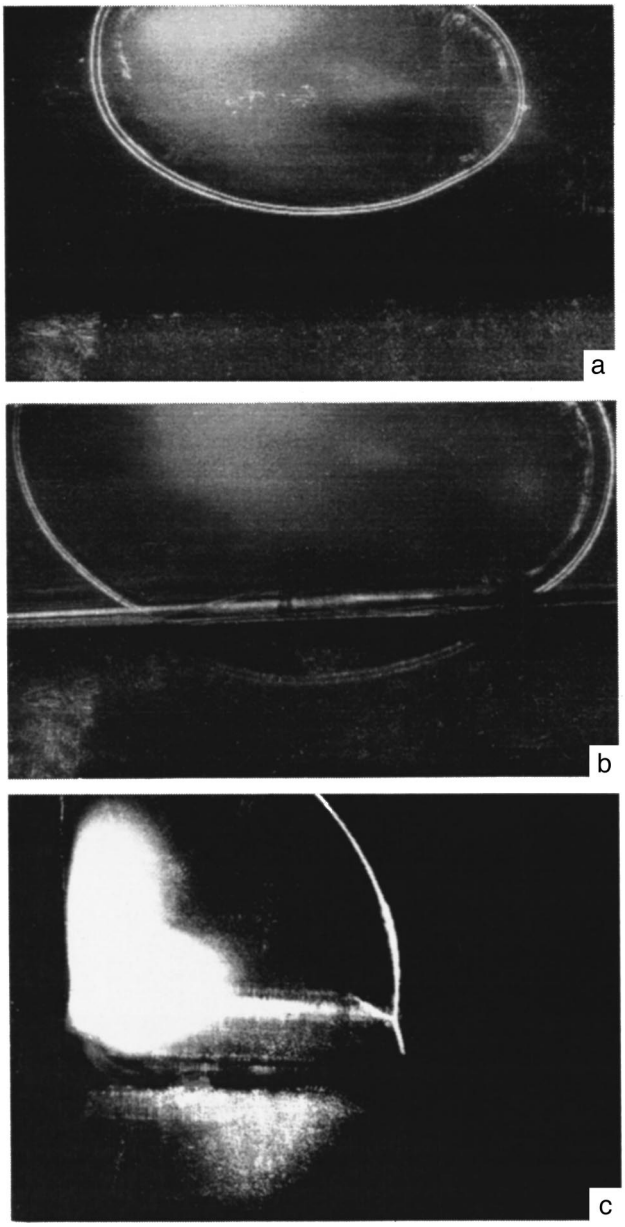

FIG. 2. Shadowgrams of shock wave at different instants of time (a,b) after breakdown, and reflection of the shock wave from the water surface for grazing incidence (c).

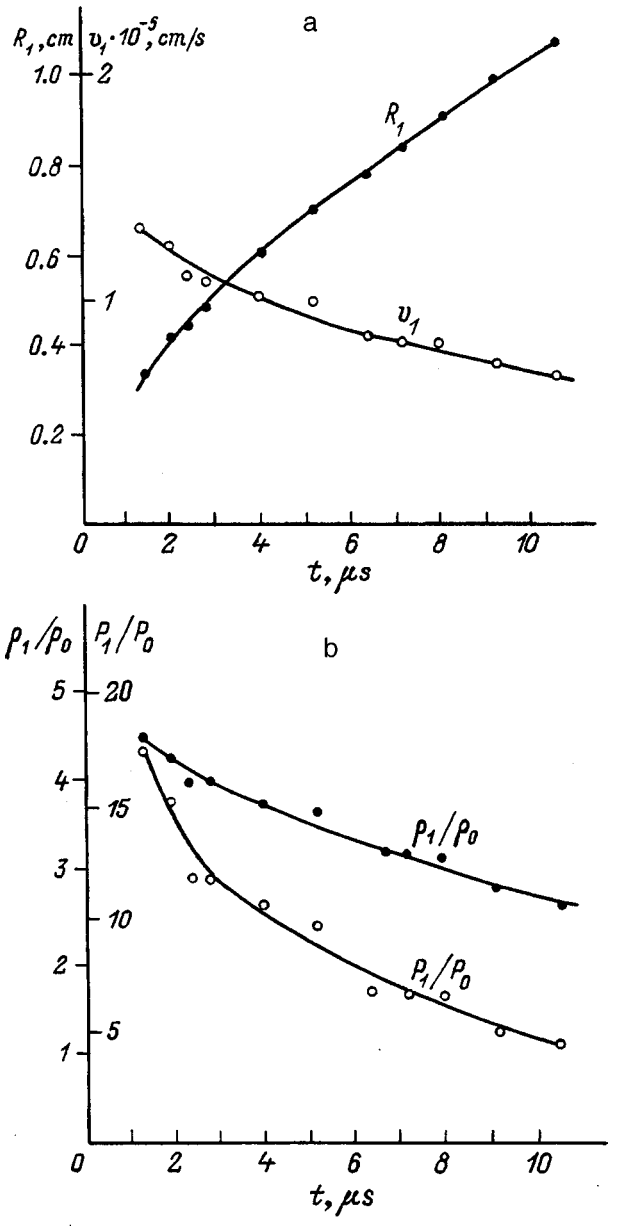

FIG. 3. a — variation with time of the small radius R_1 and velocity v_1 of the shock wave front, b — variation with time of the pressure P_1 and density ρ_1 in the shock wave front in air.

$$P_1 - P_0 = 2\rho_{01}(v_1^2 - a_{01}^2)/\gamma + 1, \tag{1}$$

$$\frac{\rho_1}{\rho_{01}} = \frac{(\gamma-1)+(\gamma+1)\dfrac{P_1}{P_0}}{(\gamma+1)+(\gamma-1)\dfrac{P_1}{P_0}}, \tag{2}$$

where P_0 and ρ_{01} are the normal pressure and density of air, a_{01} is the speed of sound in air, and γ is the adiabatic exponent (for air $\gamma = 1.4$).

In principle the pressure in the shock front in air P_2 can be calculated from the formula

$$P_2 - P_0 = \rho_{02} v_2 (v_2 - a_{02})/m \tag{3}$$

(where m is an empirical coefficient and ρ_{02} is the normal density of water) if the amount by which the speed of the front, v_2, exceeds the speed of sound in water, a_{02}, is known. However, in our experiments the speed of the front in water, measured from the shadowgrams, displayed hardly any difference from the speed of sound. Therefore, to determine the pressure in the shock front in water, and also to study the pressure and density fields of the shock wave in more detail, in air as well as in water, we used the method of holographic interferometry.

INTERFERENCE-HOLOGRAPHIC STUDY OF SHOCK WAVES

Holographic interferograms were recorded by the double-exposure method according to the scheme of focused

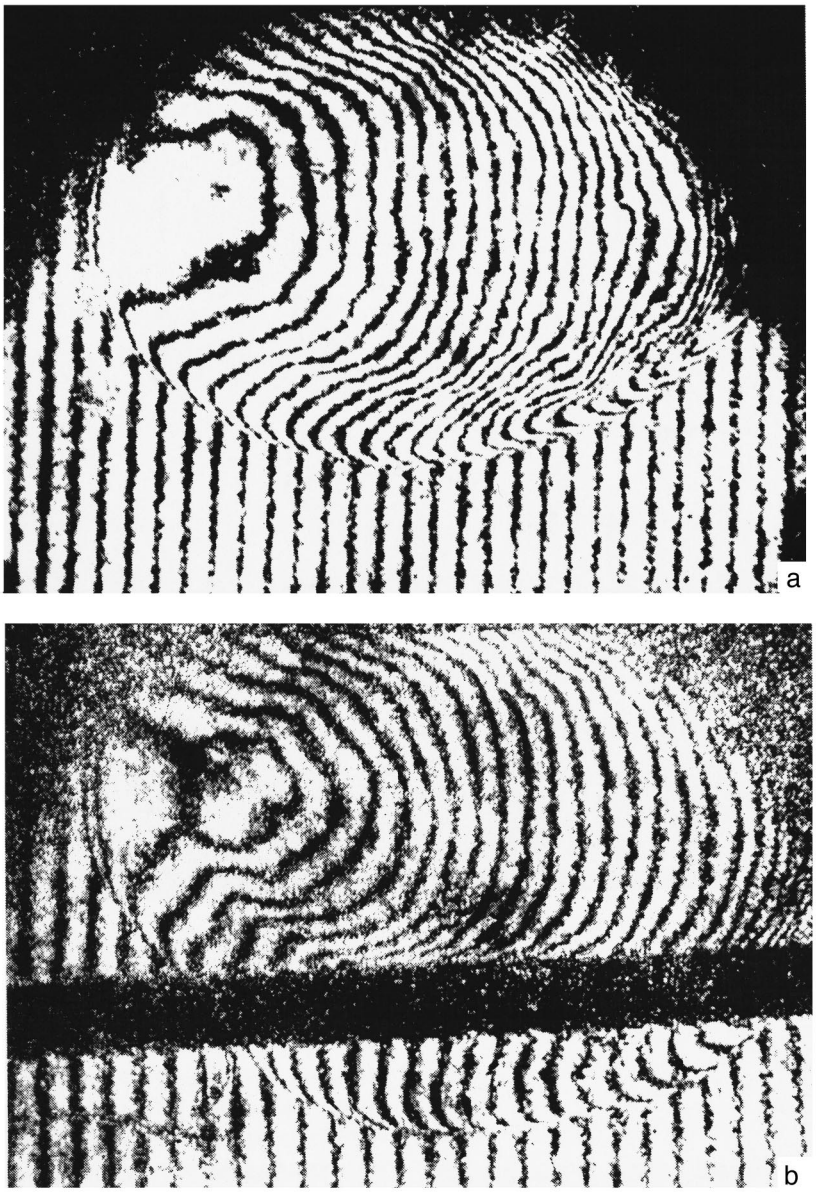

FIG. 4. Holographic interferograms of a shock wave in air.

images. The optical circuit of the setup is described in detail in Ref. 4. Figure 4 displays typical holographic interferograms of a shock wave in air (a — $t=4.7\ \mu$s) and after its penetration into water (b — $t=6\ \mu$s). On the front of the shock wave, where the gas density varies abruptly, there is a jump in the interference bands, which can lead to an error in determining their shifts by some integer number of bands. To eliminate this ambiguity, we estimated the magnitude of the jump of the bands near the shock front using data on the shock-wave parameters obtained by the shadow method.

For cylindrical symmetry, the relation between the band shift $k(x)$ and the radial density distribution $\rho(r)$ is given by the Abel transform

$$k(x)=\frac{2(n_0-1)}{\lambda}\int_x^R \frac{\left[\dfrac{\rho(r)}{\rho_0}-1\right]rdr}{\sqrt{r^2-x^2}}, \tag{4}$$

where x is distance measured from the symmetry axis of the object, n_0 is the index of refraction of air under normal conditions, and λ is the wavelength of the probe radiation.

The density profile in the immediate vicinity of the shock front can be approximated to first order by the expression

$$\rho(r)=\rho_1-\frac{\partial\rho}{\partial r}\Big|_{r=R}(R-r), \tag{5}$$

where ρ_1 is the density in the shock front, determined from shadow measurements, $\delta\rho(r)/\delta r$ is the density gradient behind the front, which can be calculated from the formula

$$\frac{1}{\rho_0}\frac{\partial\rho}{\partial r}=A(M)\frac{dM}{dr}+B(M)K \tag{6}$$

obtained by solving the system of gas-dynamic equations describing gas flow in the vicinity of the shock front.[1] Here $M=v_1/a_{01}$ is the Mach number and K is the curvature of the shock front. The functions $A(M)$ and $B(M)$ depend only on

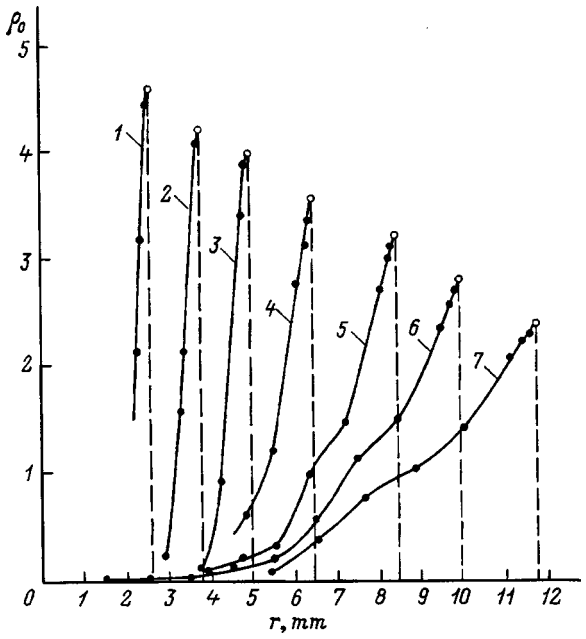

FIG. 5. Density profiles behind a shock wave front in air. t, μs: *1* — 1.1, *2* — 2.1, *3* — 3, *4* — 4.7, *5* — 7.8, *6* — 10, *7* — 13; r is the distance from the symmetry axis.

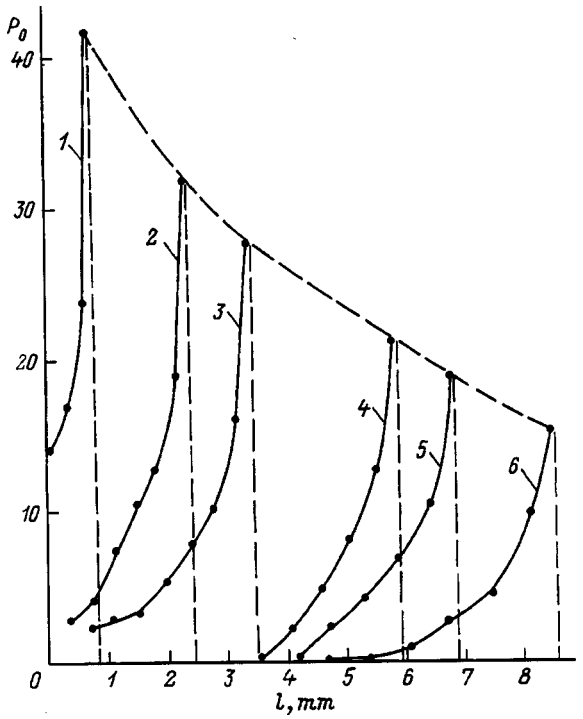

FIG. 6. Pressure profiles behind a shock wave front in water. t, μs: *1* — 4.6, *2* — 5.6, *3* — 6.6, *4* — 8.0, *5* — 8.8, *6* — 10; l is the distance from the water surface.

the shock-front parameters, which can be determined from the shadow measurements [a more detailed calculation of these functions and a derivation of formula (6) will be published later].

Substituting the approximation (5) in formula (4) and carrying out the integration, it is possible to calculate the course of the interference bands in the immediate vicinity of the shock front. The arrangement of the bands beyond the front is determined directly from the interferogram, after which, with the help of the inverse Abel transform it is possible to find the radial density distribution $\rho(r)$.

Results of interference measurements of $\rho(r)$ in a shock wave in air for a sequence of times are plotted in Fig. 5 (in the given sequence of experiments water was not present in the cuvette). It may be noted that local values of the index of refraction in inner regions of the laser spark, measured from interferograms corresponding to early stages of the process ($t<3$ μs) turned out to be less than unity, which indicates the presence of a significant electron density and prohibited us from determining the air density in the central regions of the laser spark from interferograms recorded in the light of one wavelength. For this reason, only values near the shock front obtained at early stages of the process are plotted in the density profiles.

The following series of interferograms, one of which is shown in Fig. 4b, was obtained with water present in the cuvette. The focal point 0 of the CO_2-laser radiation was located a distance $h\sim 6$ mm from the water surface (Fig. 1b). The band shifts in the interferograms in water were substantially smaller than for the waves in air, and ambiguities in the determination of the band shifts near the shock front did not arise. At the same time, interpretation of the interferograms in the given case was complicated by a disruption of the axial symmetry associated with an increase in the curvature of the wave front caused by the change in the wave speed as the wave entered the water. Nevertheless, the wave after entering the water can be considered as an axially symmetric object, whose $0'$ axis is shifted relative to its position 0 in air. It follows from the law of refraction (Snell's law) in the small-angle approximation that the $0'$ axis is offset from the water surface by the distance $h'=hv_1/v_2$, where v_1 and v_2 are the wave speeds in air and in water at the moment the wave penetrates the water. Processing the interferograms employing the assumptions made above, we obtained profiles of the pressure behind the wave front in air (Fig. 6) for different times reckoned from the time of breakdown in air (in the geometry of our experiment the wave entered the water at $t\approx 4$ μs).

DISCUSSION OF RESULTS

It is of definite interest to compare the observed rate of attenuation of an underwater shock wave (Fig. 6) with the empirical dependence of the pressure in the front on the radius of the wave, determined from a study of large-scale underwater explosions. As is well known,[1] the pressure in the front of a spherical underwater shock wave should attenuate as $r^{-1.13}$. The corresponding dependence is plotted in Fig. 6 by the dashed line. To calculate it we used the effective radius

$$R_{\text{ef}}=\frac{2R_1R_2}{R_1+R_2}, \tag{7}$$

where R_1 and R_2 are the principal radii of curvature of the shock wave.

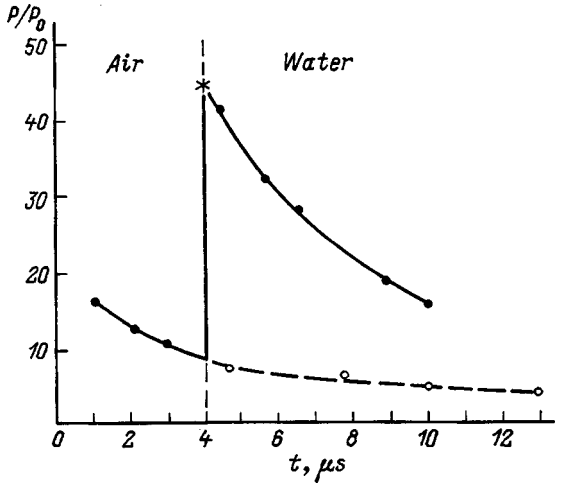

FIG. 7. Variation with time of the pressure in a shock wave front as the shock wave passes from air to water; * denotes the value of the pressure calculated according to the Izmaĭlov formula.

It can be seen that the given curve essentially coincides with the envelope of the pressure profiles.

It follows from the density and pressure profiles plotted in Figs. 5 and 6 that the shock waves both in air and in water have a very steep front. The spatiotemporal resolution of our interferograms, determined by the duration of the probe laser pulse (30 ns), affords only an approximate estimate of the growth time of the front, which does not exceed 50 ns.

From a comparison of the curves in Figs. 5 and 6 we constructed the time dependence (Fig. 7) of the pressure in the front of a shock wave traveling from air to water (solid lines). The dashed line plots the variation of the pressure in the wave front in the absence of water. As can be seen from Fig. 7, the pressure in the front grows by roughly fourfold as it passes through the air–water interface.

From hydrodynamics we know that the pressure in the wave front of a wave that has passed into water at normal incidence can be calculated from the Izmaĭlov formula[1]

$$P_2-P_0=2(P_1-P_0)+\frac{\dfrac{\gamma+1}{\gamma-1}(P_1-P_2)^2}{(P_1-P_0)+\dfrac{2\gamma}{\gamma-1}P_0}, \tag{8}$$

which for our case ($P_1\sim 9.5$ atm, $\gamma=1.4$) gives $P_2\sim 46$ atm, which is in fair agreement with the value $P_2=42$ atm obtained from experiment (especially if we take into account that due to the shadow of the meniscus the point at which the given quantity was measured was located a distance ~ 1 mm from the surface.

In summary, the combined application of the shadow method and the interference-holographic method has enabled us to obtain a quite complete picture of the variation of the parameters of a shock wave as it passes from air into water. It turns out that the nature of the attenuation of the pressure in the shock front, and also the amplification of the pressure in the passage of the wave from air into water can be described by simple empirical formulas established in a study of real underwater explosions. Possibly, laboratory studies similar to those reported in the present work can be used with certain corrections to model hydrodynamic processes accompanying large-scale explosions.

[1] Yu. S. Yakovlev, *Hydrodynamics of Explosions* [in Russian], Sudpromgiz, Leningrad, 1961.
[2] A. P. Dmitriev, G. V. Dreĭden, Yu. I. Ostrovskiĭ, and M. Étinberg, Zh. Tekh. Fiz. **53**, 311 (1983) [Sov. Phys. Tech. Phys. **28**, 191 (1983)].
[3] I. I. Komissarova, G. V. Ostrovskaya, V. N. Filippov, and E. N. Shedova, Zh. Tekh. Fiz. **67**, 138 (1997) [Tech. Phys. **42**, 247 (1997)].
[4] I. I. Komissarova, G. V. Ostrovskaya, V. N. Filippov, and E. N. Shedova, Zh. Tekh. Fiz. **62**, 34 (1992) [Sov. Phys. Tech. Phys. **37**, 130 (1992)].

Translated by Paul F. Schippnick

On the onset of free convection in vertical channels with cross sections in the form of circular or annular sectors

V. A. Kazhan

Moscow State University of Environmental Engineering, 127550 Moscow, Russia
(Submitted July 10, 1996)
Zh. Tekh. Fiz. **68**, 44–47 (August 1998)

The appearance of convection in a liquid filling a long vertical channel with ideally heat-conducting walls is considered in the case where the temperature of the liquid, originally at rest, decreases linearly with height and the cross section of the channel has the shape of a circular or annular sector. The critical Rayleigh numbers and eigenfunctions of the boundary problems are written out for the critical motions. The appearance of convection is also examined in a liquid-saturated porous medium filling a vertical channel with ideally heat-conducting or insulating walls and having a cross section in the form of a circular sector. © *1998 American Institute of Physics.* [S1063-7842(98)00708-9]

In the theory of free thermal convection in the Boussinesq approximation a class of solutions is known describing stationary convection in a liquid in a long vertical channel of constant cross section, heated from below, whose convection flow is everywhere directed along the channel.[1,2] Motions of the liquid of such kind are possible only if the Rayleigh number is equal to one of the eigenvalues of some boundary value problem. The system of equations and the boundary conditions constituting the statement of this problem make it possible to calculate the maximum temperature gradient (constant along the channel), which if exceeded the hydrostatic state of the liquid becomes unstable, and convection arises.

A number of works have considered the appearance of convection in channels of circular, annular, elliptical, and rectangular cross section in a vertical liquid layer between parallel infinite planes, and also the appearance of convective filtration, analogous in its physical mechanism, in a liquid-saturated porous medium located in vertical channels (see, for example, the literature cited in Ref. 2).

In the present paper we examine the appearance of convection in channels with cross sections having the shape of circular and annular sectors whose opening angle takes some sequence of values. We also consider the appearance of convective filtration in a porous medium filling a channel with a cross section having the shape of a circular sector.

1. Let a long vertical channel with cross section in the form of a circular sector of radius a and opening angle 2α be filled with a liquid at rest, whose temperature falls off linearly with height. We introduce the cylindrical coordinate system r, φ, z, whose z axis is directed opposite the force of gravity and passes through the vertex of the cross section of the channel and for which the ray $\varphi=0$ bisects the angle between the two line segments of the boundary of the cross section. In dimensionless variables, z-independent critical motions with velocity **v** having only a vertical component are described by the system of equations[2]

$$\Delta v + RT = \text{const}, \quad \Delta T + v = 0,$$

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2}, \tag{1}$$

where T is the perturbation of the initial linear temperature profile; $R = g\beta A a^4 (\nu\chi)^{-1}$ is the Rayleigh number; g is the acceleration due to gravity; β, ν, χ are respectively the coefficients of thermal expansion, kinematic viscosity, and thermal diffusion of the liquid; and finally $A>0$ is the vertical temperature gradient.

In the limiting case of ideally heat-conducting walls the boundary conditions have the form

$$v=0, T=0 \quad \text{for } r=1, -\alpha \leqslant \varphi \leqslant \alpha$$

$$\text{and} \quad \text{for } 0 \leqslant r \leqslant 1, \varphi = \pm\alpha. \tag{2}$$

In addition, the flux of liquid through any cross section of the channel should be equal to zero.

Eliminating the temperature perturbation from system (1), we have

$$\Delta\Delta v - Rv = 0.$$

It is not hard to write out a solution of this equation satisfying the requirement that the total flux of liquid along the channel be equal to zero,

$$v(r,\varphi) = \sin \lambda\varphi [C_1 J_\lambda(\gamma r) + C_2 I_\lambda(\gamma r)], \quad \gamma = R^{\frac{1}{4}}, \tag{3}$$

where $\lambda>0$, C_1 and C_2 are constants, and $J_\lambda(\gamma r)$ and $I_\lambda(\gamma r)$ are Bessel functions of the first kind.

From the physical point of view the temperature perturbation is caused by convection of the liquid, so that in the calculation of T it is necessary to take into account the solution of just the inhomogeneous equation obtained upon substitution of expression (3) in the second equation of system (1). As a result, we obtain

$$T(r,\varphi) = \gamma^{-2} \sin \lambda\varphi [C_1 J_\lambda(\gamma r) - C_2 I_\lambda(\gamma r)]. \tag{4}$$

Next, after substituting solutions (3) and (4) in the first equation of system (1), we find that the constant in this equation—the longitudinal pressure gradient[2]—is equal to zero.

From the boundary conditions on the flat walls of the channel (2) we find $\lambda=\pi m/\alpha$, $m=1,2,3,\ldots$. It is easy to see that the quantity m determines the azimuthal structure of the corresponding critical motion.

Upon substituting solutions (3) and (4) into the boundary conditions (2) in the cylindrical part of the wall of the channel we arrive at a system of two linear homogeneous algebraic equations in C_1 and C_2. The condition for the existence of a nontrivial solution of this system is that its determinant be equal to zero. This condition takes the form $J_\lambda(\gamma)I_\lambda(\gamma)=0$. Since for real $\lambda>0$ the zeros of the function $I_{\lambda(x)}$ are complex, the spectrum of critical Rayleigh numbers $R^{(0)}_{\lambda,s}$, $s=1,2,3,\ldots$, is determined by the zeros of the function $J_\lambda(x)$: we have $R^{(0)}_{\lambda,s}=j^4_{\lambda,s}$, where $j_{\lambda,s}$ is the sth root of the equation $J_\lambda(x)=0$.

The eigenfunctions of problem (1) and (2) corresponding to these eigenvalues are written as follows:

$$v_{\lambda,s}(r,\varphi)=J_\lambda(j_{\lambda,s}r)\sin\lambda\varphi,$$

$$T_{\lambda,s}(r,\varphi)=j^{-2}_{\lambda,s}v_{\lambda,s}(r,\varphi). \tag{5}$$

We will consider channels differing from one another in the size of the dihedral angle between the flat walls: $\alpha_n=2\pi(2n+1)^{-1}$, $n=1,2,3,\ldots$. For such channels for even m the order of the Bessel function is equal to an integer $\lambda=m(n+\frac{1}{2})$, and for odd m it is equal to a half integer, so that for odd m the solutions (5) are expressed in terms of elementary functions.[3] In both cases the eigenvalues are easily calculated with the help of tables of zeros of the Bessel functions of half-integer and integer orders.[3,4]

For

$$\lambda=m\left(n+\frac{1}{2}\right)$$

and fixed n, the smallest critical Rayleigh number $R^{(0)}_{\lambda,s}$ obtains for $m=1$ and is equal to $R^{(0)}_{n+\frac{1}{2},1}=j^4_{n+\frac{1}{2},1}$, the first root of the equation $J_{n+\frac{1}{2}}(x)=0$. Tabulated data[4] allow us to find $R^{(0)}_{n+\frac{1}{2},1}$ for each α_n for $n=1,2,3,\ldots,20$.

From Table I, which gives the values of $R^{(0)}_{n+\frac{1}{2},1}$ for various values of n, it can be seen that as α_n decreases, the threshold for the appearance of convection grows. Physically, this is explained by an increase (as a consequence of a closing together of the ideally heat-conducting walls) in the intensity of decay of the temperature perturbations arising in the liquid.

By way of an example, Fig. 1 displays a diagram of critical motions realized successively as the levels of the critical Rayleigh numbers in a channel with $\alpha_1=2\pi/3$ are raised. The radii and arcs marked off in these figures inside the given sector are node lines [on which expressions (5) vanish] dividing the regions of upwelling and downwelling fluxes.

The corresponding study for channels with cross section in the form of a circular sector with opening angle $2\alpha_n$,

TABLE I. Smallest eigenvalues $R^{(0)}_{n+\frac{1}{2},1}$, $R^{(p)}_{n+\frac{1}{2},1}$ for a channel with cross section in the form of a circular sector, and value of the coefficient $\sigma_{n+\frac{1}{2},1}$ of the powers ε in the expression for the relative correction for an annular sector.

n	α_n	$R^{(0)}_{n+\frac{1}{2},1}$	$\sigma_{n+\frac{1}{2},1}$	$R^{(p)}_{n+\frac{1}{2},1}$
1	$\frac{2\pi}{3}$	408	28	20.19
2	$\frac{2\pi}{5}$	1103	108	33.22
3	$\frac{2\pi}{7}$	2384	338	48.83
4	$\frac{2\pi}{9}$	4483	960	66.95
5	$\frac{2\pi}{11}$	7662	2560	87.53
6	$\frac{2\pi}{13}$	12214	6534	110.52
7	$\frac{2\pi}{15}$	18465	16155	135.89
8	$\frac{2\pi}{17}$	26766	38952	163.60
9	$\frac{2\pi}{19}$	37500	91943	193.65
10	$\frac{2\pi}{21}$	51079	213631	226.01

where $\alpha_n=\pi/n$, $n=1,2,3,\ldots$, takes an analogous tack. Note that in this case for $n=1,2,3$ the smallest critical Rayleigh numbers and diagrams of the corresponding critical motions can be found among the results presented in Ref. 2 for a channel of circular cross section.

2. Let us consider the question of the appearance of convection in a vertical channel with ideally heat-conducting walls and having a cross section in the form of an annular sector with opening angle $2\alpha_n$, $\alpha_n=2\pi(2n+1)^{-1}$, $n=1,2,3,\ldots$, outer radius a and inner radius equal to εa, $\varepsilon<1$. In this case, in addition to boundary conditions (2) the additional condition that the eigenfunctions vanish at $r=\varepsilon$ is imposed. Since the origin of the cylindrical coordinate system does not belong to the region under consideration, the representation of the desired solution of system (1) involves

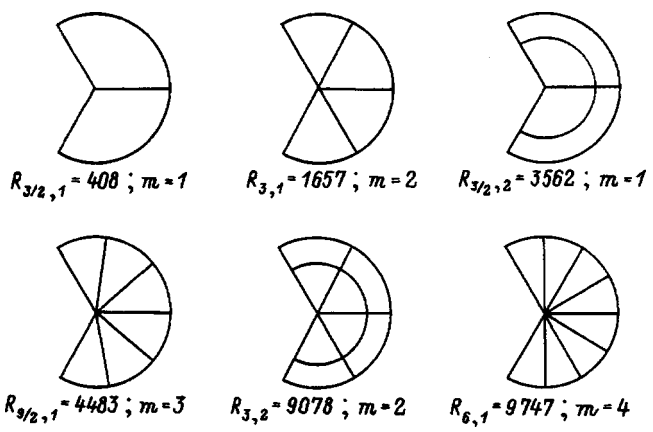

FIG. 1. Diagram of critical motions in a channel with $\alpha_1=2\pi/3$.

Bessel functions of the second kind $Y_\lambda(\gamma r), K_{\lambda(\gamma r)}$ as well as of the first kind $J_\lambda(\gamma r), I_\lambda(\gamma r)$, where as before

$$\lambda = m\left(n+\frac{1}{2}\right),$$

$m=1,2,3,\ \ldots\ ,$

$$v(r,\varphi)=\sin\lambda\varphi[C_1J_\lambda(\gamma r)+C_2I_\lambda(\gamma r)+C_3Y_\lambda(\gamma r)+C_4K_\lambda(\gamma r)],$$

$$T(r,\varphi)=\sin\lambda\varphi[C_1J_\lambda(\gamma r)-C_2I_\lambda(\gamma r)+C_3Y_\lambda(\gamma r)-C_4K_\lambda(\gamma r)], \quad (6)$$

where $\gamma = R^{\frac{1}{4}}$.

The existence of a nontrivial solution of the system of four linear homogeneous algebraic equations obtaining upon substitution of expressions (6) in the boundary conditions for $r=1$ and $r=\varepsilon$ require that the determinant of this system be equal to zero:

$$[J_\lambda(\gamma)Y_\lambda(\varepsilon\gamma)-J_\lambda(\varepsilon\gamma)Y_\lambda(\gamma)] \times[I_\lambda(\gamma)K_\lambda(\varepsilon\gamma)-I_\lambda(\varepsilon\gamma)K_\lambda(\gamma)]=0.$$

For real γ the expression inside the second pair of brackets does not vanish, so the eigenvalues of the problem under consideration are defined as the roots $\gamma_{\lambda,s}$, $s=1,2,3,\ \ldots$ of the transcendental equation

$$J_\lambda(\gamma)Y_\lambda(\varepsilon\gamma)-J_\lambda(\varepsilon\gamma)Y_\lambda(\gamma)=0. \quad (7)$$

The eigenvalues $R_{\lambda,s}=\gamma_{\lambda,s}^4$ correspond to the eigenfunctions

$$v_{\lambda,s}(r,\varphi)=\sin\lambda\varphi[J_\lambda(\gamma_{\lambda,s})Y_\lambda(\gamma_{\lambda,s}r)-Y_\lambda(\gamma_{\lambda,s})J_\lambda(\gamma_{\lambda,s}r)],$$

$$T_{\lambda,s}(r,\varphi)=\gamma_{\lambda,s}^{-2}v_{\lambda,s}(r,\varphi).$$

With regard to the case of small ε we calculate the smallest ($m=1$, $s=1$) critical Rayleigh number corresponding to some fixed n. We set $\gamma_{n+\frac{1}{2},1}=j_{n+\frac{1}{2},1}+\delta_{n+\frac{1}{2},1}$, where $j_{n+\frac{1}{2},1}$ is the first root of the equation $J_{n+\frac{1}{2}}(x)=0$. Taking $\delta_{n+\frac{1}{2},1}$ to be small, we expand the functions $J_{n+\frac{1}{2}}(\gamma_{n+\frac{1}{2},1})$ and $Y_{n+\frac{1}{2}}(\gamma_{n+\frac{1}{2},1})$ in Eq. (7) in Taylor's series in the vicinity of the point $j_{n+\frac{1}{2},1}$, and we expand the functions $J_{n+\frac{1}{2}}(\varepsilon\gamma_{n+\frac{1}{2},1})$ and $Y_{n+\frac{1}{2}}(\varepsilon\gamma_{n+\frac{1}{2},1})$ in power series. Taking only the leading terms of the expansion of the left-hand side of Eq. (7), we have

$$\frac{1\cdot3\cdot5\ldots(2n-1)}{(\varepsilon j_{n+\frac{1}{2},1})^{n+\frac{1}{2}}}J'_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1})\,\delta_{n+\frac{1}{2},1} = \frac{(\varepsilon j_{n+\frac{1}{2},1})^{n+\frac{1}{2}}}{1\cdot3\cdot5\ldots(2n+1)}Y_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1}). \quad (8)$$

We also make use of the expression[3] for the Wronskian for the spherical Bessel functions of the first kind $j_n(x)$, $y_n(x)$

$$j_n(x)y_n'(x)-j_n'(x)y_n(x)=x^{-2}, \quad (9)$$

where

$$j_n(x)=\sqrt{\frac{\pi}{2x}}J_{n+\frac{1}{2}}(x), \quad y_n(x)=\sqrt{\frac{\pi}{2x}}Y_{n+\frac{1}{2}}(x).$$

Since $J_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1})=0$, from Eq. (9) we find, setting $x=j_{n+\frac{1}{2},1}$,

$$Y_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1})=-\frac{2}{\pi j_{n+\frac{1}{2},1}J'_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1})}.$$

Substituting this expression in Eq. (8), we obtain

$$\delta_{n+\frac{1}{2},1}=\frac{2\qquad j_{n+\frac{1}{2},1}^{2n-1}}{\pi(2n+1)(1\cdot3\cdot5\ldots(2n-1)J'_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1}))^2}.$$

As a result, in the linear approximation in $\delta_{n+\frac{1}{2},1}$ we have

$$R_{n+\frac{1}{2},1}=(j_{n+\frac{1}{2},1}+\delta_{n+\frac{1}{2},1})^4=R^{(0)}_{n+\frac{1}{2},1}(1+\sigma_{n+\frac{1}{2},1}\varepsilon^{2n+1}),$$

where

$$\sigma_{n+\frac{1}{2},1}=\frac{8\qquad j_{n+\frac{1}{2},1}^{2n-1}}{\pi(2n+1)(1\cdot3\cdot5\ldots(2n-1)J'_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1}))^2}.$$

Thus in the case of ideally heat-conducting walls the stability threshold of the hydrostatic state of the liquid in a channel with cross section in the form of an annular sector with outer radius a and inner radius εa is raised in comparison with a channel having a cross section in the form of a circular sector of radius a with the same opening angle. For given n and ε the relative correction

$$(R_{n+\frac{1}{2},1}-R^{(0)}_{n+\frac{1}{2},1})/R^{(0)}_{n+\frac{1}{2},1}=\sigma_{n+\frac{1}{2},1}\varepsilon^{2n+1}$$

can be calculated with the help of tables[3,4] of the values of $j_{n+\frac{1}{2},1}$ and $J'_{n+\frac{1}{2}}(j_{n+\frac{1}{2},1})$. As an example, in Table I we give the numerical values of the coefficients $\sigma_{n+\frac{1}{2},1}$ for some values of n.

3. Let us now consider the appearance of convection in a liquid-saturated porous medium filling a vertical channel with cross section in the form of a circular sector when the temperature inside the channel decreases linearly with height and the liquid is initially at rest. From the equations of convective filtration of an incompressible liquid in a porous medium it follows that the z-independent motions whose velocities are vertical are described by a system of differential and algebraic equations[2]

$$\Delta T+v=0, \quad R^{(p)}T-v=0, \quad R^{(p)}=\frac{g\beta a^2}{\nu\chi}KA, \quad (10)$$

where K is the permeability coefficient of the porous medium.

By virtue of the absence in system (10) of derivatives of the velocity, a boundary condition on v is not required. Thus, for ideally heat-conducting walls the first equality of boundary conditions (2) must be eliminated. For the case $\alpha_n=2\pi(2n+1)^{-1}$, $n=1,2,3,\ \ldots$, in analogy to the preceding case we find that the eigenvalues are given by the roots $j_{\lambda,s}$, $s=1,2,3,\ \ldots$, of the equations $J_\lambda(x)=0$, where

$$\lambda = m\left(n+\frac{1}{2}\right), \quad m = 1,2,3,\ldots .$$

In this case we obtain $R^{(p)}_{\lambda,s} = j^2_{\lambda,s}$. The eigenfunctions have the form

$$T_{\lambda,s}(r,\varphi) = J_\lambda(j_{\lambda,s}r)\sin\lambda\varphi. \tag{11}$$

The smallest eigenvalues for some n (corresponding to the root $j_{n+\frac{1}{2},1}$) are given in Table I.

In the case of insulating walls the boundary conditions are written in the following form:

$$r = 1, -\alpha \leqslant \varphi \leqslant \alpha: \quad \frac{\partial T}{\partial r} = 0;$$

$$0 \leqslant r \leqslant 1, \varphi = \pm\alpha: \quad \frac{\partial T}{\partial \varphi} = 0. \tag{12}$$

It is not hard to find that for $\alpha_n = \pi(2n+1)^{-1}$, $n = 0, 1, 2, \ldots$, the eigenvalues of problem (10) and (12) as before are equal to the squares of the roots of the equation $J_\lambda(x) = 0$, but in contrast to the preceding case now

$$\lambda = (2m+1)\left(n+\frac{1}{2}\right), \quad m = 0,1,2,\ldots .$$

Thus, all eigenvalues of this problem are expressed, according to Eq. (11), in terms of Bessel functions of half-integer order. For fixed n the smallest eigenvalue is found for $m = 0$.

Note that the case $n = 0$ corresponds to a circular channel, inside which is found an insulating radial partition running from the cylindrical wall to the axis of the channel. For such a geometry the smallest eigenvalue $R^{(p)}_{\frac{1}{2},1} = 9.87$. As for the cases $n > 0$, for any n for a channel with $\alpha_n = \pi(2n+1)^{-1}$ and having insulating walls the smallest eigenvalue is equal to the smallest eigenvalue for a channel with $\alpha_n = 2\pi(2n+1)^{-1}$ and having ideally heat-conducting walls (see Table I).

[1] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed. [Pergamon Press, Oxford, 1987; Nauka, Moscow, 1988, 736 pp.].
[2] G. Z. Gershuni and E. M. Zhukhovitskiĭ, *Convective Stability of Incompressible Liquids* [in Russian], Nauka, Moscow (1972), 392 pp.
[3] M. Abramowitz and I. A. Stegun (Eds.), *Handbook of Mathematical Functions* [Dover, New York, 1965; Nauka, Moscow, 1979, 832 pp.].
[4] F. W. Olver (Ed.), *Bessel Functions. Part III. Zeros and Associated Values*, Royal Society Mathematical Tables, Vol. 7 [Cambridge University Press, Cambridge, 1960; Vyp. 44, Vychisl. Tsentr Akad. Nauk SSSR, Moscow, 1967, 95 pp.].

Translated by Paul F. Schippnick

Modeling of breakdown of a gas by electrons of the boundary layer during irradiation of metal targets by picosecond laser pulses

A. V. Ivlev, M. A. Yakovlev, and A. N. Bordenyuk

N. É. Bauman Moscow State Technical University, 107005 Moscow, Russia
(Submitted February 28, 1997)
Zh. Tekh. Fiz. **68**, 48–53 (August 1998)

Results are presented from a numerical simulation of the breakdown of a dense inert gas by boundary-layer electrons upon irradiation of a metal target by high-power picosecond laser pulses. It is shown that taking the electric field of the boundary layer into account leads to a substantial increase in the concentration of the seed electrons near the target surface, which accelerates the ionization process. The dependence of the breakdown time on the electric field intensity of the incident wave and the concentration of the gas atoms is derived. © *1998 American Institute of Physics.* [S1063-7842(98)00808-3]

INTRODUCTION

One of the most important results of the effect of emission electrons on the processes of interaction of electromagnetic radiation with condensed matter is the effect of low-threshold rapid ionization of a gas at high pressure ($p \approx 100$ atm). As experiments have shown,[1] the intensity of electromagnetic irradiation in this case is several orders of magnitude lower than the threshold intensity necessary for breakdown of the gas far from the target.

A numerical simulation of the ionization of a gas by emission electrons was performed by Mazhukin *et al.*[2,3] According to their results, at a radiant intensity $I_{em} \sim 10^{13}$ W/m^2 and nitrogen pressure $p \sim 100$ bar during a time ~ 1 ns a region of highly ionized plasma should be observed to form near the target surface, screening it from the incident radiation.

This is an appropriate point at which to make some remarks about the physical model, on which the calculations in Refs. 2 and 3 were based. As Mazhukin *et al.* note in their later work,[4] their previous work neglected the effect of unneutralized positive charge forming on the metal surface and creating a strong electrostatic field. This field attracts emission electrons back to the surface, thanks to which an electron boundary layer (EBL) is formed near the metal surface. This layer has a significant influence on the penetration of high-frequency electromagnetic radiation into the metal.[5,6] Note that an EBL exists near the surface of any conductor; however, at temperatures $T \sim 10^2 - 10^3$ K the influence of this layer on processes near the conductor is inconsequential. The point is that at such temperatures electrons in the conductor are highly degenerate and the electron concentration falls off with distance from the surface very rapidly, $n_e \propto z^{-2} \exp(-\beta z)$, where β^{-1} is a quantity of the order of the mean interelectron distance in the metal.[7] This makes it possible to consider the EBL of the degenerate electrons as being as thin as desired. However, with this rapid falloff of the electron concentration the degree of degeneracy falls, and when the Fermi energy $E_F(n_e)$ becomes of the order of kT, the electrons pass over to the classical state. Further decrease of the electron concentration follows a much smoother law[8,9]

$$n_e(z) = n_0 \left(1 + \frac{z}{\sqrt{2} D} \right)^{-2}, \tag{1}$$

where $D = (\varepsilon_0 kT / e^2 n_0)^{1/2}$ is the Debye screening radius and n_0 is the boundary-layer concentration, defined by the condition of removal of degeneracy: $E_F(n_0) \simeq kT$, i.e., $n_0 \propto T^{3/2}$. Thus, growth of the electron temperature is accompanied by an increase in the size of the EBL region where the concentration falls off according to dependence (1). At the same time, the boundary-layer value of the concentration n_0 also grows, and when temperatures of the order of the Fermi energy in the metal are reached, all the EBL electrons pass over to a classical state. In this case their distribution obeys dependence (1), and n_0 becomes of the order of the electron concentration in the metal. Consequently, upon intense nonequilibrium heating of the electron component of the metal an extended layer of seed electrons having a high concentration can be formed near its surface. The rate of ionization processes due to EBL electrons should therefore be substantially higher than in the model of free thermal-electron emission in the metal. In addition, the electric field of the EBL can strongly alter the profile of the electron and ion concentrations near the surface.

In addition to the indicated regularities, it should also be noted that Refs. 2 and 3 used the radiative transfer equation; however, as will be shown below, a region of highly ionized plasma is formed in regions of maximum amplitude of the standing electromagnetic wave—therefore, it is necessary to employ Maxwell's equations in the calculations.

MATHEMATICAL MODEL

This work presents results of a numerical simulation of breakdown of a dense gas by EBL electrons upon irradiation of a metal target by high-power laser pulses of picosecond duration. In light of the fact that the initial electron concentration on the surface is very large, it is of interest to answer the question of what should the threshold values of the near-

surface gas density and laser radiation intensity be, such that very rapid near-surface ionization of the gas can take place, thereby screening the irradiated target. By the phrase ''very rapid ionization'' we mean breakdown of the gas, whose onset time τ_I is much less than the characteristic electron–lattice relaxation time in a conductor $\tau_l \sim 10^{-10}$ s, i.e., in the regime of very rapid ionization the lattice temperature at the onset of strong surface screening remains at the same order-of-magnitude level as it started with, as a result of which the target material is not damaged in any way. For this reason, in the present work we consider substantially higher radiant intensities than in Refs. 2 and 3. The problem is treated in a one-dimensional formulation. The calculations make use of the following assumptions.

1. The gas above the target is inert; therefore, ionization and recombination processes cannot lead to the formation of molecular compounds. Formation of molecular ions of the sort A_2^+ is not taken into account, since under the conditions of the problem the coefficient of dissociative recombination involving such ions is substantially less than the coefficient of impact–radiative recombination in ternary collisions.[10]

2. As will become clear from the calculations that follow, the degree of ionization of the gas up to the onset of screening of the target does not exceed 10^{-2}. Therefore, the number of excited atoms in the gas is small and ionization of atoms only from the ground state was taken into account in the calculations.

To describe the ionization kinetics of the gas, we solved a system of equations combining the heat conduction equations for the electron temperature, the continuity equations for the electron and ion components, and Maxwell's equations for the electric field of the electromagnetic wave and the field of the unneutralized space charge. The mechanism of ionization of the gas due to the multiphoton photoelectric effect was not taken into account, since the characteristic onset time of breakdown under the conditions of the problem is of the order of $10^{-9}-10^{-10}$ s, which is much greater than the characteristic time of ionization due to EBL electron impact.[10]

The heat conduction equation in the region $z<0$ (metal) has the form

$$C_m \frac{\partial T_e}{\partial t} = \frac{\partial}{\partial z}\left[\chi_m \frac{\partial T_e}{\partial z}\right] - \alpha(T_e - T_l) + \kappa_r \kappa_i k_0 \varepsilon_0 c |E_0|^2 \exp(2k_0 \kappa_i z), \quad (2)$$

where T_e and T_l are the electron temperature and lattice temperature of the metal, C_m and χ_m are the specific heat and thermal conductivity of the electrons, κ_r and κ_i are the real and imaginary part of the complex refractive index in the metal, E_0 is the amplitude of the wave field at $z=0$, and α is the coefficient of heat exchange of the electrons with the lattice. For typical metals, where $n_m \sim 10^{28}\mathrm{m}^{-3}$, the coefficient $\alpha \sim 10^{16}\mathrm{W}\cdot\mathrm{m}^{-3}\cdot\mathrm{K}^{-1}$ (Ref. 11).

The equation of heat conduction in the gas ($z>0$) has the form[12]

$$\frac{3}{2}k\frac{\partial T_e}{\partial t} = \frac{1}{n_e}\frac{\partial}{\partial z}\left[\chi_e \frac{\partial T_e}{\partial z}\right] - \frac{3m}{M}k(T_e - T_a)\nu_e - \left(I + \frac{3}{2}kT_e\right)\nu_I + \frac{e^2|E|^2\nu_e}{2m(\omega^2+\nu_e^2)}, \quad (3)$$

where M is the mass of an atom of the gas, I is the ionization potential, ν_e is the electron collision frequency in the gas, equal to the sum of the electron–ion and electron–atom collision frequencies, and ν_I is the ionization frequency calculated according to the classical Thomson formula.[10]

The electron–atom collision frequency was calculated using the well-known temperature dependence of the transport cross section of electron scattering on the inert-gas atoms.[13]

Since the duration of the processes under consideration under the conditions of the problem is much less than the electron–lattice relaxation time in a typical metal $\tau_l \sim 10^{-10}$ s, the lattice temperature and the temperature of the heavy gas component will vary only weakly during the time calculated. For example, even for the mean electron temperature $T_e \sim 10$ eV at the time $t \sim 10^{-13}$ s the lattice temperature will vary by an amount of the order of 10^2 K. For this reason, the lattice and gas temperatures were taken to be unvarying in the calculations. In addition, the radiative thermal conductivity was not taken into account in the calculation of the heat transfer processes, since the plasma arising as a result of rapid ionization is optically transparent (the mean free path of a photon in it $l_{ph} \sim (\omega/\omega_p)^2(\omega/\nu_e)\lambda \gg \lambda$, where ω_p is the electron plasma frequency and λ is the wavelength of the radiation).

The boundary conditions for Eqs. (2) and (3) have the form

$$\frac{\partial T_e}{\partial z} = 0, \quad z = -l_m, l_a,$$

$$\left[\chi \frac{\partial T_e}{\partial z}\right] = 0, \quad z = 0, \quad (4)$$

where l_m and l_a are the boundaries of the region under consideration in the metal and in the gas (formally $l_m, l_a \to \infty$).

The equations of continuity of the electron component and the ion component have the form

$$\frac{\partial n_{e,i}}{\partial t} = \frac{\partial}{\partial z}\left[D_{e,i}\frac{\partial n_{e,i}}{\partial z} \pm \mu_{e,i} E_z n_{e,i}\right] + \nu_I n_e - \beta_p n_e n_i - \beta_{tr} n_e^2 n_i, \quad (5)$$

where β_p and β_{tr} are respectively the coefficients of photorecombination and impact–radiative recombination, and $D_{e,i}$ and $\mu_{e,i}$ are the diffusion coefficient and the mobility.[12]

In Eq. (5) we have left out the divergence term $\mathrm{div} n_{e,i}\mathbf{u}$, since the mechanism of diffusional transport plays the main role in rapid ionization processes.[2] We represent the electron concentration in the form of a sum $n_e = n_{e(\mathrm{ion})} + n_{e(\mathrm{lay})}$, where $n_{e(\mathrm{ion})}$ and $n_{e(\mathrm{lay})}$ satisfy the following equations:

$$\frac{\partial n_{e(\mathrm{ion})}}{\partial t}=\frac{\partial}{\partial z}\left[D_e\frac{\partial n_{e(\mathrm{ion})}}{\partial z}+\mu_e E_z n_{e(\mathrm{ion})}\right]+\nu_I n_e-\beta_p n_e n_i-\beta_{tr} n_e^2 n_i,$$

$$\frac{\partial n_{e(\mathrm{lay})}}{\partial t}=\frac{\partial}{\partial z}\left[D_e\frac{\partial n_{e(\mathrm{lay})}}{\partial z}+\mu_e E_z n_{e(\mathrm{lay})}\right]. \tag{6}$$

By virtue of the fact that the concentration of the EBL electrons $n_{e(\mathrm{lay})}$ near the metal surface is much greater than the concentration of the electrons $n_{e(\mathrm{ion})}$ formed as a result of ionization, the boundary value of n_e at $z=0$ is equal to n_0 according to Eq. (1). This allows us to write the boundary conditions for Eqs. (6) in the following form:

$$\begin{aligned} z=0:&\quad n_{e(\mathrm{lay})}=n_0,\\ &\quad D_e\frac{\partial n_{e(\mathrm{ion})}}{\partial z}+\mu_e E_z n_{e(\mathrm{ion})}=0,\\ z=l_a:&\quad D_e\frac{\partial n_{e(\mathrm{lay})}}{\partial z}+\mu_e E_z n_{e(\mathrm{lay})}=0,\\ &\quad D_e\frac{\partial n_{e(\mathrm{ion})}}{\partial z}+\mu_e E_z n_{e(\mathrm{ion})}=0. \end{aligned} \tag{7}$$

The boundary conditions for the ion component n_i have a form analogous to the conditions for $n_{e(\mathrm{ion})}$. The initial conditions for $n_{e(\mathrm{ion})}$ and n_i are: $n_{e(\mathrm{ion})}|_{t=0}=0$, $n_i|_{t=0}=0$. The initial condition for $n_{e(\mathrm{lay})}$ corresponds to the electron distribution (1) at the initial temperature T_0. Note that the setup time of the distribution (1) near the metal surface is equal in order of magnitude to $D/v_e\simeq\omega_p^{-1}(n_0)\sim 10^{-14}-10^{-15}$ s (v_e is the thermal speed of the electrons), i.e., the boundary-layer concentration $n_0(T)$ is able to ''track'' variations of the temperature occurring over substantially longer times.

The distribution of the longitudinal electric field E_z is found from the equation

$$\frac{\partial E_z}{\partial z}=-\frac{e}{\varepsilon_0}(n_e-n_i),$$

which at each time step obeys the boundary condition

$$E_z|_{z=0}=\frac{e}{\varepsilon_0}\int_0^{l_a} n_{e(\mathrm{lay})}dz$$

(here it has been taken into account that the integral of $n_{e(\mathrm{ion})}-n_i$ is equal to zero).

The spatial distribution of the electric field of the electromagnetic wave is the solution of Maxwell's equations. However, by virtue of the fact that ν_I^{-1}, the characteristic time of variation of the parameters of the problem (concentration, temperature, etc.), is much larger than ω^{-1}, the field distribution can be found by solving the stationary wave equation

$$\frac{d^2E}{dz^2}+k_0^2\varepsilon' E=0,$$

$$\varepsilon'=1-\frac{e^2 n_e}{m\varepsilon_0(\omega^2+\nu_e^2)}-i\frac{e^2 n_e}{m\varepsilon_0(\omega^2+\nu_e^2)}\frac{\nu_e}{\omega}, \tag{8}$$

at each time step, where $k_0=\omega/c$.

The given equation assumes that the development of parametric instabilities in the nascent plasma, accompanied by decay of the electromagnetic waves into plasma and ion–sound waves,[14] will take place much slower than the development of ionization. In order that this condition be fulfilled, it is necessary that the growth rate γ_d of the decay instability of the electromagnetic wave, which is equal in order of magnitude to $\gamma_d\sim(eE/m\omega c)\omega_p$, be much smaller than the ionization frequency ν_I. If the field strength of the wave field $E\leqslant 5\times 10^9$ V/m and the concentration $n_a\geqslant 10^{27}\mathrm{m}^{-3}$ while the temperature $T\sim I$ (the parameter values that were used in the calculations), then $\gamma_d\leqslant 10^{12}\mathrm{s}^{-1}$ and $\nu_I\geqslant 10^{13}\mathrm{s}^{-1}$. Consequently, under the conditions under consideration collisions are the primary mechanism of radiative dissipation, which makes it possible in Eq. (8) to use the ''traditional'' dependence for the complex dielectric constant ε'.

Equation (8) obeys the following boundary conditions:

$$z=0,\ l_a:\quad [E]=0,\quad \left[\frac{dE}{dz}\right]=0, \tag{9}$$

where the wave field has the form

$$z<0:\quad E=\frac{1}{2}\exp[k_0\kappa_i z]\{E_0\exp[i(\omega t+k_0\kappa_r z)]+\mathrm{c.c.}\},$$

$$0<z<l_a:\quad E=\frac{1}{2}\{E_a(z,t)\exp[i\omega t]+\mathrm{c.c.}\},$$

$$z>l_a:\quad E=\frac{1}{2i}E_\infty\{\exp[i(\omega t-k_0 z)]-\exp[-i(\omega t-k_0 z)]\}+\frac{1}{2}\{R\exp[i(\omega t-k_0 z)]+\mathrm{c.c.}\}. \tag{10}$$

Representing the wave amplitude and the dielectric constant in the form $E_0=E_{01}+iE_{02}$, $R=R_1+iR_2$, $E_a=E_{a1}+iE_{a2}$, $\varepsilon'=\varepsilon_1+i\varepsilon_2$, with the help of boundary conditions (9), we obtain the following relations at the boundaries of the region:

$$\begin{aligned} &\left.\frac{dE_{a1}}{dz}\right|_{z=0}-k_0\kappa_i E_{a1}(0)+k_0\kappa_r E_{a2}(0)=0,\\ &\left.\frac{dE_{a2}}{dz}\right|_{z=0}-k_0\kappa_r E_{a1}(0)-k_0\kappa_i E_{a2}(0)=0,\\ &\left.\frac{dE_{a1}}{dz}\right|_{z=l_a}-k_0 E_{a2}(l_a)=2k_0E_\infty,\\ &\left.\frac{dE_{a2}}{dz}\right|_{z=l_a}+k_0 E_{a1}(l_a)=0. \end{aligned} \tag{11}$$

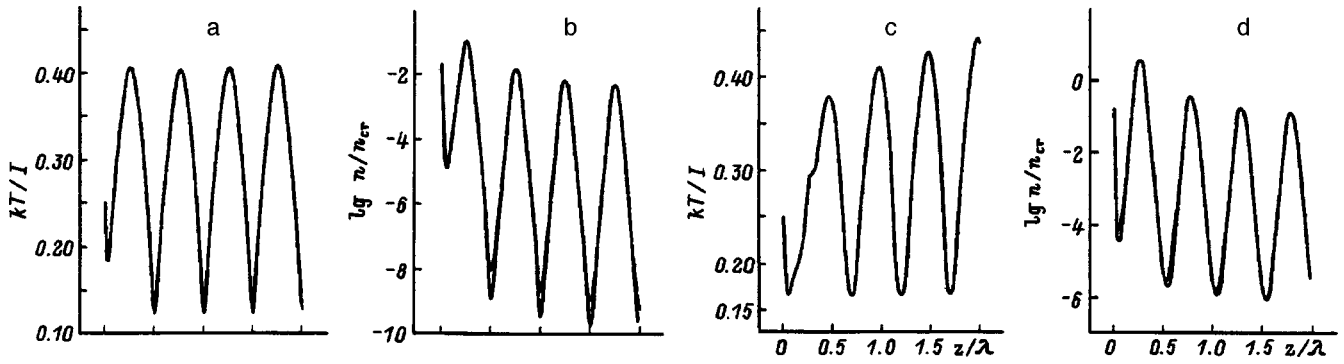

FIG. 1. Distribution of the electron temperature T_e, and also the electron and ion concentrations $n_{e,i}$ at different times $t=7.0\times10^{-13}$ (a,b); 1.0×10^{-12} s (c,d). The gas concentration was $n_a=3\times10^{27}$ m^{-3}; after an incident wave with $E_\infty=2\times10^9$ V/m.

In this case, the wave equation (8) splits into two equations

$$\frac{d^2E_{a1}}{dz^2}+k_0^2(\varepsilon_1E_{a1}-\varepsilon_2E_{a2})=0,$$

$$\frac{d^2E_{a2}}{dz^2}+k_0^2(\varepsilon_1E_{a2}+\varepsilon_2E_{a1})=0. \qquad (12)$$

To solve this system of equations numerically, we used the finite-element method. A difference approximation of the heat conduction and continuity equations was developed with the help of a first-order conservative scheme in τ and h. The diffusion terms were written in implicit form, while the term μEn in the continuity equation was written in explicit form.

RESULTS OF NUMERICAL CALCULATIONS AND DISCUSSION

In our calculations we chose the following numerical values of the parameters of the problem: $\lambda=1.06$ μm; radiative intensity $I_{em}\sim3\times10^{15}-3\times10^{16}$W/m^2; refractive indices of the metal at the given wavelength $\kappa_r=1.5$, $\kappa_i=10.1$ (corresponding to a concentration of the conduction electrons in the metal $n_m=4\times10^{28}$m^{-3}); gas concentration $n_a=(10^{27}-10^{28})$m^{-3} (pressure $p=40-400$ atm); ionization potential $I=15.8$ eV (corresponding to argon); the initial temperature T_0 was varied from 300 to 3000 K, which did not have a noticeable effect on the results; dimensions of the calculation region $l_m=10\kappa_i^{-1}k_0^{-1}$, $l_a=2\lambda$.

Figure 1 displays profiles of the electron temperature T_e for two different times (Figs. 1a and 1c) and profiles of the electron and ion concentrations n_e and n_i (Figs. 1b and 1d) corresponding to a gas concentration $n_a=3\times10^{27}$m^{-3} and incident wave field $E_\infty=2\times10^9$V/m. As can be seen, for the large values of n_a and E_∞ and early times considered, the mechanism of ambipolar diffusion has no qualitative effect on the spatial distribution of the temperature and concentration. Their distribution clearly reproduces the profile of the intensity $|E_a|^2$ of the standing electromagnetic wave, having period $\lambda/2$ at the initial stage of ionization, while the profiles of n_e and n_i are nearly indistinguishable. The electron temperature at the given stage of ionization grows rapidly until its value reaches a few tenths of I, which leads to a significant increase in the rate of the ionization processes. The temperature hardly varies with time beyond this point, and the electron concentration increases rapidly in the regions of the maxima of T_e (this ionization stage is clearly visible in Figs. 1a and 1b). After the value of n_e at the maxima becomes larger than the critical concentration $n_{\rm cr}$, the spatial distribution of the field E_a loses its periodicity while its amplitude begins to decrease with further growth of n_e, which leads to a rapid falloff of the temperature near the surface of the target. The rate of decrease of the temperature at the beginning of this stage lies within the limits 10^4 to 10^5 deg/ps for the example shown in Fig. 1. The electron concentration reaches its critical value $n_{\rm cr}$ starting in the region of the first intensity maximum (nearest to the surface of the target), after which this happens near the other maxima (this trend is easy to make out by comparing Figs. 1b and 1d). Simultaneous with this, the layers of plasma of above-critical density increase in thickness, which then leads to their coalescence and total screening of the irradiated target.

Figure 2 plots some points reflecting the dependence of the characteristic breakdown time τ_I on n_a and E_∞. The

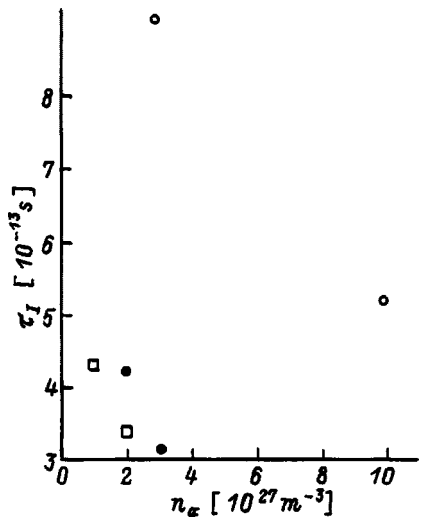

FIG. 2. High-frequency ($\omega^2\gg\nu_e^2$) dependence $\tau_I=\tau_I(E_\infty,n_a)$. E_∞, V/m: ○ — 2×10^9, ● — 3×10^9, □ — 4×10^9.

breakdown time τ_I is defined as the time at which the electron–ion collision frequency ν_{ei} becomes larger than the electron–atom collision frequency ν_{ea}. Such an arbitrary definition of the onset of breakdown makes it possible to record the onset of strong ionization of the gas regardless of its concentration, i.e., to track the degree of ionization of the gas without regard for the absolute electron concentration. Calculating for several combinations of E_∞ and n_a, it is possible to obtain curve families $\tau_I = \tau_I(n_a, E_\infty)$. Using these dependences it is easy to estimate any one of the parameters τ_I, n_a, and E_∞ after first assigning the values of the other two. It is possible to make such an estimate irrespective of the the frequency of the electromagnetic wave if instead of the variable E_∞ we use E_∞/ω (in the high-frequency case).

If the intensity of the incident wave field is less than some threshold value, then screening of the target will not take place. For example, for $n_a = 10^{27}\text{m}^{-3}$ and $E_\infty = 2 \times 10^9\text{V/m}$ the temperature at the maxima reaches the value $0.4I$ after less than a picosecond and its growth then continues at a significantly lower rate [the electron concentration at this time does not exceed $(0.2-0.3)n_{\text{cr}}$]. Therefore, if the electron concentration reaches its critical value n_{cr}, it is after a time comparable with the electron–lattice relaxation time in the metal τ_l.

CONCLUSION

On the basis of our numerical calculations we can draw the following conclusions. When a metal target is irradiated by electromagnetic pulses, as a result of nonequilibrium heating of the electron component a boundary layer of emission electrons is formed above the target surface. Under the condition of high gas density and high radiation intensity, formation of the electron boundary layer leads to the appearance in the gas of regions of highly ionized plasma, screening the target from the incident radiation. The duration of this process for gas density $n_a \sim 10^{27} - 10^{28}\text{m}^{-3}$ and radiative intensity $I_{em} \sim 3\times10^{15} - 3\times10^{16}\text{W/m}^2$ ($E_\infty \sim 3\times10^9\text{V/m}$) lies within the limits of a few tenths of picoseconds to several picoseconds. Our calculations yield a relation between the threshold values of E_∞ and n_a for which the regime of very rapid ($\tau_I \ll \tau_l$) ionization of the gas is realized.

[1] A. A. Uglov, N. N. Rykalin, and M. M. Nizametdinov, Zh. Éksp. Teor. Fiz. **69**, 722 (1975) [Sov. Phys. JETP **42**, 367 (1975)].
[2] V. I. Mazhukin, A. A. Uglov, and B. N. Chetverushkin, Dokl. Akad. Nauk SSSR **246**, 1338 (1979) [Sov. Phys. Dokl. **24**, 443 (1979)].
[3] V. I. Mazhukin, A. A. Uglov, and B. N. Chetverushkin, Fiz. Khim. Obrab. Mater., No. 6, 73 (1979).
[4] V. I. Mazhukin, A. A. Uglov, and B. N. Chetverushkin, Kvantovaya Elektron. **10**, 679 (1983) [Sov. J. Quantum Electron. **13**, 419 (1983)].
[5] A. V. Ivlev, K. B. Pavlov, and M. A. Yakovlev, Zh. Tekh. Fiz. **64**, 50 (1994) [Tech. Phys. **39**, 888 (1994)].
[6] A. V. Ivlev and M. A. Yakovlev, Zh. Tekh. Fiz. **65**, 142 (1995) [Tech. Phys. **40**, 368 (1995)].
[7] A. B. Gupta and K. S. Singwi, Phys. Rev. B **15**, 1801 (1977).
[8] Yu. B. Afanas'ev and A. P. Kanavin, Kvantovaya Elektron. **10**, 2267 (1983) [Sov. J. Quantum Electron. **13**, 1473 (1983)].
[9] K. B. Pavlov and M. A. Yakovlev, Izv. Akad. Nauk SSSR, Ser. Fiz., No. 1, 84 (1989).
[10] Yu. P. Raĭzer, *Physics of Gas Discharges* [in Russian], Nauka, Moscow (1987), 592 pp.
[11] S. I. Anisimov, Ya. S. Imas, G. S. Romanov *et al.*, *Action of High-Power Radiation on Metals* [in Russian], Nauka, Moscow, 1970.
[12] Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* [Academic Press, New York, 1966-1967; Nauka, Moscow, 1966, 688 pp.].
[13] L. J. Kieffer, At. Data **2**, 293 (1971).
[14] V. P. Silin, *Parametric Action of High-Power Radiation on a Plasma* [in Russian], Nauka, Moscow (1973), 296 pp.

Translated by Paul F. Schippnick

System for measuring collective scattering spectra for thermonuclear plasma diagnostics

L. V. Lubyako, E. V. Suvorov, A. B. Burov, A. M. Shtanyuk, Yu. A. Dryagin, L. M. Kukin, and N. K. Skalyga

Institute of Applied Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia
(Submitted May 6, 1997)
Zh. Tekh. Fiz. **68**, 54–62 (August 1998)

The design and arrangement of a detection system for measuring spectra, using a 140 GHz gyrotron as the probe-radiation source, developed for investigating collective Thomson scattering in plasma are discussed. The measurement procedure and examples of the results obtained are presented. [S1063-7842(98)00908-8]

As is well known (see, for example, Refs. 1 and 2), the spectra of electron density fluctuations, whose frequencies ω and wave numbers k are related by the relation

$$\omega \cong k v_i, \tag{1}$$

where v_i are the characteristic ion velocities, carry information about the ion distribution function.

A widely used method of investigating the characteristics of plasma turbulence (specifically, the spectra of electron density fluctuations) is collective scattering of monochromatic probe radiation by the plasma. For collective scattering, the well-known resonance conditions hold:

$$\omega_s = \omega_i + \omega_p, \quad \mathbf{k}_s = \mathbf{k}_i + \mathbf{k}_p, \tag{2}$$

where $\omega_i, \mathbf{k}_i, \omega_p, \mathbf{k}_p$ and $\omega_s, \mathbf{k}_s$ are the frequencies and wave numbers of the probe radiation, the plasma turbulence, and the scattered radiation, respectively.

Thus, the spatio-temporal spectra of the scattered radiation carry information about the plasma turbulence spectra. For the scattered signal spectrum to be determined by the resonance relations (2), the so-called Salpeter parameter must satisfy the relation

$$1/k_p^2 r_d^2 \gg 1, \tag{3}$$

where r_d is the electronic Debye radius.

The relations (2) and (3) determine the admissible collective-scattering geometry for the known frequency of the probe-radiation source, while for the chosen scattering geometry and the expected parameters of the ion component of the plasma the relation (1) determines the required frequency range of spectral analysis. The problems of ion distribution function (specifically, the ion temperature) diagnostics by the collective scattering method have been discussed for many years in the literature.[1–8] The main difficulty here is due to the extremely low value of the scattering coefficient of equilibrium fluctuations, which for typical conditions of a thermonuclear plasma does not exceed 10^{-23} 1/Hz. This makes it necessary to use a powerful source of probe radiation and a quite sensitive detector.

Historically, progress in experimental investigations of the ion distribution function depended on the development of powerful sources of radiation, among which optical and infrared lasers were the leading sources for a long time. However, for optical sources the condition (3) can be satisfied only for very small scattering angles (essentially forward scattering), which makes it much more difficult to measure and interpret the collective scattering spectra. Good results from local measurements of the ion temperature were obtained with a D_2O laser (385 μm, 0.5 J with 1.4 μs pulse duration) as the source of probe radiation.[4] Nonetheless, the limited pulse duration makes the use of this source very problematic for important applications such as α-particles diagnostics of a thermonuclear plasma, where the scattered signal is another several orders of magnitude weaker. In this connection, the millimeter range is quite promising. Quite powerful generators (gyrotrons), used for electron-cyclotron heating in a quasicontinuous operating regime, have been developed for this range. This gives hope that the signal/noise ratio can be substantially increased by averaging over a sufficiently long measurement time.[5]

In the present paper we present a system, developed at the Institute of Applied Physics of the Russian Academy of Sciences, for measuring collective scattering spectra that is intended to be used together with a gyrotron, which provides powerful probe radiation at a frequency of 140 GHz. The system is intended for use in medium-size controlled thermonuclear fusion installations (specifically, we have in mind the W7-AS stellarator — Garching, Germany, and the FTU tokamak — Frascati, Italy) with ion temperature ranging from several hundreds of eV to several keV. The spectral analysis range makes it possible to measure the collective scattering spectra of thermal density fluctuations in hydrogen or deuterium plasma in different geometries (scattering angle interval from 60 to 180°).

The collective-scattering system (Fig. 1) includes a source of powerful probe radiation — a gyrotron, a transmission line for the powerful probe radiation, a receiving-transmitting antenna system which determines the scattering geometry and makes it possible, in principle, to change the position of the scattering volume inside the plasma filament, a transmission line for the scattered radiation, a detection system for measuring spectra, and a data acquisition, storage, and processing system.

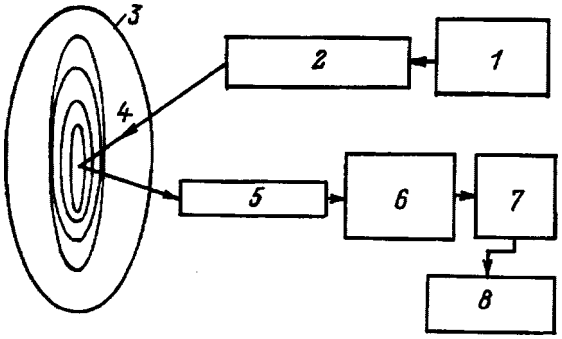

FIG. 1. Block diagram of the system for measuring collective scattering: *1* — Gyrotron, *2* — high-power line, *3* — vacuum chamber, *4* — plasma, *5* — low-power waveguide line, *6* — HF part of the detector, *7* — IF spectrum analyzers, *8* — digital processing system.

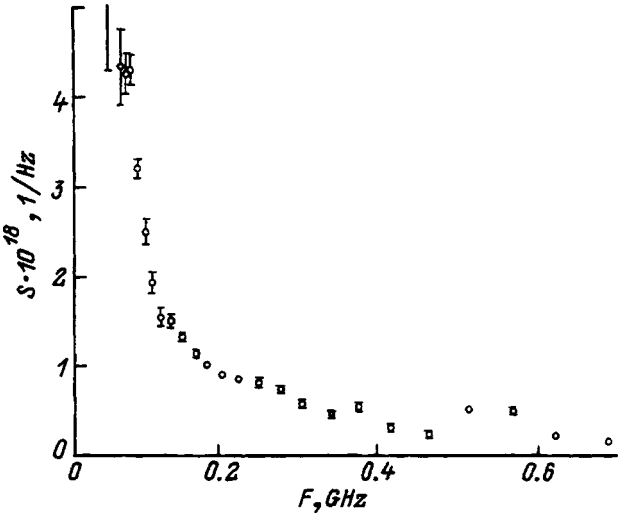

FIG. 2. Relative power density of the gyrotron noise.

REQUIREMENTS FOR THE DETECTION SYSTEM

The main requirements for the detection system are dictated chiefly by the expected characteristics of the scattered signal. Referring the reader to the literature for the details,[1,2] we note that in the case of scattering by thermal fluctuations the signal investigated consists of noise whose spectrum lies in a limited frequency band and is symmetric with respect to the frequency of the probe radiation. The shape of the spectrum depends on the scattering geometry and the gas composition, the density, and the electron and ion temperatures of the plasma. The width of the spectrum is maximum for hydrogen with close-to-backward scattering. At temperatures of the order of several keV the upper limiting frequency for hydrogen plasma is ≈ 1 GHz, which gives the required analytical band of the detection system. The magnitude of the scattered signal intercepted by the receiving antenna varies from several tens to hundreds of eV, depending on the scattering geometry and the plasma parameters.[8–10] This does not require detection apparatus with ultrahigh sensitivity and makes it possible, by using a superheterodyne detection scheme, to provide a wide dynamical range for the values of the measured signals with adequate spectral resolution.

It should be kept in mind, however, that in reality external factors, such as parasitic penetration of some portion of the powerful probe radiation to the input of the receiving antenna[9] as well as the characteristic cyclotron radiation of the plasma, can limit the sensitivity of the receiving system. The thermal cyclotron radiation of the plasma is, as a rule, stronger than the expected scattered signal, but its contribution can be decreased to several eV, if the magnetic field in the installation does not correspond to cyclotron resonance at the frequency of the probe radiation. Moreover, since it is difficult to expect the cyclotron noise spectrum to have any features in the quite narrow analytical band, even a weak scattered signal can be distinguished against its background by means of accumulation and averaging over a sufficiently long period of time.

The most serious obstacle for reliable measurement of collective-scattering spectra could be ''parasitic getting through'' of the probe radiation to the input of the receiving antenna, both as a result of multiple reflections from the chamber walls and on account of the side lobes of the transmitting and receiving antennas. Sufficiently strong parasitic radiation can cause overloading of the detection system or even damage to the mixer at the input of the detection system. To prevent this the receiver must be shielded by a system of attenuators, which together with the parasitic signal also weaken the scattered signal, degrading its ratio to the detector noise, or by notch filters, which decrease the degree to which radiation at the frequency of the probe radiation ''gets through.''

The ratio between the characteristic gyrotron noise and the useful signal is determined by the strength of this noise in the useful signal band and by the degree of decoupling between the transmitting and receiving antennas. The plasma cyclotron noise can be taken into account simply by subtracting out a constant noise component, whose absolute magnitude need not be well known, whereas subtracting out the gyrotron noise is a more complicated procedure. It is difficult to expect the gyrotron noise spectrum to be constant within the useful signal band (Fig. 2). For this reason, in order to ''subtract it out'' very careful measurements of the gyrotron noise spectrum and quite accurate measurements of the parasitic signal strength at the input to the detection channel in the collective-scattering regime must be performed. Experimental investigations[7,11–13] have shown that the parasitic radiation level depends strongly on the scattering geometry and the plasma parameters. As one can see from the Fig. 2, for decoupling of the order of 40 dB between the transmitting and receiving antennas and for probe radiation power of about 0.5 MW the noise level at frequencies displaced by 100–150 MHz from the center of the generation line can be of the order of 100 eV in terms of the noise equivalent temperature.

BLOCK DIAGRAM OF THE SPECTRUM MEASURING SYSTEM

The measuring system (Fig. 3) is assembled in the manner of a superheterodyne multichannel radiometer with input signal modulation. The following basic parts can be distinguished according to functional characteristics: a wide-band

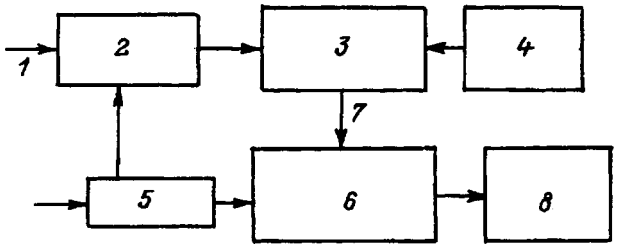

FIG. 3. Block diagram of the measurement of the spectra: *1* — Waveguide input, *2* — detector shielding system, *3* — heterodyne frequency converter, *4* — system for stabilizing the heterodyne frequency, *5* — synchronizer, *6* — IF spectrum analyzers, *7* — low-loss coaxial line, *8* — digital processing system.

heterodyne converter equipped with a system for shielding the mixer from parasitic gyrotron radiation; a system for stabilizing the heterodyne frequency; intermediate-frequency (IF) spectrum analyzers; a multichannel analog-to-digital converter; and, a computer-based data acquisition and processing system. In addition, a square-law detector for measuring the parasitic radiation power entering the input channel is connected, through a waveguide coupler, at the input of the detection system.

The heterodyne converter enables transferring the spectrum of the signal being investigated from the high-frequency (HF) region (of the order of 140 GHz) into the IF region 50–1200 MHz, where spectral analysis is performed. The spectral analysis problem is solved with the aid of two multichannel systems performing parallel analysis, enabling operation in real-time with of the order of 1 ms resolution. The IF scanning spectrum analyzer includes a main filter system with 10% relative spectral resolution. The second analyzer for the IF spectrum is a ''lens'' system, which makes it possible to investigate with a higher spectral resolution any section of the spectrum with a width of 100 MHz within the entire scan band.

A characteristic feature of the frequency converter is the ''quasihomodyne'' operating regime, which is made possible by two-band operation of the mixer (without suppression of the mirror-image band) and by the equality of the heterodyne and gyrotron frequencies. The use of both bands is optimal from the standpoint of sensitivity, since the scattered-signal spectrum is symmetric relative to the probe signal frequency. To preserve a single-valued relation between the output signal spectrum and the IF signal the relative detuning of the heterodyne, detector, and gyrotron frequencies must not exceed the minimum spectral resolution. The system for stabilizing the heterodyne frequency provides a relative frequency stability of 10^{-8}, which is obviously less than the spectral resolution of the device, so that detuning can appear because of either regular drift of the gyrotron frequency during the pulse or inadequate pulse-to-pulse repeatability of the gyrotron operating regime.

Investigations of several gyrotrons (A,C,E), which in principle can be used for collective scattering on W7-AS, showed (Fig. 4) that each gyrotron investigated is characterized by its own characteristic time dependence of the frequency and this dependence repeats from pulse to pulse with an accuracy of not worse than 10 MHz in identical gyrotron

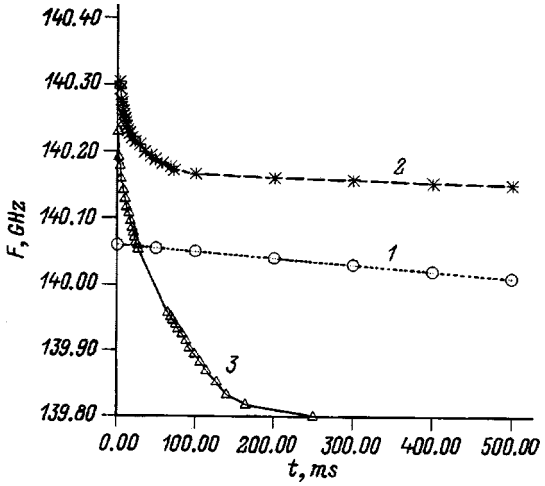

FIG. 4. Variation of the gyrotron frequency inside the generation pulse. *1* — gyrotron *A*, *2* — gyrotron *E*, *3* — gyrotron *C*.

operating regimes. This makes it possible to use for collective scattering a prechosen interval for the generation pulse during which the frequency drift does not exceed several MHz. From the standpoint of measurements the most favorable sections of the radiation pulse appear to be the sections with the minimum curvature on the frequency–time characteristic, i.e., several hundreds of milliseconds after generation commences. However, in the first place, the pulse-to-pulse repeatability degrades somewhat on these sections and, in the second place, the duration of the powerful pulse injected into the chamber is limited to several tens of milliseconds in the case that the radiation is not absorbed by the plasma, and such operating regimes are of interest from the standpoint of decreasing the level of the cyclotron radiation background. On this basis the gyrotron C is of no interest for collective scattering, while gyrotron A is most promising.

The choice of the superheterodyne radiometer scheme makes it possible to satisfy the sensitivity requirement even for the average parameters of the mixer and IF amplifier, whose conversion losses and noise temperature determine the sensitivity of the detector. Specifically, the components used made it possible to obtain at 140 GHz a noise temperature ≈ 0.6 eV referred to the input (with ≈ 0.2 eV at the mixer input).

As is well known, the radiometric sensitivity determined as the minimum detectable increment to the antenna temperature is determined by the relation

$$\Delta T_{\min} = \frac{2T_n}{\sqrt{\Delta f \cdot \tau}}, \tag{4}$$

where Δf is the width of the spectral channel and τ is the averaging time.

In an experiment the total noise temperature of the measuring system, referred to the antenna input, consists of the characteristic radiation of a plasma with brightness temperature T_{pl}, the gyrotron noise, characterized by the temperature $T_{\text{g.n}}$, in the ''wings'' of the generation line (Fig. 2) and the characteristic noises of the receiving system with temperature $T_{\text{r.n}}$

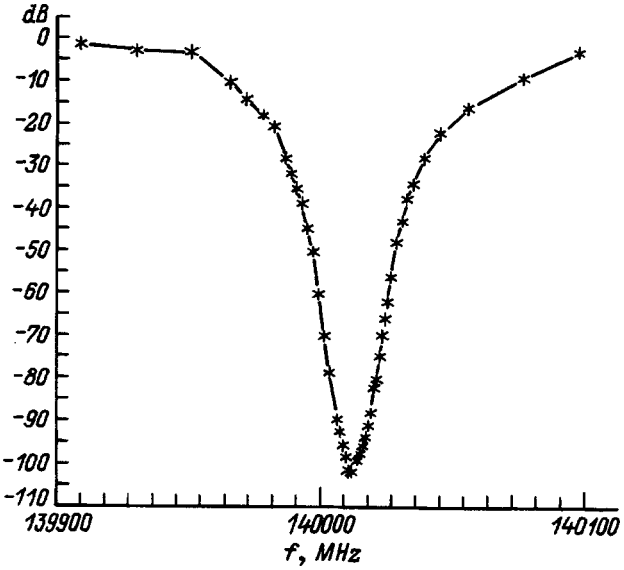

FIG. 5. Frequency characteristic of the notch filter based on a waveguide section.

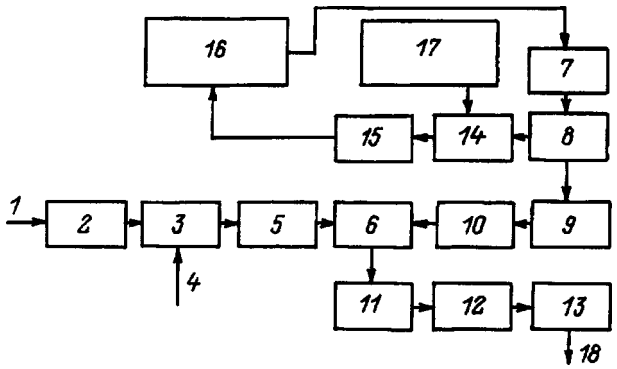

FIG. 6. Block diagram of the HF part of the receiver: *1* — Input waveguide, *2* — notch filter, *3* — modulator, *4* — 10 kHz gated voltage, *5* — rectifier, *6* — mixer, *7* — heterodyne, *8* — coupler, *9* — rectifier, *10* — band-pass filter, *11* — IF amplifier, *12* — HF filter, *13* — second IFA module, *14* — PLL mixer, *15* — IFA PLL, *16* — PLL unit, *17* — synthesizer, *18* — IF output.

$$T_n = T_{pl} + L_a T_{\text{g.n}} + L_l^{-1} T_{\text{r.n}}, \tag{5}$$

where L_a is the coupling coefficient between the receiving and transmitting antennas and L_l is the transfer factor of the channel from the receiving antenna to the mixer input.

To give an idea of the order of magnitudes, we note that a sensitivity of 1 eV with an averaging time of 1 ms corresponds to cyclotron noise of the order of 50 eV in narrow channels ($\Delta f \approx 5$ MHz) and of the order of 200 eV in wide channels ($\Delta f \approx 100$ MHz) or intrinsic detector noise with total damping in the detection channel of the order of 20 dB for narrow channels or 27 dB for wide channels. Finally, the intrinsic gyrotron noise gives a similar limit on the sensitivity in channels displaced approximately by 100 MHz from the center of the generation line with decoupling of the order of 40 dB between the receiving and transmitting antennas.

SHIELDING SYSTEM

A key element of collective-scattering experiments performed with a gyrotron as a source is the problem of shielding the receiver from parasitic radiation. The level of this radiation, even taking account of antenna and transmission line losses, is sufficient to disrupt the normal operation of the detector. To solve this problem, a shielding system was developed which employs a multiresonator notch filter and switched electronic *p–i–n* attenuators. This system solves two problems: attenuation of the parasitic gyrotron radiation, and gating of the detected signal to prevent overloading of the detection system at the moments when the gyrotron generation line falls outside the suppression band of the notch filter.

The system employs a mechanically tunable filter based on a dominant-mode waveguide with 20 cylindrical resonators operating on the H_{013} mode.[14] This filter (Fig. 5) gives maximum suppression of up to 100 dB and not less than 40 dB in a band of the order of 50 MHz. The characteristic outside the ± 50 MHz band relative to the center of the suppression line can be treated as an additional perturbation of the ''through'' transfer function of the entire system as a whole; thus, 3 dB damping corresponds to frequencies of approximately ± 60 MHz, while distortions introduced outside a ± 90 MHz band do not exceed 2 dB. Damping on the wings of the line of the notch filter leads to a small additional loss of sensitivity in the low-frequency (LF) channels. To eliminate the influence of the temperature instability, the filter is placed in a thermostatically controlled unit.

The second component of the shielding system, making it possible to block the input of the detection system at a level of not less than 40 dB and open it only for the measurement time, is a *p–i–n* attenuator that can be controlled by a specially formed gate-voltage pulse. Inside the gate this attenuator operates as a modulator. The controlling voltage is formed with the aid of a synchronizer, which permits varying the triggering time and gate duration over wide limits.

HIGH-FREQUENCY DETECTION SYSTEM

The HF part of the detection system for measuring spectra, which is assembled as a separate unit, well-screened from external radiation, includes the following (Fig. 6): a waveguide *p–i–n* modulator, a ferrite rectifier, a mixer, two IF amplifiers between which are inserted an HF filter with a limiting frequency of 45 MHz, a heterodyne (backward wave tube) and a ferrite rectifier and band-pass filter with an approximately 300 MHz band placed in the heterodyne channel. Moreover, the HF part of a phase-locked loop is also placed there. A notch filter, which is a component of the circuit shielding the detector from the parasitic gyrotron signal, and a coupler for measuring the parasitic gyrotron radiation power are placed at the mixer input in front of the modulator.

In the HF unit, the signal entering from the transmission line for the scattered radiation is amplitude-modulated with a frequency of 10 kHz and enters a balanced mixer through the rectifier 1. The losses in the modulator do not exceed 1 dB, and the direct losses in the rectifier equal 3 dB with 27 dB decoupling. The signal from the heterodyne — a backward wave tube (140 GHz, 10 mW) — is also fed into the mixer

through a rectifier 2 and a band-pass filter. The filter gives additional suppression of the heterodyne noise of at least 5–8 dB in the 300 MHz band centered to the heterodyne frequency. The dc operating regime of the mixer is maintained with a current source (not shown in the diagram) and is monitored according to the measured voltage on its diodes. The use of nonreciprocal devices in the input circuits, though this involves a degradation of sensitivity, is justified by the fact that it makes it possible to solve the parasitic signal problem inherent to modulation radiometers.

The IF signal from the mixer output is amplified by a wide-band low-noise amplifier based on bipolar transistors. The band of the IF amplifier is from 45 MHz to 1.3 GHz, the integral noise temperature equals 210 K, and the total (for two modules) gain equals 56 dB. To improve the dynamic characteristic of the IF amplifier at low frequencies, an HF filter with a cutoff frequency of 45 MHz and a rate of fall off of 10 dB/MHz is inserted between the IF amplifier modules. The noise temperature of the detector as a whole, taking account of losses in the notch filter, the modulator, and the rectifier, i.e., referred to the input of the unit, was equal to about 5000 K.

The IF signal flows from the output of the HF unit along a coaxial cable with low losses into the spectrum analyzers.

The frequency of the heterodyne is stabilized with phase-locked loop (PLL). The PLL system includes the following components (Fig. 6): a waveguide coupler, a harmonic mixer, IF-amplifier PLL, RCH6-02 frequency synthesizer for the range 4–8 GHz, phase detector, and a scheme for controlling the frequency of the backward wave tube. The HF components of the PLL are placed directly in the microwave unit of the spectrum analyzer. The stabilization system includes only one PLL unit, which was made possible by the use of a high harmonic number ($n=21$) and a high frequency of the reference oscillator (6678 MHz). The heterodyne frequency is stabilized in the entire generation band of the backward wave tube (of the order of 20%), and in the synchronization mode it is set by the frequency of the reference oscillator — frequency synthesizer

$$f_{\text{het}} = 2l \cdot f_{\text{syn}} - 250 \text{ MHz}. \tag{6}$$

The stability of the heterodyne frequency is determined by the synthesizer parameters and in our case is not less than 5×10^{-8}; this stability is much more than enough for the spectral analysis system described above. The hold-in band of the PLL system equals 600 MHz and the lock-in band equals 250 MHz. The minimum frequency tuning step is 210 kHz.

INTERMEDIATE-FREQUENCY SPECTRUM ANALYZERS

The IF scanning spectrum analyzer and an additional high-resolution spectrum analyzer were assembled as individual units. Both devices implement parallel analysis and differ only by the parameters of the filter systems and the presence of a second heterodyne in the auxiliary spectrum analyzer.

The scanning spectrum analyzer (Fig. 7), intended for analyzing a signal in a band from 50 MHz to 1.2 GHz with a

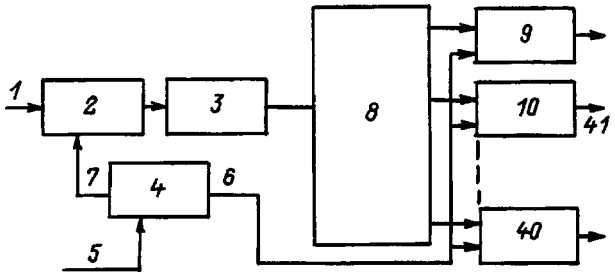

FIG. 7. Block diagram of scanning IF spectrum analyzer: *1* — Input, *2* — *p–i–n* attenuator, *3* — IFA, *4* — synchronizer, *5* — external gating signal, *6* — 10 kHz reference voltage output, *7* — gate, *8* — 32–channel multiplexer, *9–40* — LF channels, *41* — output to ADC.

constant relative resolution of 10%, employs a specially developed 32-channel multiplexer, consisting of a chain of active wide-band intermediate-frequency amplifiers (IFAs), eight cascades of active power dividers based on field-effect transistors, 32 intermediate-frequency band-pass filters, and 32 amplitude detectors. In addition, it has a wide-band output, decoupled with the help of an attenuator, for connecting an auxiliary spectrum analyzer.

The IF filters are assembled in a three-loop arrangement with capacitive coupling based on lumped elements and tuned to the following frequencies: 48, 54, 59, 66, 73, 80, 90, 102, 111, 123, 136, 153, 172, 187, 205, 227, 252, 280, 308, 344, 379, 418, 464, 514, 570, 625, 692, 766, 845, 940, 1033, 1155 MhZ; the band of each filter equals approximately 10% of the average frequency of the filter. The decoupling between neighboring channels is not less than 20 dB.

The IF analyzer channel also includes a voltage-controlled intermediate-frequency *p–i–n* attenuator and a wide-band IFA, which are placed in front of the multiplexer. The transfer factor required to provide the dynamic range of the IFA is set by an attenuator. In addition, this attenuator is used for additional gating synchronously with the HF modulator.

Each LF channel of the analyzer includes a low-noise LF amplifier, a synchronous filter, an attenuator, a synchronous detector, an integrating RC circuit, and an output video amplifier. The time constant is chosen to be 1 ms based on the impossibility of using a modulation frequency above 10 kHz because of the limited speed of *p–i–n* modulator.

The use of a system of attenuators makes it possible to obtain a wide dynamical range for the measured signals appearing at the input of the detector system: from hundredths of an eV, which is the threshold sensitivity because of the intrinsic noise of the detector system, up to 10^8 eV. This made it possible to use the measuring system not only for the direct purpose for which it was intended — measurements of the ion temperature, but also for investigating nonequilibrium plasma turbulence as well as for real-time measurements of the frequency drift of the gyrotron radiation.

The second spectrum analyzer (Fig. 8) is intended for investigating the fine structure of the spectrum and makes it possible to observe with a resolution of 5.5 MHz the spectrum in a 100 MHz band, centered to one of the frequencies of the analytical band. It is implemented in an arrangement with second heterodyning and includes the following: an in-

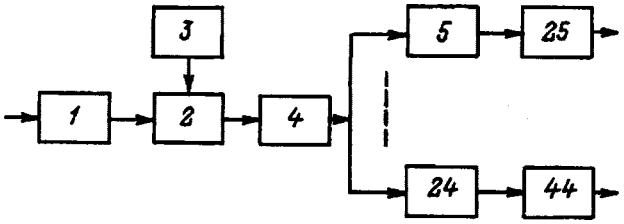

FIG. 8. Block diagram of a ''lens'' spectrum analyzer: *1* — HF filter, *2* — mixer, *3* — second heterodyne, *4* — band-pass IFA; *5–24* — system of 20 narrow-band filters, *25–44* — LF channels.

put preselector, a mixer, a second heterodyne, a converted-signal selector, IFA2, a 20-channel multiplexer, and a system of 20 selective channels. The second intermediate frequency equals 250 MHz. The analytical band is tuned by switching the heterodyne frequency with a step of 100 MHz together with corresponding switching of the preselector filters. The LF part of the unit is similar to that used in the scanning spectrum analyzer.

MEASUREMENT PROCEDURE

The voltage at the output of each spectral channel of the analyzer can be written as

$$u_i = K_i T_{ai} + u_i^0, \tag{7}$$

where K_i is the total transfer factor of the i-th receiving channel from the input of the receiving antenna to the analog output of the spectrum analyzer; T_{ai} is the antenna temperature, averaged over the band of the i-th receiving channel; and, u_i^0 is a constant shift of the ''instrumental zero.''

Generally speaking, in each channel K_i and u_i^0 depend on the choice of the dynamical range, determined by the settings of the attenuators. For this reason, to measure the signal at the input of the receiving antenna it is necessary to perform calibration measurements, which make it possible to determine these quantities in each channel for each choice of the dynamical range. The ideal method of calibrating the entire detection system as a whole would be to place two noise sources with known and quite strongly differing temperatures in front of the input of the receiving antenna. With such an absolute calibration the quantities K_i and u_i^0 are determined uniquely in each channel, which makes it possible to perform an absolute measurement of the effective temperature of the radiation in each channel in the collective-scattering regime. A good calibration source for a horizontal orientation of the directional pattern of the receiving antenna is an optically thick, resonant plasma layer emitting at the second cyclotron harmonic, such that the radiation temperature of the layer is equal to its physical electronic temperature. An empty vacuum chamber, whose radiation temperature can be assumed to be equal to the temperature of the surrounding medium, can be used as the second reference source. If the effective temperature of the cyclotron radiation and of the measured signal are much greater than this quantity, then its exact value is unimportant and the effective temperature of the measured signal can be determined as follows:

$$T_{ai} = \frac{u_i - u_i^{(k)}}{u_i^{(\mathrm{ECE})} - u_i^{(k)}} T_{\mathrm{ECE}}, \tag{8}$$

where u_i, $u_i^{(k)}$, and $u_i^{(\mathrm{ECE})}$ correspond to the values of the output signal in the i-th channel for collective scattering, for measurements of the signal from an empty chamber, and for calibration according to the cyclotron radiation, respectively.

For a different orientation of the directional pattern of the receiving antenna the cyclotron radiation can be used only for relative calibration (programmed equalization of the transfer factors in all channels of the spectrum analyzer) on the assumption that the effective temperature of the radiation is constant within the analytical band. The absolute value of the measured signals is obtained here with a low accuracy, which is determined either by estimating the cyclotron radiation level in the given geometry or by estimating the losses in the receiving channel and absolute calibration with a ''black body'' source connected to the receiver input.

CONCLUSIONS

The first experiments, which were performed on the W7-AS stellarator, on collective scattering using a 140 GHz gyrotron in the detection system described above for measuring spectra were of a preliminary character and made it possible to understand the need to update a number of components and fundamental solutions. Specifically, it became clear that deeper suppression of the parasitic gyrotron signal is needed. It was necessary to reject open quasioptical components in the line transporting the scattered signal and to provide careful screening of the waveguide channel and the HF part of the detector and to introduce a more effective system for gating the signal being investigated. To improve the stability of the parameters of the detector the radiometer circuit was changed (to modulation instead of compensation) and thermostatting of the HF part of the radiometer was introduced. For purely illustrative purposes, we present here a number of examples.

The first results on collective scattering were obtained on W7-AS in the backscattering geometry using the gyrotron A; in addition, the probe and scattered beams were close to the equatorial plane. In this geometry, in the absence of absorption of the probe radiation by the stellarator plasma, the level of the scattered gyrotron radiation was sufficiently high so that scattering by thermal fluctuations could be detected only for certain optimal positions of the receiving and transmitting antennas by subtracting the gyrotron noise from the general signal. An example of such a difference signal and gyrotron noise spectrum is shown in Fig. 9.

When intense absorption of the probe radiation occurred (the extraordinary wave at the second cyclotron harmonic was used as the probe radiation), the scattered signal could not be distinguished against the background consisting of the plasma cyclotron radiation. Another example of the spectrum of the scattered radiation is shown in Fig. 10. This spectrum was obtained with scattering by nonthermal lower-hybrid turbulence in the same back-scattering geometry, and the level of the scattered signal exceeded by several orders of magnitude the cyclotron radiation of the plasma, the gyrotron

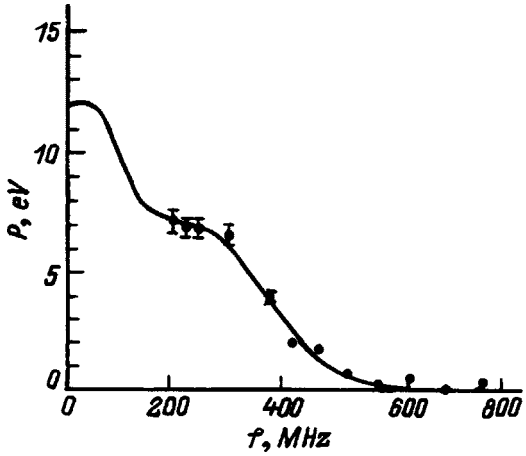

FIG. 9. Experimental (dots) and fitted (solid line) scattering spectra of thermal fluctuations of the density.

noise, and the level of the signal scattered by thermal fluctuations. Good frequency resolution was obtained by using a ''lens'' system of spectral analysis. Figure 11 shows the spectrum obtained for scattering by thermal fluctuations, using gyrotron E as the source of the probe radiation. This spectrum can be used for local determination of the ion temperature, since it is obtained in a 90°-scattering geometry from a volume with a size of the order of 4 cm. A characteristic feature of the new geometry is good decoupling between the transmitting and receiving antennas — the level of the gyrotron noise at the input of the receiving antenna is approximately two orders of magnitude lower than the level of the detected signal. The spectrum presented is distinguished against the cyclotron radiation background, whose effective radiation temperature is two to three times higher than that of the scattered signal.

The system for measuring collective-scattering spectra with the introduction of strong additional attenuation of the signal makes it possible to investigate also the frequency drift during the gyrotron pulse. As an example, Fig. 12 shows the evolution of the generation frequency of gyrotron E obtained in this manner.

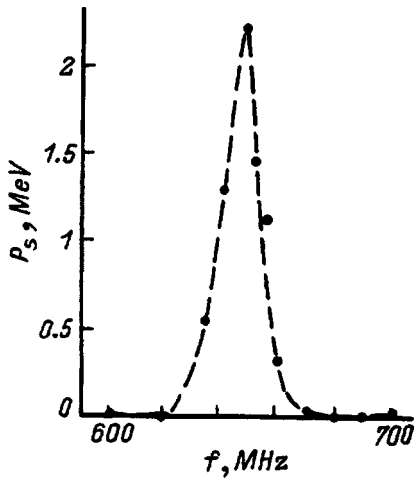

FIG. 10. Scattering spectrum of lower-hybrid turbulence.

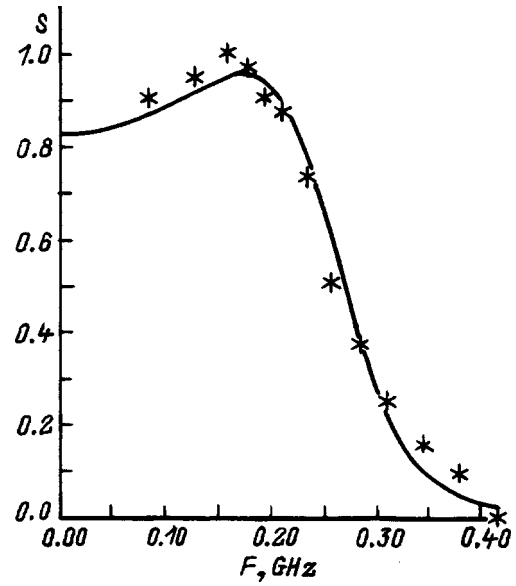

FIG. 11. Experimental (asterisks) and computed (line) scattering spectra of thermal fluctuations.

These investigations have demonstrated that it is possible in principle to use a gyrotron as a source of powerful radiation for collective scattering. This is very important not only and not so much for measurements of ion temperature, but also for α-particle diagnostics under the conditions of a thermonuclear reaction. Today the gyrotron is the only real source of powerful long-duration probe radiation, whose use opens up the prospect of local measurements of the collective scattering spectra of α particles with quite high sensitivity due to the high ''radiometric gain.'' The principles, which were checked on the W7-AS setup, of the system for measuring collective scattering spectra can be transferred without any serious modifications to the α-particle diagnostics system. At the same probe-radiation frequency the ''collectiveness'' condition for scattering is satisfied for any imaginable geometry. For reasonable widths of the spectrum of the measured signal (3–4 GHz) the scattering angle should not be much greater than 30°. An appreciable decrease of the parasitic signal level can be expected in a large reactor-scale setup, but since the scattered signal is expected to be weak, it appears that a notch filter will remain a necessary component of the detection system. Since the duration

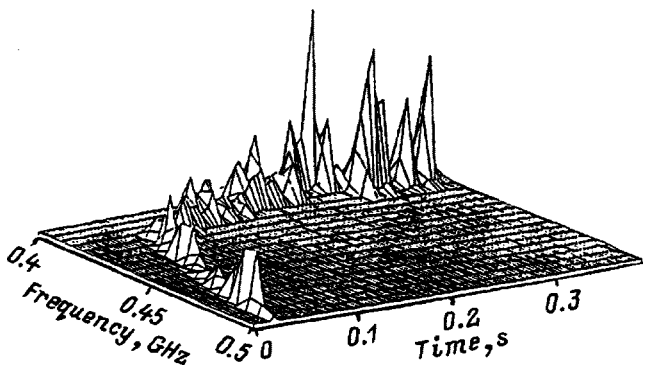

FIG. 12. Evolution of the radiation spectrum of gyrotron E.

of the probe-signal pulse should be maximum (of the order of seconds) in order to increase sensitivity, the problem of developing either a quite wide-band notch filter or a filter with an electronically controlled characteristic to compensate the drift of the gyrotron frequency in a long pulse will arise.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-19455-a).

[1] J. Sheffield, *Plasma Scattering of Electromagnetic Radiation* [Academic Press, New York, 1975; Atomizdat, Moscow, 1978].
[2] M. L. Levin and S. M. Rytov, *Theory of Equilibrium Fluctuations in Electrodynamics* [in Russian], Nauka, Moscow, 1967.
[3] J. A. Hoekzema, A. E. Costley *et al.*, in *Proceedings of the 15th Symposium on Fusion Technology*, Utrecht, 1988, p. 314.
[4] F. Orsitto and P. Buratti, ''Analysis of collective Thomson for alpha particle diagnostics on tokamaks,'' ENEA rep. 1988.
[5] R. Benhn, D. Dicken *et al.*, J. Appl. Phys. **69**, No. 4 (1991).
[6] P. Woskoboinikov, D. R. Cohn *et al.*, Digest XII Intern. Conf. on IR&MM Waves, Orlando, Florida, N CH2490-1 (1987).
[7] L. V. Lubyako, A. G. Luchinin, G. S. Nusinovich *et al.*, Fiz. Plazmy **18**, 211 (1992) [Sov. J. Plasma Phys. **12**, 112 (1992)].
[8] Yu. Ya. Brodskiĭ, L. V. Lubyako, S. I. Nechuev *et al.*, in *Abstracts the Interdisciplinary Cpnference on Instruments, Technology, and Propagation of millimeter and Submillimeter Waves* [in Russian], Kharkov, 1992, p. 109.
[9] E. V. Suvorov, in *Proceedings of the International School of Plasma Physics ''Piero Caldirola,''* Villa Monastero-Varenna (Italy), 1991, pp. 79–98.
[10] E. V. Suvorov, A. B. Burov, Yu. A. Dryagin *et al.*, in *Proceedings of the International Workshop ''Strong Microwaves in Plasma,''* Nizhniĭ Novgorod, 1993, pp. 172–184.
[11] A. N. Kuftin, L. V. Lubyako, and N. K. Skalyga, in *Abstracts the Interdisciplinary Cpnference on Instruments, Technology, and Propagation of millimeter and Submillimeter Waves* [in Russian], Kharkov, 1992, p. 35.
[12] E. V. Suvorov, Yu. A. Dryagin, S. E. Filchenkov *et al.*, in *Ninth Joint Workshop on ECE and ECRH*, World Scientific Publishing Co. Pte LTD, Singapore, 1995, pp. 501–508.
[13] E. V. Suvorov, V. Erckmann, and E. Holzhauer, Plasma Phys. Controlled Fusion Fusion **37**, 1207 (1995).
[14] Yu. Dryagin, N. Skalyga, and T. Geist, Int. J. IR & MM Waves **17**, 1199 (1996).

Translated by M. E. Alferieff

Conductance of shock-compressed metal–insulator–metal structures based on polymethyl methacrylate

N. D. Semkin and K. E. Voronov

Samara State Aerospace University, Samara, Russia
(Submitted February 24, 1997)
Zh. Tekh. Fiz. **68**, 63–66 (August 1998)

The problem of the conductance of shock-compressed metal–insulator–metal (MIM) film structures based on polymethyl methacrylate is solved by numerical and analytical methods. Analytical expressions are obtained for the conductance of a MIM structure impacted by solid particles moving with velocities exceeding 1 km/s and for the output voltage of a capacitor detector. The numerical and analytical calculations of the conductance (current pulse) of a shock-compressed structure are in qualitative agreement with one another. © *1998 American Institute of Physics.* [S1063-7842(98)01008-3]

The study of the processes occurring when high-velocity particles interact with metals,[1,2] semiconductors, and insulators,[3,4] is of great scientific and practical interest in connection with the fact that such materials are used as transducers of the physicochemical parameters of the particles[5,6] and as elements for monitoring the serviceability of the structural components of space vehicles.[7] The sharp increase in the conductance of shock-compressed MIM (metal–insulator–metal) structures is due to both compression of the material and thermal ionization.[8] Particle transducers based on polymethyl methacrylate possess a high impact sensitivity, reliability, and stability of characteristics under the conditions in space. To determine the conductance of shock-compressed insulators in MIM structures as a function of time and the particle parameters, a system of equations consisting of the equations of hydrodynamics and electrodynamics is solved using the equations of state for a metal (aluminum) and an insulator (polymethyl methacrylate)

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{W}) = 0, \tag{1}$$

$$\frac{\partial(\rho \mathbf{W}_i)}{\partial \rho} + \operatorname{div}(\rho W_i \mathbf{W}) + V\rho = 0, \tag{2}$$

$$\frac{\partial(\rho E)}{\partial E} + \operatorname{div}(\rho E \mathbf{W}) + \operatorname{div}(\rho \mathbf{W}) = 0, \tag{3}$$

$$p = p(\rho; E - W^2/2), \quad i = 1,2,\ldots, \tag{4}$$

$$p = a\rho^3; \quad T = \frac{p - a(\rho^3 - \rho_0^3)}{4\rho N}, \tag{5}$$

which are, respectively, the laws of conservation of mass, momentum, and energy and the equations of state of aluminum and the insulator.

The temperature T is determined from a three-term equation of state,[8] and the quantities p, E, $\mathbf{W}$, ρ, and T are found by solving the equations by the method of large particles.[9] Here we are considering the case of a collision of a particle with a MIM structure, deposited on a semi-infinite barrier,[7] consisting of the top plate (aluminum) of a capacitor, i.e., a crater is formed without mechanical damage to the insulator.[4,7] The equation of state for the aluminum plate is used in the form[10]

$$p = p_c(v) + \frac{\gamma(v,\varepsilon)}{v}[\varepsilon - \varepsilon_c(\tau)], \tag{6}$$

where

$$p_c(v) = \frac{C_{mn}}{v}(\sigma_c^m - \sigma_c^n), \quad \sigma_c = \frac{v_{00}}{v}, \quad \sigma = \frac{v_0}{v},$$

$$v = \frac{1}{\rho}; \quad \varepsilon_c(v) = c_{mn}\left(\frac{\sigma_c^m}{m} - \frac{\sigma_c^n}{n}\right) + \varepsilon_s,$$

$$\gamma(v,\varepsilon) = \frac{2/3 + (\gamma_0 - 2/3)(1 + \sigma_m^{2\alpha})\sigma^{\alpha}/(\sigma^{2\alpha} + \sigma_m^{2\alpha})}{1 + (\varepsilon + \varepsilon_c(v))/\varepsilon_{\alpha}},$$

$v_0 = 0.369\ \mathrm{cm^3/g}$, $v_{00} = 0.361\ \mathrm{cm^3/g}$, $\gamma_0 = 2.19$, $\sigma_m = 0.95$; $\alpha = 8$, $m = 1.8$, $n = 0.6$, $c_{mn} = 23.399\ \mathrm{J/g}$; $\varepsilon_{\alpha} = 3.06\ \mathrm{kJ/g}$, $\varepsilon_s = 12.2\ \mathrm{kJ/g}$, $[p] = 10^4$ bar; $c_p = 1.425 \times 10^3\ \mathrm{J/(kg \cdot deg)}$.

Polymethyl methacrylate with the equation of state (5), taken from Ref. 11, is used as the insulator.

The law of variation of the electrical conductivity of a shock-compressed MIM structure is taken in the form[8]

$$\sigma = \sigma_0 \exp(-\Delta E/kT), \tag{7}$$

where $\sigma_0 = 70591\ \Omega \cdot \mathrm{cm}$, and for polymethyl methacrylate we take $\Delta E = 1.2\ \mathrm{eV} \approx \mathrm{const}$ in the particle velocity range 1–10 km/s. The pressure in the region of the insulator depends on the thickness of the top plate of the MIM structure.

The quantity σ_0 and ΔE are found by substituting T_1, σ_1, T_2, and σ_2 into Eq. (7) and solving the system of equations (2). The values of U, E, p, and ρ, characterizing the behavior of the materials as a function of time, were determined by solving the system of equations (1)–(5). The program also permits calculating the temperature field in the cells. The solutions were used for calculations of the electrical conductivity and the conductivity, which depends on the temperature of the shock-compressed conduction channel of

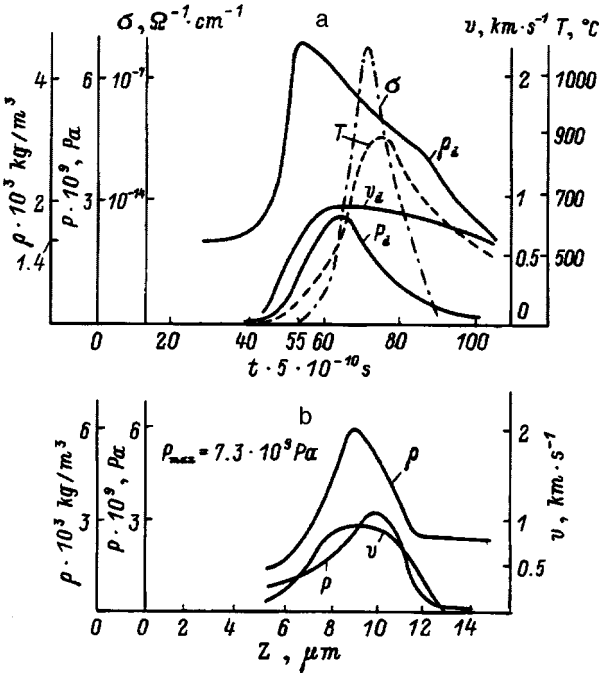

FIG. 1. a — Pressure, velocity, density, and temperature of the shock-compressed insulator versus time; b — pressure, velocity, and density of the insulator versus the longitudinal coordinate.

the MIM structure. Given the temperature σ is determined from the known function $\sigma=\sigma(T)$ or given σ and T the empirical dependence of σ on T can be determined. Figures 1a and b show the computed curves of the pressure, density, velocity, temperature, and electric conductivity of the shock-compressed insulator as functions of time and the longitudinal coordinate Z, respectively. The top plate, the insulator, and the bottom plate of the structure were 20, 1.5, and 20 μm thick, respectively, while the particles were 1.5 μm in size.

The computational results show that the film MIM structure is highly sensitive to shock action. Thus, for a 1.5 μm particle moving with velocity 5 km/s the electrical conductivity of the shock-compressed channel changes by 7 orders of magnitude. The temperature $T=900\,^{\circ}\mathrm{C}$, $\sigma=10^{-6}\,\mathrm{A}$, $\rho=4.5\times10^{3}\,\mathrm{kg/m^3}$, and $p_d=3\times10^{9}\mathrm{Pd}$, which in order of magnitude corresponds to the experimental results of Ref. 7.

There is no justification for using the solution of the systems of equations (1)–(5) for processing information from the MIM-based transducers under real operating conditions because of the large volume of the calculations and greater complexity of the apparatus. A simple model is required for calculating the current (voltage) pulse from the MIM structure as a function of time, the particle parameters, and the structural parameters of the transducer. An analytical expression for calculating the current in the conduction channel of a shock-compressed MIM structure was obtained in the form

$$J=\frac{U}{D^2}\int\int_{(V)}\int\sigma(\mathbf{r},t)dV, \tag{8}$$

by solving the equations of electrodynamics for an equivalent circuit of the transducer. Here J is the current through the MIM structure (total current through the shock-compressed capacitor), U is the voltage on the plates of the structure, V is the volume of the region of integration encompassed by the shock wave, $\sigma(\mathbf{r},t)$ is the electrical conductivity, d is the thickness of the insulator, and $\mathbf{r}$ and t are the coordinates and time.

The total conductance G of the conduction channel equals

$$G=\int\int_{(V)}\int\sigma(\mathbf{r},t)dV\cdot d^{-2}. \tag{9}$$

If the conductivity σ is a function of time t only, then Eq. (9) can be written in the form

$$G=\sigma(t)V(t)d^{-2}. \tag{10}$$

The volume $V(t)$ can be calculated approximately according to the equation

$$V(t)=S(t)d, \tag{11}$$

where $S(t)$ is the area of the spot formed by the shock wave on the back side of the insulator.

In this case, substituting expression (11) into Eq. (10) gives

$$G=\sigma(t)S(t)d^{-1}. \tag{12}$$

At $t=0$, $S(0)=0$ and $G(0)=0$, i.e., the conductance of the conduction channel itself equals zero. However, the conductance of the entire structure consists of the conductance of the rest of the insulator. Let S_g be the area of the entire surface of the insulator and S_c the area of the conduction channel. Then the total conductance will equal

$$G_g=\sigma(t)S_c(t)d^{-1}+\sigma_0'(S_g-S_c(t))d^{-1}, \tag{13}$$

where σ_0' is the conductivity of the undisturbed part of the insulator at room temperature.

The expression can be put into the form

$$G_g=G_0[1+(\sigma(t)/\sigma_0'-1)\cdot S_c(t)/S_g], \tag{14}$$

where G_0 is the conductance of the structure before the shock.

We shall make a number of assumptions which do not change the qualitative character of the processes occurring in the presence of a shock:

1) A spherical wave propagates away from the point of contact of the particle with the target. The amplitude of the wave can be calculated in accordance with the shock adiabat of the barrier material.

2) The distribution of the parameters from the contact point to the shock wave front is uniform.

3) The phase transition occurring during the passage of the shock waves is neglected.

4) The damping of the shock wave will be determined on the basis of the law of conservation of energy.

From the theory of shock waves the pressure on the front of the shock waves is

$$P=\frac{\rho_r W_0^2}{(1-\alpha_M)(1+\sqrt{\rho_r/\rho_m})^2}, \tag{15}$$

where ρ_r and ρ_m are the initial densities of the particle and target and W_0 is the collision velocity.

During the propagation of the shock wave the particle pumps in kinetic energy for a period of time t_0, so that the specific internal energy ε is constant in time. The distance H from the top of the contact surface to the free surface at the moment when the particle and the target no longer interact with one another will equal

$$H=R_r\sqrt{\rho_r/\rho_m}=R_r a, \tag{16}$$

where R_r is the characteristic particle size.

We estimate the time t_0 according to the formula

$$t_0=R_r(a+1)/W_0. \tag{17}$$

After the particle is spent, for $t>t_0$ the parameters of the shock wave decay as a result of an increase in the volume of the material. On the basis of the law of conservation of energy

$$E=\int_{(V)}\rho\varepsilon\,dV, \tag{18}$$

where E is the internal energy stored in a volume of radius H over a time t_0 and ρ is the density of the shock-compressed matter.

Since the quantities ρ and ε do not depend on the coordinates, we shall write the relation (18) in the form $E=\rho\varepsilon V(t)$. Since the wave is hemispherical, we have $V(t)=2/3\pi R^3(t)$, where $R(t)$ is the law of motion of the shock wave front. It is obvious that

$$R(t)=\int_{t_0}^{t}D(t')dt'+H, \tag{19}$$

where $D(t')$ is the velocity of the shock wave.

Assuming that for times close to T_0 the velocity of the shock wave is practically constant, we obtain

$$R(t)=D(t-t_0)+H. \tag{20}$$

The velocity D of the shock wave is related with the particle parameter as

$$D=\frac{W_0}{(1-\alpha_M)(1+a)}. \tag{21}$$

Substituting expressions (21) and (20) into Eq. (18) we obtain

$$E=\rho_M\varepsilon(t)\frac{2}{3}\pi\left[\frac{W_0(t-t_0)}{(1-\alpha_M)(1+a)}+H\right]^3. \tag{22}$$

Since

$$E=\frac{\rho_M}{2}\left(\frac{W_0}{1+a}\right)^2\frac{2}{3}\pi H^3=\text{const},$$

expression (22) makes it possible to determine the variation of the specific internal energy $\varepsilon(t)$. For Plexiglas the relation between the temperature T at the shock-wave front and the specific internal energy ε is given by expression (11)

$$T=k\varepsilon, \tag{23}$$

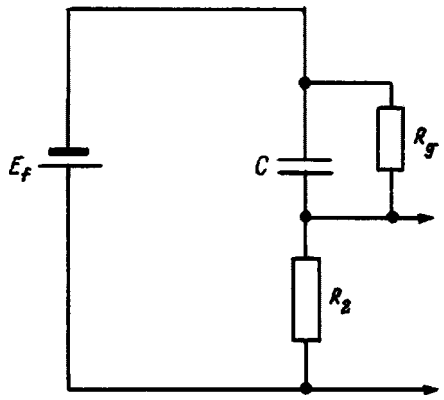

FIG. 2. Equivalent circuit of the shock-compressed MIM structure.

where $k=4.582\times10^{-8}$ K·g/erg is the coefficient of proportionality.

Substituting expression (23) into Eq. (22), we obtain the law of cooling of the insulator

$$T=T_0+\frac{E}{2/3\pi\rho_M[W_0(t-t_0)[(1-\alpha_M)(1+a)]^{-1}+H]^3k}. \tag{24}$$

We shall now determine the arrival time of the shock wave at the back surface of the insulator from the moment of impact

$$t_d=t_0+\frac{(L+d-H)(1-\alpha_M)(1+a)}{W_0}, \tag{25}$$

where d is the thickness of the insulator, located at a distance L from the source of the shock wave.

At this moment current starts to flow through the MIM structure. Switching to the new variable $t\to t+t_d$, we rewrite expression (24) in the form

$$T=T_0+\frac{E}{2/3\pi\rho_M[W_0t[(1-\alpha_M)(1+a)]^{-1}+L+d]^3k}, \tag{26}$$

where

$$\frac{2}{3}\pi\rho_M\left[\frac{W_0t}{(1-\alpha_M)(1+a)}+L+d\right]^3k=N.$$

We shall determine the area S_k of the spot on the back side of the insulator as

$$S_k=\pi[(Dt+L+d)^2-(L+d)^2]. \tag{27}$$

Substituting expressions (26) and (27) into Eq. (14), using expression (7) for the conductivity of the MIM structure, we obtain

$$G_g=G_0\left[1+\left(\frac{\sigma_0\exp(-\Delta E[T_0+EN^{-1}]^{-1})}{\sigma_0\exp(-\Delta ET_0^{-1})}-1\right)\times\frac{\pi[(Dt+L+d)^2-(L+d)^2]}{S_g}\right]. \tag{28}$$

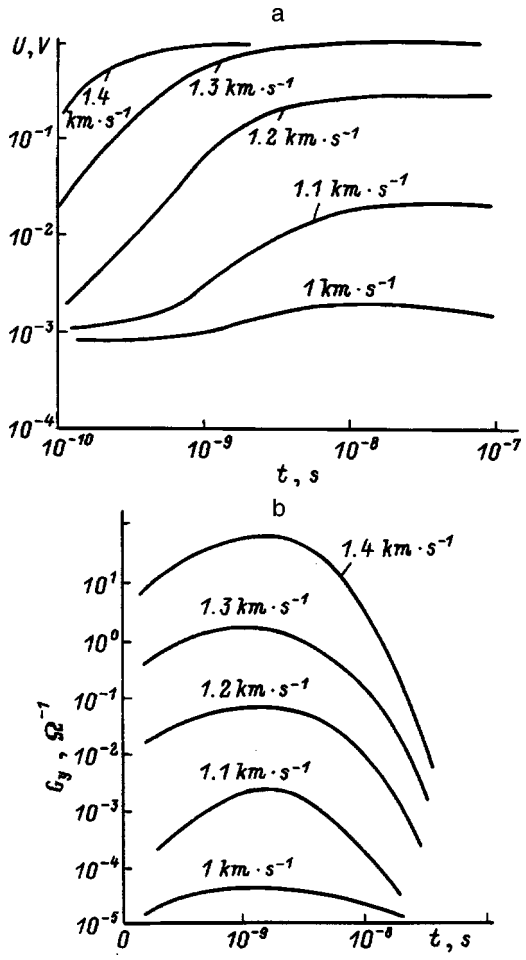

FIG. 3. a — Voltage of the shock-compressed capacitor versus time; b — conductance of the shock-compressed MIM structure versus time for different values of the particle velocity.

The expression (28) describes the time-dependence of the conductance of the MIM structure as a function of the particle parameters and the structural parameters of the structure. To obtain the voltage (current) pulse from the shock-compressed MIM structure, let us study its equivalent circuit (Fig. 2). Solving the Kirchhoff equation for the electrical circuit, we obtain a first-order linear differential equation for the output voltage

$$U'_{\text{out}}+\left(\frac{G_g(t)}{G_2}+1\right)\frac{U_{\text{out}}}{t_x}=\frac{G_g(t)E_f}{G_2 t_x}, \tag{29}$$

where $t_x=R_2C$, $G_2=1/R_2$, C is the capacitance of the MIM structure (capacitor), E_f is the voltage of the power supply, and R_2 is the load resistance.

The initial condition for Eq. (29) has the form at $t=0$

$$U_{\text{out}}(0)=U_{\text{st}}=\frac{R_2E_f}{R_2+R_g}=\frac{G_gE_f}{G_g+G_2},$$

where R_g is the resistance of the conducting channel.

The solution of Eq. (29) can be reduced to quadratures:

$$U=\exp\left(-\int_0^t\frac{G(t')}{G_2}+1\right)\frac{dt'}{t_x}$$
$$\times\left\{U_{\text{st}}+\frac{E_f}{t_x}\int_0^t\frac{G(x)}{G_2}\times\left[\exp\int_0^{t'}\left(\frac{GG''}{G_2}+1\right)dt''\right]dt'\right\}. \tag{30}$$

The results of solving Eq. (30) are presented in Fig. 3a (output voltage) and in Fig. 3b (conductance of the MIM structure). The results show satisfactory order-of-magnitude agreement with the experimental data[7] and the results obtained by solving the system of equations (1)–(5). Therefore the model proposed for calculating the conductance of a shock-compressed MIM structure (output signal) can be used, for example, for designing transducers of physical chemical parameters of high-velocity particles.

[1] V. N. Nikolaevskiĭ (Ed.), *High-Velocity Impact Phenomena*, Mir, Moscow, 1973.
[2] N. A. Inogamov, A. B. Konstantinov, S. I. Anisimov, and S. B. Zhitin, Zh. Éksp. Teor. Fiz. **99**, 1699 (1991) [Sov. Phys. JETP **72**, 948 (1991)].
[3] N. Pailer, J. Kissel, and E. A. Schneider, Space Sci. Instrum. **4**, 85 (1978).
[4] N. D. Semkin, S. M. Semenchuk, and T. Ya. Yusupov, Pis'ma Zh. Tekh. Fiz. **14**(6), 517 (1990) [Sov. Tech. Phys. Lett. **14**(3), 230 (1990)].
[5] R. Z. Sagdeev, I. Kissel *et al.*, Pis'ma Astron. Zh. **12**, No. 8 (1986) [Sov. Astron. Lett. **12**, 254 (1986)].
[6] L. S. Novikov, N. D. Semkin, and V. S. Kulikauskas, Fiz. Khim. Obrab. Mater., No. 6, 49 (1989).
[7] N. D. Semkin, Izv. Vyssh. Uchebn. Zaved. Priborostroenie **29**, 60 (1986).
[8] Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, Vols. 1 and 2 [Academic Press, New York, 1966, 1967; Nauka, Moscow, 1966, 607 pp.].
[9] O. M. Belotserkovskiĭ and Yu. M. Davydov, *The Method of Large Particles in Gas Dynamics* [in Russian], Nauka, Moscow (1982), 181 pp.
[10] L. V. Al'tshuller *et al.*, Zh. Éksp. Teor. Fiz. **38**, 791 (1960) [Sov. Phys. JETP **10**, 573 (1960)].
[11] S. Drapats and K. W. Michel, Z. Naturforsch., No. 29, 870 (1974).

Translated by M. E. Alferieff

Radiation-stimulated diffusion in solids

V. A. Stepanov

Physics and Power Engineering Institute State Science Center, 249020 Obninsk, Russia
(Submitted May 27, 1997)
Zh. Tekh. Fiz. **68**, 67–72 (August 1998)

Irradiation of solids produces a microscopic nonequilibrium state in which the vibrational energy distribution function of the atoms deviates from the thermodynamically equilibrium function. Expressions are obtained for the nonequilibrium distribution function and for the frequencies of activational transitions of atoms out of a potential well. It is shown that the radiation stimulation of diffusion processes involves a deviation of the temperature dependences of the frequencies of transitions of the atoms out of positions of equilibrium from the Arrhenius law. Under subthreshold irradiation conditions the rate of diffusion processes is higher for atoms whose vibrations thermalize over long times and depends linearly on the irradiation intensity. Under above-threshold irradiation conditions the characteristics of cascade regions in solids — their sizes and the vibrational excitation energy of the atoms — can be determined by comparing the computed and experimental temperature dependences of the diffusion coefficient. © *1998 American Institute of Physics.* [S1063-7842(98)01108-8]

INTRODUCTION

The kinetics of different activational processes in solids under irradiation can differ substantially from the kinetics of such processes in the absence of irradiation. Such differences are most strongly manifested experimentally as diffusion acceleration (radiation-stimulated diffusion) due to a factor of 10^2-10^{10} increase of the diffusion coefficient of the intrinsic and impurity atoms in metals, semiconductors, and wide-gap insulators.

Different mechanisms are invoked to explain this phenomenon. However, they cannot be used in all cases of radiation-stimulated diffusion. Examples are the mechanism of diffusion acceleration due to ''radiation jarring'' of solids accompanying the relaxation of the metastable states that are formed[1] or the crowdion mechanism of diffusion acceleration as a result of the displacement of entire groups of atoms.[2] Such notions are enlisted to explain radiation phenomena in metals under above-threshold irradiation conditions. The processes leading to the formation of defects in insulators as a result of the decay of electronic excitations are examined in Refs. 3 and 4. A mechanism of radiation-stimulated diffusion due to electronic charge transfer on impurities in semiconductors under subthreshold irradiation conditions is proposed in Ref. 5.

The most general approach from the standpoint of different types of radiation action and materials appears to be an approach based on analysis of the influence of vibrational excitations of atoms on statistical processes in solids. Indeed, aside from the contribution of radiation to defect formation by the impact mechanism and the formation of long-lived electronic excitations, a major fraction of the energy is converted directly or by collisional and electronic relaxation over times $10^{-15}-10^{-16}$ s into atomic vibrational energy. One can talk about a microscopic nonequilibrium state, characterized by the deviation of the vibrational energy distribution function of the atoms from a thermodynamically equilibrium function. The thermalization times of atomic vibrations (greater than 10^{-12} s) are much longer than the characteristic times of other microscopic dynamical radiation phenomena in solids, so that the need to take account of the influence of microscopic nonuniformity on diffusion processes seems obvious.

NONEQUILIBRIUM DISTRIBUTION FUNCTION

Let the subsystem of particles characterized by an energy distribution function $f(E)$ interact with a radiation flux I in a manner so that the energy transfer to the particles in the subsystem in the energy interval $(\varepsilon,\varepsilon+d\varepsilon)$ is determined by the differential cross section $d\sigma=\kappa(\varepsilon)d\varepsilon$. We shall use the approach of Refs. 6 and 7 to derive the kinetic equation for the distribution function $f(E)$. Let us represent the change produced in the distribution function by the interaction with the flux I as

$$\int_0^\infty i\kappa(\varepsilon)\left[\exp\left(\frac{-\varepsilon d}{dE}\right)f(E)-f(E)\right]d\varepsilon. \tag{1}$$

Here $\exp(-\varepsilon d/dE)$ is a shift operator

$$\exp\left(\frac{-\varepsilon d}{dE}\right)f(E)=f(E-\varepsilon)U(E-\varepsilon), \tag{2}$$

where

$$U(x)=\begin{cases}0, & x<0,\\ 1, & x>0.\end{cases}$$

Let the excited particles relax with thermalization time τ. We shall employ the τ-approximation, according to which the relaxation time does not depend on the excitation level of the particles. The equation for the distribution function is

$$\frac{\partial f(E)}{\partial t}=-\frac{1}{\tau}(f(E)-f_0(E))+I\int_0^\infty \kappa(\varepsilon)\times\left[\exp\left(\frac{-\varepsilon d}{dE}\right)f(E)-f(E)\right]d\varepsilon, \quad (3)$$

where $f_0(E)$ is the equilibrium distribution function.

We obtain for the nonequilibrium distribution function in a stationary state the expression

$$f(E)=\frac{1-\alpha}{1-\alpha/\sigma\int\kappa(\varepsilon)\exp\left(\frac{-\varepsilon d}{dE}\right)d\varepsilon}f_0(E), \quad (4)$$

where $\sigma=\int\kappa(\varepsilon)d\varepsilon$ is the total interaction cross section and $\alpha=\tau\sigma I/(1+\tau\sigma I)$.

Expanding expression (4) in a series we obtain

$$f(E)=(1-\alpha)\sum_{n=0}^{\infty}\left(\frac{\alpha}{\sigma}\right)^n\left[\int\kappa(\varepsilon)\exp\left(\frac{-\varepsilon d}{dE}\right)d\varepsilon\right]^n f_0(E) \quad (5)$$

or

$$f(E)=(1-\alpha)\sum_{n=0}^{\infty}\left(\frac{\alpha}{\sigma}\right)^n\int\kappa(\varepsilon_1)\ldots\kappa(\varepsilon_n)\times U(E-\varepsilon_1-\ldots\varepsilon_n)f_0(E-\varepsilon_1-\ldots\varepsilon_n)\times d\varepsilon_1\ldots d\varepsilon_n. \quad (6)$$

PROBABILITY OF A TRANSITION THROUGH AN ENERGY BARRIER

Let a particle whose energy is characterized by the distribution function $f(E)$ move in a potential well whose profile remains unchanged during the motion of the particle. Over a period of the oscillations the particle will leave the potential well if its energy is higher than the energy barrier W. Therefore, we shall estimate the frequency of transitions per unit time as

$$\omega=\nu\int_W^\infty f(E)dE, \quad (7)$$

where ν is the vibrational frequency of the particle in the potential well. In the equilibrium case $f_0(E)=Ae^{-E/T}$ and

$$\omega_0=\nu\ \exp\ \frac{W}{T}. \quad (8)$$

Under the conditions of microscopic nonequilibrium we obtain from Eq. (6)

$$\omega=\nu(1-\alpha)A\sum_{n=0}^{\infty}\left(\frac{\alpha}{\sigma}\right)^n\int_0^\infty\int_W^\infty\exp(\varepsilon_1+\ldots+\varepsilon_n-E)\times U(E-\varepsilon_1-\ldots-\varepsilon_n)\kappa(\varepsilon_1)\ldots\kappa\times(\varepsilon_n)dEd\varepsilon_1\ldots d\varepsilon_n=\omega_0(1-\alpha)\sum_{n=0}^{\infty}\left(\frac{\alpha}{\sigma}\right)^n\times\int\begin{cases}\exp((\varepsilon_1+\ldots+\varepsilon_n)/T), & W>\varepsilon_1+\ldots+\varepsilon_n\\ \exp(W/T), & W<\varepsilon_1+\ldots+\varepsilon_n\end{cases}\times\kappa(\varepsilon_1)\ldots\kappa(\varepsilon_n)d\varepsilon_1\ldots d\varepsilon_n. \quad (9)$$

The expression obtained for the probability of transitions of excited particles through an energy barrier can be used to calculate activational transitions under conditions of external excitations of atoms in solids, where the function $\kappa(\varepsilon)$ determines the energy transfer to the atoms.

The case $\kappa(\varepsilon)=\sigma\cdot\delta(\varepsilon-\varepsilon_0)$, where the differential cross section is described by a Dirac delta function, corresponds to resonance laser action and is studied in Ref. 8. The condition

$$\kappa(\varepsilon)=\sigma/\varepsilon_0, \quad (10)$$

with $0<\varepsilon<\varepsilon_0$ can obtain in solids irradiated by fluxes of high-energy particles (electrons, ions, neutrons, γ rays). Here σ is the total scattering cross section and ε_0 is the maximum energy transferred to the atoms. The atoms acquiring energy above the threshold energy leave the potential well by a mechanism different from an activational mechanism. For this reason, we shall assume that ε_0 does not exceed the threshold energy of a displacement of atoms in a solid. We shall also assume that the excitation energy of the atoms is still quite high and greater than the energy barriers for elementary activational processes in solids. As a rule, the threshold displacement energy for metals exceeds 15 eV, whereas for nonmetals it can assume values ranging from 6 eV (in InSb) up to 80 eV (in diamond),[9] while the activation energies do not exceed 2–3 eV. For $\varepsilon_0>W$ and under the condition (10) the integral appearing in Eq. (9) can be written in the form

$$J_n=\int_V\exp((\varepsilon_1+\ldots\varepsilon_n)/T)d\varepsilon_1\ldots d\varepsilon_n+\exp(W/T)\times\int_0^{\varepsilon_0}d\varepsilon_1\ldots d\varepsilon_n-\exp(W/T)\int_V d\varepsilon_1\ldots d\varepsilon_n, \quad (11)$$

where V is the volume in a n-dimensional space bounded by the coordinate planes and the hyperplane $\varepsilon_1+\ldots+\varepsilon_n=W$.

Performing the integration we obtain

$$J_n=(-T)^n+\varepsilon_0^n\exp\left(\frac{W}{T}\right)-(-T)^n\exp\left(\frac{W}{T}\right)\sum_{k=0}^{n}\frac{(-W/T)^k}{k!}. \quad (12)$$

Substituting expression (12) into Eq. (9) gives

$$\omega=\omega_0(1-\alpha)\sum_{n=0}^{\infty}\left[\left(\frac{-\alpha T}{\varepsilon_0}\right)^n+\exp\frac{W}{T}\alpha^n-\exp\frac{W}{T}\left(\frac{-\alpha T}{\varepsilon_0}\right)^n\sum_{k=0}^{n}\frac{(-W/T)^k}{k!}\right]=\omega_0\left\{(1-\alpha)+\exp\frac{W}{T}\left(1+\frac{\alpha T}{\varepsilon_0}-(1-\alpha)\exp\frac{\alpha W}{\varepsilon_0}\right)\right\}\times\left(1+\frac{\alpha T}{\varepsilon_0}\right)^{-1}. \quad (13)$$

For small values of α expression (13) becomes

$$\omega=\omega_0\left\{1+\alpha\exp(W/T)\frac{\varepsilon_0-W+T}{\varepsilon_0}\right\}. \quad (14)$$

From Eq. (14) follows the condition under which the microscopic nonequilibrium must be taken into account in order to calculate the probability of particle transitions through the barrier W:

$$\alpha\exp\frac{W}{T}>1. \quad (15)$$

SUBTHRESHOLD RADIATION-STIMULATED DIFFUSION

The diffusion coefficient D in solids is proportional to the hopping frequency ω of the atoms

$$D=bl^2\omega, \quad (16)$$

where l is the hopping length between stable neighboring positions of the atoms and b is a geometric factor.

According to Eqs. (14) and (16), the diffusion coefficient under microscopic nonequilibrium conditions is determined by the expression

$$D=D_e\left\{1+\alpha\exp(W/T)\frac{\varepsilon_0-W+T}{\varepsilon_0}\right\}, \quad (17)$$

where D_e is the diffusion coefficient under equilibrium conditions and ε_0 is the maximum energy transferred to the diffusing atoms by the particles of the radiation flux.

At high temperatures $D\sim D_e$. At low temperatures, where the condition (15) is satisfied, the diffusion coefficient is virtually temperature-independent and is proportional to the parameter α of nonequilibrium. This is an obvious result, showing that at low temperatures the probability of transitions of atoms out of a potential well is proportional to the probability of excitation of the atoms with energy greater than the energy barrier.

There are a large number of experimental results on subthreshold radiation-stimulation of diffusion, primarily in semiconductors. In the case of the irradiation of semiconductors microscopic nonequilibrium could be due to the excitation of atomic vibrations in the process of relaxation of the radiation-induced electronic states. An example is the result obtained in an experiment on the stimulation of diffusion in silicon by x rays.[10] After irradiation with intensity $I=5\times10^{12}\,\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}$ at temperatures $\sim300\,\mathrm{K}$ the concentration profiles of the distribution of gold in silicon corresponded to diffusion coefficients under equilibrium conditions at temperatures higher by 300 °C. It is easy to show that the relation (15) holds under such irradiation conditions. For diffusion activation energy $W=1.5\,\mathrm{eV}$, the x-ray scattering cross section $\sigma=10^{-19}\,\mathrm{cm}^2$ and the lifetime of vibrational excitations on impurity atoms $\tau=10^{-11}\,\mathrm{s}$, the parameter $\alpha\sim\tau\sigma I=5\times10^{-18}$, and $\alpha e^{W/T}>10^7$.

We note that from Eq. (17) follows a linear dependence of the diffusion coefficient on the irradiation intensity. Such a dependence has been observed, for example, in experiments on γ-ray stimulated diffusion of copper in cadmium sulfide.[11] In the opinion of the author of Ref. 11, the radiation-stimulated diffusion coefficient is proportional to the rate of change of the charge state of the impurity. Taking this into account, the acceleration of diffusion is due to vibrational excitations of impurity atoms as a result of electronic recombination processes. To make an accurate calculation of the parameter α and radiation-stimulated diffusion, the character of the electron-vibrational relaxation must be taken into account.

The dependence of the parameter α of nonequilibrium on the thermalization time τ of the atoms and on the cross section σ of the interaction of the atoms with the radiation flux leads to the possibility of selective stimulation of the diffusion of atoms in solids. The thermalization time of light impurity atoms is longer than that of the main lattice atoms. In Ref. 12 it was established experimentally that irradiation with γ rays, x rays, and heavy ions selectively stimulates hydrogen migration in semiconductors and insulators. The large cross section for freeing of hydrogen from traps ($10^{-15}\,\mathrm{cm}^{-2}$) indicates that the process occurs by a subthreshold mechanism. In Ref. 12 it is noted that a vibrationally excited bond containing hydrogen possesses local vibrational quanta with energy 0.2–0.5 eV outside the phonon spectrum of the crystal, and generation of three to five phonons by the multiplet mechanism is required to dissipate this energy. As a result of this, the diffusion process is activated by the excitation of the vibrations of hydrogen-containing bonds.

POINT DEFECT CONCENTRATION AND DIFFUSION DUE TO ABOVE-THRESHOLD IRRADIATION

The concentration of vacancies and interstitial atoms in simple irradiated materials in the case of above-threshold irradiation are determined by the standard macroscopic equations (see, for example, Ref. 13)

$$\frac{\partial C_i}{\partial t}=K-K_1(C_iC_v-C_{ei}C_{ev})-P_i\omega_i(C_i-C_{ei}),$$

$$\frac{\partial C_v}{\partial t}=K-K_1(C_iC_v-C_{ei}C_{ev})-P_v\omega_v(C_v-C_{ev}), \quad (18)$$

where K_1 is a coefficient characterizing the rate of mutual recombination of point defects; $P_{i,v}$ are the probabilities of capture of defects on sinks; $\omega_{i,v}$ are the hopping frequencies of defects; and, $C_{ei,ev}$ are the thermodynamically equilibrium concentrations of the corresponding defects.

In the stationary case the solution of Eqs. (18) is

$$C_i=\frac{2K(1+\beta)}{P_i\omega_i\eta}\left\{\left[1+\frac{\eta}{(1+\beta)^2}\right]^{1/2}-1\right\}+C_{ei},$$

$$C_v=\frac{P_i\omega_i}{P_v\omega_v}+C_{ev}, \quad (19)$$

where $\eta=4KK_1/(P_iP_v\omega_i\omega_v)$ and $\beta=K_1C_{ev}/(P_i\omega_i)$.

If the relation (19) is used neglecting the microscopic nonequilibrium to calculate the defect concentrations, then an unphysical result is obtained at low temperatures: The vacancy concentration is greater than the total number of atoms. Taking account of the microscopic nonequilibrium

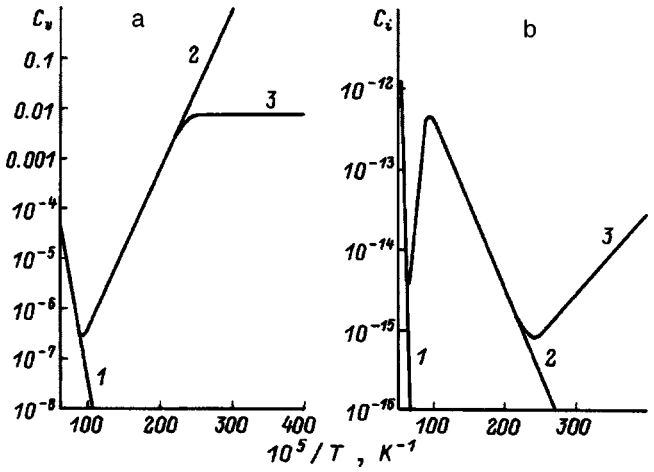

FIG. 1. Computed stationary concentrations of vacancies (a) and interstitials (b) in nickel versus the reciprocal of the temperature under equilibrium conditions (*1*) and under reactor irradiation conditions neglecting (*2*) and taking account of (*3*) the microscopic nonequilibrium with $\alpha=1.8\times10^{-15}$ (*3*).

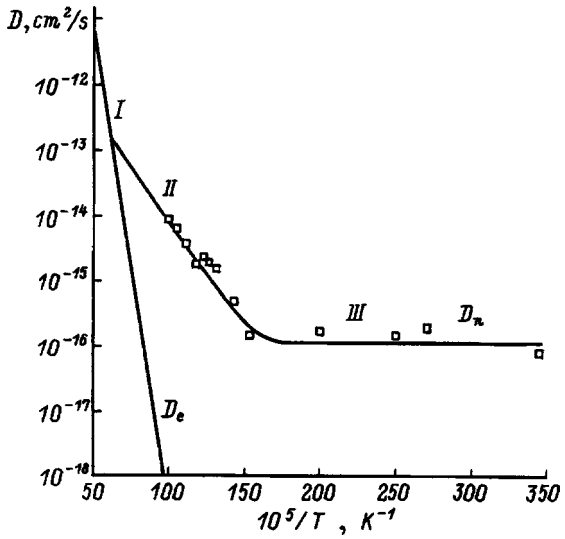

FIG. 2. Computed coefficients of diffusion of nickel in nickel versus the reciprocal of the temperature under equilibrium conditions (D_e) and under irradiation conditions (D_n) with $\alpha=10^{-10}$. □ — experimental values of the diffusion coefficient in the case of irradiation with Ni^- ions (300 keV, $I=6.3\times10^{12}\,cm^{-2}\cdot s^{-1}$).[14]

consists in using expression (14) instead of expression (9) to calculate the defect hopping frequency. In so doing, it can be assumed that the thermodynamically equilibrium defect concentrations do not change and their temperature dependences are described by the Arrhenius relations $C_{ev,i}=\exp(-V_{\nu,i}/T)$. Indeed, using expression (6) we obtain for the average energy of the nonequilibrium subsystem

$$\langle E\rangle=(1-\alpha)\sum_{n=0}^{\infty}\left(\frac{\alpha}{\sigma}\right)^n\int\kappa(\varepsilon_1)\ldots\kappa(\varepsilon_n)\times U(E-\varepsilon_1-\ldots\varepsilon_n)f_0(E-\varepsilon_1-\ldots\varepsilon_n)Ed\varepsilon_1\ldots\varepsilon_n$$
$$=\langle E\rangle_0+(1-\alpha)\sum_{n=0}^{\infty}\left[\frac{\alpha}{\sigma}\right]^n n\sigma^{n-1}\int\varepsilon\kappa(\varepsilon)d\varepsilon$$
$$=\langle E\rangle_0+\langle\varepsilon\rangle\frac{\alpha}{1-\alpha},\qquad(20)$$

where $\langle E\rangle_0=\int Ef_0(E)dE$ is the average energy of the equilibrium system (thermostat temperature); $\langle\varepsilon\rangle=\int\varepsilon\kappa(\varepsilon)d\varepsilon/\sigma$ is the average energy transferred to the particles in the subsystem.

Under real irradiation conditions (accelerator and reactor) α does not exceed 10^{-10} and $\langle\varepsilon\rangle$ is less than 100 eV. This means that under microscopic nonequilibrium conditions the average energy of a system of particles and together with it all thermodynamic quantities also vary by not more than 10^{-4}%.

Figure 1 shows the computed stationary concentrations of vacancies and interstitial atoms in nickel versus the reciprocal of the temperature neglecting and taking into account the microscopic nonequilibrium with parameter $\alpha=1.8\times10^{-15}$. In the calculations it was assumed that $K_i\simeq\omega_i$, $P_i=P_\nu=10^{-7}$ (annealed metal), $K=10^{-6}$ da/s (reactor irradiation), the defect formation and migration energies $V_i=5$ eV, $V_v=1.63$ eV, $V_{\nu m}=1.25$ eV, and $V_{im}=0.2$ eV, and the defect vibrational frequencies $\nu_v=5\times10^{13}\,s^{-1}$ and $\nu_i=10^{13}$ s. Analysis of the temperature dependences of the defect concentrations displayed in Fig. 1 shows that when the microscopic nonequilibrium is neglected, the stationary vacancy concentrations are substantially underestimated and the interstitial concentrations are substantially overestimated for the case of reactor irradiation in the temperature range below 400 K.

The parameter α of nonequilibrium under above-threshold irradiation conditions, where cascades of displacements appear, can be determined as the ratio of the total energy introduced by radiation into the solid, minus the energy of the defects and electronic centers (for insulators) produced, to the average excitation energy $\varepsilon_0/2$ of the atomic vibrations in the cascade region, i.e., we assume that all of the excess energy is expended on the excitation of atomic vibrations. In the absence of processes leading to the formation of electronic radiation centers (metals),

$$\alpha\cong2[E_m-g(V_i+V_v)]\frac{I\sigma\tau}{\varepsilon_0}\cong2K\tau\frac{2E_d}{\varepsilon_0},\qquad(21)$$

where E_m is the energy transferred to the atoms by the particles of the radiation flux; E_d is the threshold energy of displacement; $g\sim E_m/2E_d$ is the cascade function (the average number of displaced atoms per primary knocked out atom); and, $K=\sigma Ig$ is the dose accumulation rate (number of displacements per primary knocked out atom, da/s).

The ratio $2E_d/\varepsilon_0$ is the average number of excited undisplaced atoms per displaced atom in the cascade region, while $g2E_d/\varepsilon_0$ is the average number of atoms excited in the entire cascade region. These parameters can be estimated using the results of experimental measurements of radiation-stimulated diffusion coefficients.

Figure 2 shows the experimental values of the diffusion coefficient of the isotope ^{63}Ni in nickel under irradiation with 300 keV ^{58}Ni ions with flux $I=6.3\times10^{12}\,cm^{-2}\cdot s^{-1}$ as a function of the reciprocal of the temperature.[14] The figure

also shows the computed dependences of the diffusion coefficient of nickel in nickel under equilibrium conditions and under microscopic nonequilibrium conditions with parameter $\alpha = 10^{-10}$. The diffusion coefficients were calculated from the formula

$$D_n = 10^{-17}(C_v \omega_v + C_i \omega_i), \tag{22}$$

where the relation (14) was used for the defect hopping frequency, while the defect density was calculated according to Eq. (19). The diffusion coefficient D_e under equilibrium conditions was also calculated according to Eq. (22), where the standard Arrhenius dependences were used for the concentrations and the hopping frequencies. The above-indicated values of the energy barriers and vibrational frequencies of defects for nickel and the dose accumulation rate $K = 3 \times 10^{-2}$ da/s ($g = 3.7 \times 10^3$) were used in the calculation.

The temperature variation of the diffusion coefficient (Fig. 2) can be divided into three sections. The Arrhenius law (*I*) holds at high temperatures. As the temperature decreases, the deviation from the Arrhenius law is due to the production of radiation defects (*II*). At low temperatures diffusion is temperature-independent (*III*), because of the microscopically nonequilibrium state of the atoms in the solid.

Using the relation (21), we shall estimate the parameters of the cascade regions in nickel under accelerator irradiation conditions. Assuming that the thermalization time of the atoms falls into the range $3-10 \times 10^{-11}$ s, we estimate the average number of excited undisplaced atoms per displaced atom as $\sim 20-50$, and we estimate the maximum excitation energy of atomic vibrations in the cascade region to be $\varepsilon_0 \approx 1.5-5$ eV with $6-20 \times 10^4$ atoms in the cascade region.

CONCLUSIONS

The irradiation of solids with different types of radiation results in a state of microscopic nonequilibrium due to a deviation of the vibrational energy distribution function of the atoms from the thermodynamically equilibrium function. Under these conditions the temperature dependences for the frequencies of transitions of the atoms out of their positions of equilibrium are different from the Arrhenius law, especially at low temperatures. This is the reason for the subthreshold radiation-stimulated diffusion processes.

The microscopic nonequilibrium state is characterized by a parameter of nonequilibrium, which is proportional to the irradiation intensity and the thermalization time of the atomic vibrations. This gives rise to a linear dependence of the subthreshold stimulation of diffusion processes on the irradiation intensity as well as to the possibility of selective stimulation of the diffusion of atoms which thermalize over long times.

Taking account of the microscopic nonequilibrium state under above-threshold irradiation conditions, the macroscopic kinetic equations for the creation–annihilation of defects can be used to calculate the defect density at low temperatures. By comparing the computed and experimental temperature dependences of the diffusion coefficient it is possible to determine the parameter of nonequilibrium, from which it is possible to obtain the important characteristics of above-threshold actions on solids — the size and the atomic vibrational excitation energy of cascade regions.

[1] V. L. Indenbom, Pis'ma Zh. Tekh. Fiz. **5**, 489 (1979) [Sov. Tech. Phys. Lett. **5**, 200 (1979)].
[2] M. Gieb, J. Heieck, and W. Schule, J. Nucl. Mater. **225**, 85 (1995).
[3] B. Lushchik, I. K. Vital, and M. A. Élango, Usp. Fiz. Nauk **122**, 223 (1977) [Sov. Phys. Usp. **20**, 489 (1977)].
[4] M. A. Élango, *Elementary Inelastic Radiation Processes* [in Russian], Nauka, Moscow (1988), 150 pp.
[5] V. T. Mak, Zh. Tekh. Fiz. **63**, 173 (1993) [Sov. Phys. Tech. Phys. **38**, 246 (1993)].
[6] P. L. Bhatnagar, E. P. Gross, and M. Krook, Phys. Rev. **94**, 511 (1954).
[7] E. P. Gross and M. Krook, Phys. Rev. **102**, 593 (1956).
[8] V. A. Stepanov, ''Laser surface microprocessing,'' Proc. Soc. Photo-Opt. Instrum. Eng. (SPIE) **1352**, 208 (1990).
[9] J. W. Corbett and J. C. Bourgoin, in *Semiconductors and Molecular Crystals*, Vol. 2, edited by J. H. Crawford and L. M. Slifkin (Plenum Press, New York, 1975), pp. 1–161.
[10] V. S. Vavilov, A. E. Kiv, and O. R. Niyazova, *Mechanisms of the Formation and Migration of Defects in Semiconductors* [in Russian], Nauka, Moscow (1981), 368 pp.
[11] V. T. Mak, Neorg. Mater. **32**, 1184 (1996).
[12] I. P. Chernov, A. P. Mamontov, Yu. I. Tyurin, and Yu. P. Cherdantsev, in *Abstracts the 8th Conference on Radiation Physics and Chemistry of Inorganic Materials* [in Russian], Tomsk, 1993, Part 1, p. 124.
[13] I. A. Akhiezer and L. N. Davydov, *Introduction to the Theoretical Radiation Physics of Metals and Alloys* [in Russian], Naukova Dumka, Kiev (1985), 142 pp.
[14] M. P. Macht, A. Muller, V. Naundorf, and H. Wollehberger, Nucl. Instrum. Methods Phys. Res. B **16**, 148 (1986).

Translated by M. E. Alferieff

Defect-density switching wave in crystals under pulsed laser irradiation

F. Kh. Mirzoev

(Submitted May 28, 1997)
Zh. Tekh. Fiz. **68**, 73–77 (August 1998)

A model is proposed for the ignition of a defect-density switching wave in a crystal by powerful laser pulses. It is shown that the switching wave arises as a result of the nonlinear dependence of the activation energy of the defect-formation process on the strain field due to defects. The conditions under which a switching wave arises and the profile, velocity, and propagation direction of the wave are discussed. © *1998 American Institute of Physics.* [S1063-7842(98)01208-2]

INTRODUCTION

It is well known that powerful laser pulses acting on crystals generate generation of point defects (PDs) (vacancies, interstices) with concentration much higher than the equilibrium concentration. According to the thermal model, the formation of point defects is related with thermal fluctuations and its probability increases with the temperature of the medium (or the radiation intensity) and with the defect density n.[1] The effect of the defect density is due to the decrease in the height E of the activation barrier for defect formation due to the appearance of an elastic strain of the medium as a result of the difference of the covalent radii of the atoms of the matrix and the defects. Under certain conditions the nonlinear character of the defect generation process can lead to instability at the surface (or in the interior region) of the crystal together with the formation of different static spatially periodic structures.[1,2] In reality, however, the concentration dependence of the source function g can have a maximum due to an increase in the activation energy at sufficiently high defect concentrations.[3] Unique phenomena can be observed in this case — the presence of two stable states of equilibrium and the propagation of a point-defect density wave front (a solution of the traveling front type) through the crystal. During the propagation of such a wave the local defect density in the crystal switches from a certain minimum value $n_{\min}$ to a maximum value $n_{\max}$.

The aim of the present paper is to analyze the critical conditions of excitation of a switching wave and to estimate its profile and propagation velocity.

AUTOWAVE MODEL

This effect can be analyzed theoretically on the basis of a distributed nonlinear system of equations describing the dynamics of an ensemble of point defects taking account of their elastic interaction. If it is assumed that the main processes controlling the behavior of the defect density in time are generation, diffusion, and recombination on centers, then we have

$$\frac{\partial n}{\partial t} = g_0 \exp[-(E+W)/kT] + D\frac{\partial^2 n}{\partial r^2} - \beta n, \tag{1}$$

$$E = E_0 - \left\langle \sum_i U(r-r_i) \right\rangle \approx E_0 - zU(\bar{r}). \tag{2}$$

The first term on the right-hand side of Eq. (1) describes the generation of point defects by thermal fluctuations ($g_0 = d^3 \nu N_0^2$, d — lattice constant; ν — atomic vibrational frequency of the order of the Debye frequency: $\nu \approx V_{\text{ph}}/d \sim 10^{14}\,\text{s}^{-1}$; N_0 — density of lattice sites, V_{ph} — sound speed), the second term describes the spatial diffusion of point defects (D — diffusion coefficient), and the third term describes recombination on centers (recombination rate $\beta = \beta_0 \exp(-W/kT)$, $\beta_0 = \rho \nu d^2$, ρ — density of recombination centers, W — activation energy for diffusion of defects, k — Boltzmann's constant, T — temperature).

The equation (2) describes the renormalization of the defect activation energy E due to the appearance of an elastic strain field from the defects themselves. Here $U(r-r_i)$ is the strain field potential produced by the i-the defect; $\bar{r}$ is the average distance between defects; z is the coordination number; and, E_0 is the defect formation energy in the absence of interaction. The interaction between defects is described in terms of the elastic strain field in the nearest-neighbor approximation.

In the general case the strain field potential around a defect in crystals can be represented as[4]

$$U(r) = \frac{\alpha_1 r^m}{\alpha_2 + \alpha_3 r^{p+m}}, \tag{3}$$

where the exponent $p = 3$ or 6 and $m = 3$ ($\alpha_{1,2,3} = \text{const}$).[5]

Far from a defect ($r \to \infty$) expression (3) gives the asymptotic form $U(r) \to r^{-p}$, while close to a defect ($r \to 0$) either $U(r) \to r^m$ asymptotically or a singularity obtains, depending on the values of the constants. In the present paper we shall confine our attention to the most interesting special case:

$$U(r) = 2U_0 \frac{r/r_0^3}{I + (r/r_0)^6}, \tag{4}$$

where U_0 and r_0 are the characteristic amplitude and radius of interaction for the strain field.

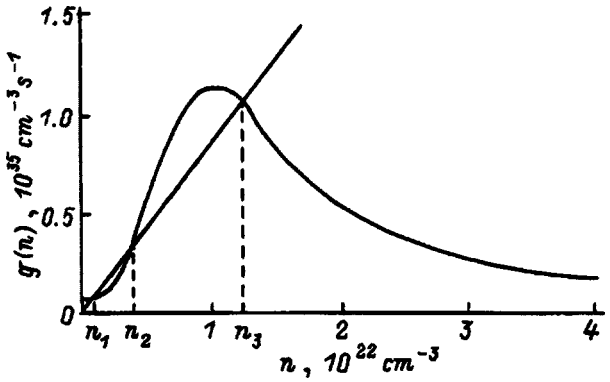

FIG. 1. Graphical solution of Eq. (7).

We note that although in Ref. 4 an expression of the type (3) was used to model the elastic interaction of defects in metals, it can be equally used for semiconductor systems.[6]

Substituting expression (4) in Eq. (2), we obtain for the activation energy for defect generation

$$E(n)=E_0-2zU_0\frac{n\Omega}{I+(n\Omega)^2}. \tag{5}$$

The derivation of expression (5) took account of the fact that $\overline{r}=n^{-1/3}$, $r_0\sim d$, and $\Omega\sim d^3$ is the volume of a defect.

At the same time as the defect formation energy decreases with the appearance of lattice strain, the defect migration energy also decreases, a result of which is that the diffusion coefficient increases and therefore the recombination rate increases. If it is assumed that $W=W_0-K\Omega_m^2 n$ (K — bulk modulus, Ω_m — activation volume for defect migration),[1] we obtain for the strain-accelerated recombination rate

$$\beta(n)=\beta\exp(K\Omega_m^2 n/kT),\quad \beta=\beta_0\exp(-W_0/kT). \tag{6}$$

To within the accuracy of the numerical values of the parameters of the model (1)–(6), the microscopic nature of the defects can be different in crystals with different short-range order structure or chemical-bond type. The equations (1)–(6) form a nonlinear dynamical system, characteristic for the physics of autowave processes in nonequilibrium media. The basic solutions, most often materializing in experiments on the observation of autowaves in such media, are solutions in the form of traveling excitation fronts and traveling pulses.[7] In many respects the system is similar to that studied in the theory of combustion[8] or the theory of nonequilibrium explosive crystallization in amorphous media.[9]

STABILITY ANALYSIS OF THE STATIONARY SOLUTIONS

In the stationary-homogeneous case we have from Eqs. (1)–(6)

$$g_0\exp\left[-\frac{E(n)}{kT}\right]=\beta_0 n. \tag{7}$$

The equation (7) can be solved graphically. In a certain range of values of the parameters of the system it can have from one to three real roots, i.e., $n=n_1$, $n_1<n_2<n_3$ (Fig. 1). Using a piecewise-linear approximation of the source function $g(n)$, it is easy to obtain

$$n_1\approx(G/3\beta_0)(1-3\sqrt{B\beta_0/G\Omega}),\quad n_2\approx G/3\beta_0,$$
$$n_3\approx(G/3\beta_0)(1+3\sqrt{B\beta_0/G\Omega}). \tag{8}$$

Here $G=g_0\exp(-E_0/kT)$, $B=2zU_0/kT$.

Let us investigate first the stability of the solutions of Eq. (1) with respect to a spatially periodic disturbance with wavelength λ

$$n'(r,t)=n_q\exp(\gamma t+iqr), \tag{9}$$

where the wave number of the disturbances $q=2\pi/\lambda$ and γ is the growth rate of the instability.

The equation linearized near the stationary solution (8) with respect to small nonuniform corrections $n'(r,t)$ can be written in the form

$$\frac{\partial n'}{\partial t}=D\frac{\partial^2 n'}{\partial r^2}+n'\frac{\partial S}{\partial n'}(n'=n_i), \tag{10}$$

where

$$S(n)\{g_0'\exp[Bn\Omega/(I+n^2\Omega^2)]-\beta n\}\exp(K\Omega_m^2 n/kT),$$
$$g_0'=G\exp(-W_0/kT).$$

Substituting expression (9) into Eq. (10) gives the dispersion relation

$$\gamma(q)=-Dq^2+\frac{\partial S}{\partial n'}(n'=n_i). \tag{11}$$

The stability of the solution is determined by the sign of the derivative in Eq. (11). It is obvious that at the points n_1 and n_3 the derivative is negative and the solution is asymptotically stable. At an intermediate point (n_2) $\partial S/\partial n'$ ($n'=n_2$)>0, if

$$n>n_*=kT/6zU_0\Omega. \tag{12}$$

The interval $0\leqslant q\leqslant q_0$ determines the region of unstable values of the wave numbers, where $q_0^2=D^{-1}(\partial S/\partial n')(n'=n_2)$. Thus, the left-hand ($n=n_1$) and right-hand ($n=n_3$) roots in Fig. 1 are stable, while the middle root ($n=n_2$) is unstable.

ESTIMATES OF THE PROFILE AND PROPAGATION VELOCITY OF A DEFECT WAVE

To analyze a solitary front traveling with constant velocity v, we switch in Eq. (1) to the self-similar variable η

$$\eta=r-vt,\quad n(r,t)=n(\eta). \tag{13}$$

Substituting expression (13) into Eq. (1) we obtain the equation

$$-v\frac{dn}{d\eta}=D\frac{d^2 n}{d\eta^2}+S(n). \tag{14}$$

The equation (14) with the boundary conditions $n(-\infty)=n_3$, $n(\infty)=n_1$, and $n'_\eta(\pm\infty)=0$ is an eigenvalue

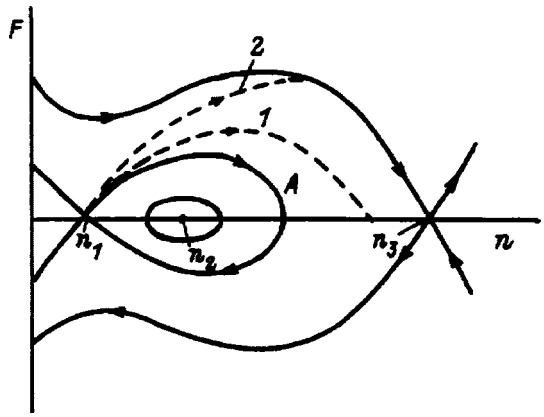

FIG. 2. Phase portrait of wave trajectories for the system of equations (15).

problem, whose spectrum determines the possible values of the velocities, while the corresponding eigenfunctions $n(\eta)$ determine the forms of the defect wave.

Setting $F=dn/d\eta$, we shall investigate Eq. (14) in the phase plane (F,n). First, we represent this equation as a system of equations

$$D\frac{dF}{d\eta}=-vF-S(n), \quad \frac{dn}{d\eta}=F. \tag{15}$$

The number of zeros of the function $S(n)$ depends on the temperature. For $T_1<T<T_2$, where T_1 and T_2 are the solutions of the equation

$$(3f^2-1)(kT)^2-3fk(3f\Delta+2zU_0)T$$
$$+3f\Delta(fE+2zU_0)=0,$$
$$f=\beta_0/\Omega g_0\Delta=E_0-W_0,$$

where Eq. (7) possesses three roots, there exists a single eigenvalue of the velocity v.[7] Figure 2 displays the characteristic form of the phase trajectories of the system of equations (15) in the $(n,dn/d\eta)$ plane for the case $v>0$. It can be shown that for $v>0$ the system of equations (15) always has a solution such that $n(\eta)\to n_3$ as $\eta\to-\infty$ and $n(\eta)\to n_1$ as $\eta\to+\infty$. Indeed, for $v=0$ any solution n satisfies the equation

$$D\left(\frac{dn}{d\eta}\right)^2+2\int_0^n S(\xi)d\xi=\text{const},$$

and the trajectory of the motion is one of the curves in Fig. 2. The phase trajectory A, described by the equation

$$D\left(\frac{dn}{d\eta}\right)^2+2\int_0^n S(\xi)d\xi=0,$$

leaves the left-hand saddle point $(n_1,0)$, goes around the point $(n_2,0)$, and returns to the point $(n_1,0)$. It does not reach the saddle point $(n_3,0)$, since there $\int_0^n S(\xi)d\xi\leqslant 0$, while $\int_0^{n_1}S(\xi)d\xi>0$.

The singular point $(n_2,0)$ (center type) corresponds to an unstable solution of Eq. (14). If $n(\xi)$ is a solution starting on the unstable manifold of the saddle point $(n_1,0)$, then from Eq. (14) we obtain that in the first quadrant

$$\frac{d}{d\eta}\left[\frac{D}{2}\left(\frac{dn}{d\eta}\right)^2+\int_0^n S(\xi)d\xi\right]=v\left(\frac{dn}{d\eta}\right)^2>0.$$

For small and positive values of v, the solution closely follows the curve A and reaches the n axis (curve *1*) at points $n<n_3$. However, if v is large, then the solution reaches the trajectory (curve *2*) given by the equation

$$\frac{D}{2}\left(\frac{dn}{d\eta}\right)^2+\int_0^n S(\xi)d\xi=\int_0^{n_1}S(\xi)d\xi.$$

Since the unstable manifold and the solution n are continuous functions of v, there exists a $v>0$ such that $n(\eta)\to n_1$ as $\eta\to+\infty$ and $n(\eta)\to n_3$ as $\eta\to-\infty$. For this solution the derivative $dn/d\eta$ approaches zero exponentially as $\eta\to\pm\infty$, since $(n_1,0)$ and $(n_3,0)$ are saddle points. Therefore the only stable stationary solution of Eq. (14) is described by a separatrix, passing from the saddle point $(n_3,0)$ into the saddle point $(n_1,0)$.

The problem of the propagation of the front of a solitary wave for the system (15) can be reduced to the boundary-value problem

$$DF\frac{dF}{dn}+vF-S(n)=0 \tag{16}$$

with the boundary conditions $F(n_1)=F(n_3)=0$.

In the general case the problem (16) is solved by numerical methods. To obtain results in an explicit form we approximate the function $S(n)$ by a cubic polynomial

$$S(n)=\theta(n-n_1)(n-n_2)(n_3-n),$$

where $\theta=\Omega^2\beta_0$.

In this case the problem (16) can be solved exactly. Setting $F=\delta(n-n_1)(n-n_3)$ in Eq. (16) and equating the coefficients of like powers of n, we have for the propagation velocity of the wave

$$v=(n_1+n_3-2n_2)\sqrt{(\theta D/2)}, \quad \delta=\sqrt{\theta/2D}. \tag{17}$$

Next, integrating $dn/d\eta=F$, we obtain the profile of the solitary wave

$$n(r,t)=n_1+\frac{n_3-n_1}{1+C\exp\left(\dfrac{r-vt}{L}\right)}. \tag{18}$$

The solution (18) corresponds to a wave of switching of the point-defect density from the minimum value n_1, which exists in the initial state $(t=0)$, to the maximum value n_3. The width of the region of switching of the defect density is

$$L=\sqrt{(2D/\theta)}/(n_3-n_1)\approx\frac{1}{6}\sqrt{kTD/zU_0g_0'\Omega}. \tag{19}$$

If at $t=0$ the wave is located at the origin of the coordinates, then $C=1$. A plot of the wave solution (18) for this case is shown in Fig. 3.

The velocity of a switching wave can be calculated by the approximate methods that are widely used in the theory of bistable active media (for the case of slow switching

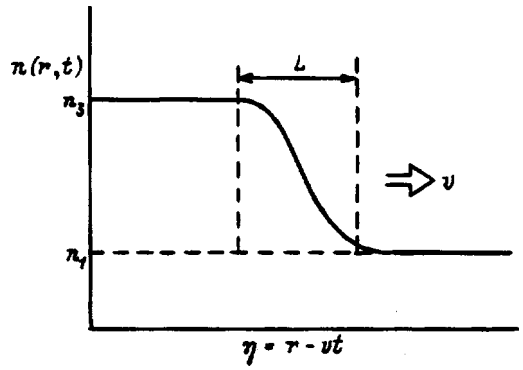

FIG. 3. A wave of switching of the point-defect density.

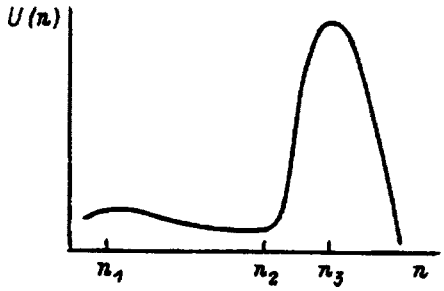

FIG. 4. The potential U versus the density n.

waves)[10] or in the theory of thermal propagation of flames (for the case of fast switching waves).[8] Let us represent Eq. (14) in the form

$$D\frac{d^2n}{d\eta^2}=-v\frac{dn}{d\eta}-\frac{dU}{dn},\quad U(n)\int_0^n S(\xi)d\xi. \tag{20}$$

The solution of Eq. (20) can be regarded formally as a trajectory of a classical particle of ''mass'' D moving with time η along the n axis in the potential $U(n)$ in the presence of friction forces, which are proportional to the particle velocity v, playing the role of a dynamic viscosity. If both sides of Eq. (20) are premultiplied by the derivative $dn/d\xi$ and integrated with respect to ξ over the range $(-\infty,+\infty)$, then taking account of the conditions $n(\eta)\to n_3$ as $\eta\to-\infty$ and $n(\eta)\to n_1$ as $\eta\to+\infty$ it is easy to obtain[10]

$$v=\int_{n_1}^{n_3}S(n)dn\left[\int_{-\infty}^{+\infty}\left(\frac{dn}{d\xi}\right)^2 d\xi\right]^{-1}. \tag{21}$$

Since the denominator in Eq. (21) contains the derivative $dn/d\eta$, which depends in turn on the velocity v of the switching wave, it is difficult to use this formula in practice. However, if the velocity of the wave is low, then $n(\eta)=n_0(\eta)$, where $n_0(\eta)$ is the profile of a standing switching wave, corresponding to $R=0$, where

$$R=\int_{n_1}^{n_3}S(\xi)d\xi=U(n_3)-U(n_1).$$

Then we obtain from Eq. (21)

$$v\approx R\left[\int_{-\infty}^{+\infty}\left(\frac{dn_0}{d\xi}\right)^2 d\xi\right]^{-1}.$$

The function $n_0(\eta)$ can be found from the equation

$$\eta=\sqrt{\frac{D}{2}}\int_{n_1}^{n}[U(n_1)-U(n)]^{-1/2}dn.$$

Therefore the velocity of slow switching waves is proportional to R.

In the case of fast switching waves, however, the potential $U(n)$ can have the form shown in Fig. 4. Following Ref. 8, we find for the velocity v

$$v\approx(2DR)^{1/2}\sqrt{\beta_0 kT\Omega/8zU_0g_0'}.$$

DISCUSSION

The solution (18) is a step moving with the velocity (17) with no change in profile and with a transitional region of characteristic width L. We note that the sign of the velocity changes at the point $n_1+n_3=2n_2$, which corresponds to the condition that the front of the switching wave is immobile. In a crystal stable, stationary zones with high and low content defect are possible.

The velocity of the wave decreases as the difference $U(n_3)-U(n_1)$ decreases and vanishes at $R=0$. If $R<0$, the switching wave moves in the opposite direction: $n(\eta)\to n_1$ as $\eta\to-\infty$ and $n(\eta)\to n_3$ as $\eta\to+\infty$, i.e., as the wave propagates, the state n_3 of the medium is replaced by the state n_1. If $R>0$, then as the wave passes, switching occurs from a low-defect state n_1 into a high-defect state n_3, and hence the wave propagating through the medium is a generation wave. However, if $R<0$, then after the wave passes the medium switches from a high-defect into a low-defect state and we have a recombination wave. Since R depends on the temperature (or the radiation intensity), the switching wave can be controlled by varying the temperature — the velocity and even the direction of motion of the wave can be changed. At low temperatures the wave will be a generation wave, while at high temperatures the wave will be a recombination wave.

To build up a defect generation wave the defect density must exceed a certain threshold value n_*, determined by Eq. (12). For the parameter characteristic for vacancy-type defects in Si $U_0=0.8\,\text{eV}$, $z=5$, $\Omega=10^{-22}\,\text{cm}^3$, $\rho=10^9\,\text{cm}^{-2}$, $D=1.5\exp(-0.95/kT)$,[13] $kT=0.03\,\text{eV}$, and $K=5\times10^{11}$ e/cm^3, we have $n_*\sim10^{19}\,\text{cm}^3$. On account of the passage of such a wave the local defect density in the crystal switches from the minimum value n_1 to the maximum value n_3. According to Eq. (17), the propagation velocity of a switching wave lies in the interval $v=10^{-3}-10^{-1}$ cm/s. An estimate of the width of the region of switching of the defect density using Eq. (19) gives $L=5\times10^{-6}$ cm. We obtain for the maximum density of defects produced by the wave the estimate $n_3\geqslant N_0$. At such point-defect densities in materials with covalent bonds in the crystal lattice (Si, Ge, Bi, and so on) the free energy of a crystal can reach the values necessary for amorphization in a solid-phase reaction regime. It is possible that in order for amorphization processes to develop not only must a critical point-defect density be reached but the point defects must then accumulate and form clusters, which can serve as centers of nucleation for the amorphous

phase. The total contribution of PDs and their clusters determines whether or not the energy state of the crystals above the energy for a supercooled liquid is attained. We note that every specific system has a set of physical factors which determine the possibility and mechanism of the transition from a crystalline into an amorphous state, and the factor determining the generation of a switching wave is one of them. It is obvious that to reach an amorphous state the local defect density must satisfy the condition $n > n_2$. The model proposed in the present paper for the ignition of a switching wave predicts realistically attainable values of the threshold defect density and could be of definite interest for describing the first stage of amorphization — attainment and preservation of the high-energy state of crystals with values of the free energy above that of a supercooled liquid. The physical picture developing at the second stage of the process — the stage of nucleation and growth of the amorphous phase — must be analyzed separately for a specific system and for specific transformation mechanisms.

Defect wave front propagation is an autowave process in a nonlinear active medium. It corresponds to a wave of switching between two stable stationary states, one of which (corresponding to the maximum defect density n_3) can be identified with the amorphous phase. The velocity and profile of this wave are determined uniquely by the characteristics of the medium and of the laser irradiation. The wave of switching of the defect density is similar to switching of the temperature in a combustion process.[8] Autowave processes of this kind have also been observed in the thermal propagation of the normal phase in high-temperature superconductor films[12] and in dielectrics during electrothermal heating on account of the nonmonotonic temperature dependence of the dielectric loss factor.[13]

I wish to thank L. A. Shelepin for a discussion of the results obtained in this work.

[1] F. Kh. Mirzoev, V. Ya. Panchenko, and L. A. Shelepin, Usp. Fiz. Nauk **166**, 3 (1996).
[2] F. Kh. Mirzoev, Kvantovaya Élektron. (Moscow) **23**, 827 (1996).
[3] F. Kh. Mirzoev and L. A. Shelepin, Pis'ma Zh. Tekh. Fiz. **22**(13), 28 (1996) [Tech. Phys. Lett. **22**(7), 531 (1996)] .
[4] Yu. A. Osip'yan (Ed.), *Computer Simulation of Defects in Metals* [in Russian], Nauka, Leningrad (1990), pp. 188.
[5] R. A. Masumura and G. Sines, J. Appl. Phys. **41**, 3930 (1970).
[6] I. M. Torrens, *Interatomic Potentials*, New York (1972), pp. 247.
[7] A. V. Vasil'ev, Yu. M. Romanovskiĭ, and V. G. Yakhno, *Autowave Processes* [in Russian], Nauka, Moscow (1987), pp. 82.
[8] Ya. B. Zel'dovich, G. I. Barenblatt, V. B. Librovich, and G. M. Makhviladze, *Mathematical Theory of Combustion and Explosion* [in Russian], Nauka, Moscow (1980), 478 pp.
[9] V. A. Shklovskiĭ and V. M. Kuz'menko, Usp. Fiz. Nauk **157**, 311 (1989) [Sov. Phys. Usp. **32**, 163 (1989)].
[10] B. S. Kerner and V. V. Osipov, *Autosolitons* [in Russian], Nauka, Moscow, 1989.
[11] B. I. Boltaks, *Diffusion and Point Defects in Semiconductors* [in Russian], Leningrad, 1972.
[12] V. A. Al'tov, Yu. M. L'vovskiĭ, and V. V. Sychev, Pis'ma Zh. Tekh. Fiz. **15**, 34 (1989) [Sov. Tech. Phys. Lett. **15**, 52 (1989)].
[13] P. N. Bondarenko, O. A. Emel'yanov, and S. N. Koĭkov, Pis'ma Zh. Tekh. Fiz. **15**, 45 (1989) [Sov. Phys. Tech. Phys. **15**, 640 (1989)].

Translated by M. E. Alferieff

Inverse problems of the nonlinear electrodynamics of high-T_c superconductors

K. P. Gaĭkovich and A. N. Reznik

Institute of Microstructure Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia
(Submitted June 10, 1997; resubmitted October 20, 1997)
Zh. Tekh. Fiz. **68**, 78–86 (August 1998)

A method is developed for reconstructing the dependence of the surface impedance of $Z_s(H)$ of high-T_c superconductors (HTSCs) on the amplitude of an ac magnetic field from measurements of the nonlinear radiophysical characteristics of microwave HTSC resonators. It is shown that the nonuniformity of the structure of the electromagnetic field in a resonator lead to integral equations for $Z_s(H)$. The corresponding integral equations are obtained for certain types of resonators. The errors of the conventional methods employing algebraic relations for $Z_s(H)$ are analyzed. The Tikhonov regularization method, developed in the theory of ill-posed inverse problems, is used to reconstruct the functions $Z_s(H)$ from the integral equations. Numerical experiments on the reconstruction of $Z_s(H)$ were performed, and the requirements for the accuracies and other characteristics of the experimental data were determined from them. [S1063-7842(98)01308-7]

INTRODUCTION

In the last few years, the nonlinear electromagnetic properties of HTSCs have become a subject of active investigations. There are several reasons for the interest which has developed in this problem. In the first place, HTSCs have promising applications in many microwave devices: resonators, filters, antennas, delay lines, multiplexers, and others.[1,2] The nonlinearity of superconductors limits the dynamical range of these devices. In the second place, such changes can serve as a means of quality control of HTSCs, which in large measure determines their nonlinear properties.[3,4] Finally, these investigations are also important for the physics of superconductors, since at present there is no consensus concerning the mechanism of nonlinearity.[5–11] The following are cited as likely reasons:[10] destruction of Cooper pairs by a microwave field, penetration of Abrikosov vortices into the mass of a superconductor or into the intergranular contact region, global heating of HTSCs by an electromagnetic field, or local heating of individual nonsuperconducting defects.

The nonlinear properties of HTSCs are ordinarily characterized by the dependence of the surface impedance $Z_s = R_s + iX_s$ (R_s, X_s — surface resistance and reactance, respectively) on the amplitude of the ac H field at the surface of the superconductor, i.e., $Z_s = Z_s(H)$. As a rule, the function $Z_s(H)$ is measured using different microwave resonators: microstrip,[6,12,13] cavity,[14] dielectric,[11,15,16] and quasi-optical.[17] The measured characteristics of the resonators are the coefficients of reflection, transmission, and absorption and their dependences on the frequency f and power P_{in} of the radiation supplied from the outside, i.e., the integral nonlinear response of the electrodynamic system as a whole is studied. The dependences $R_s(H)$ and $X_s(H)$ are usually determined from the broadening Δf_B of the frequency characteristic and the shift Δf_0 of the central frequency. A characteristic feature of all resonators listed above is nonuniformity of the distribution of the fields H over the surfaces of the HTSCs. At the same time, the algebraic relations between Δf_B, Δf_0 and $R_s(H)$, $X_s(H)$ which are valid for superconductors which are uniform with respect Z_s were used in works on the diagnostics of the nonlinear properties of HTSCs.[4–6,15,18] The nonlinearity of a superconductor and the nonuniformity of H lead to integral equations for $Z_s(H)$. Ignoring this fact results in serious errors in determining $Z_s(H)$.

In the present paper a new approach, based on solving the integral equations for $Z_s(H)$, is developed for the problem of the diagnostics of the nonlinearity of HTSCs with the aid of microwave resonators. The physical and practical advantages of this approach are discussed, integral equations for certain types of resonators are derived, the errors of the conventional methods are analyzed, numerical schemes for inverting the integral equations are constructed, and the requirements for accuracy and degree of discretization of the experimental data are determined on the basis of numerical experiments on the reconstruction of $Z_s(H)$.

NONLINEAR SURFACE IMPEDANCE OF HTSCS

In linear electrodynamics Z_s is determined as the coefficient of proportionality between the tangential components of the amplitudes of the electric field E_τ and magnetic field $H_\tau = H$ at the surface of a medium

$$E_\tau = Z_s H. \tag{1}$$

For a medium with good conductivity Z_s characterizes the conducting properties of the medium, which are referred to a point on the surface, i.e., it is a parameter of the material. The nonlinearity of a conductor leads to a self-action effect of the field on the frequency of the pump wave that can likewise be described by the relation (1) by means of the surface impedance, which is now a function of the field, i.e., $Z_s = Z_s(H)$. It is this effect that is ordinarily used as a test for nonlinearity of HTSCs, which is judged according to the

character of the dependence $Z_s(H)$ (we assume that the HTSC is a locally nonlinear medium, i.e., Z_s depends on H only at the same point of the surface). On the basis of this definition $Z_s(H)$ is still a characteristic of the conducting microwave properties of the HTSC material.

The methods[4–6,15,18] used to determine $Z_s(H)$ are based on the equation

$$\Delta f=\Delta f_0+\frac{i}{2}\Delta f_B=\frac{i}{8\pi W}\int_S H^2 Z_s d^2 r, \tag{2}$$

where the field H has a zero phase; W is the energy stored in the resonator; the integration extends over the surface of the HTSC; Δf_B, Δf_0 are determined relative to the values corresponding to an ideal conductor ($Z_s=0$). For impedance Z_s uniform over the surface of the HTSC, expression (2) becomes

$$\Delta f=\frac{i}{2}GZ_s, \tag{3}$$

where

$$G=\frac{1}{4\pi W}\int_S H^2 d^2 r \tag{4}$$

is a geometric factor, which can be calculated for each specific resonator or measured by calibration.

The equation (3) can be used, quite justifiably, to study the linear characteristics of HTSCs.

The situation is different in the case of a nonlinear resonator, where the nonuniformity of the field structure results in nonuniformity of the distribution of Z_s over the surface of the HTSC, since $Z_s=Z_s(H)$. In this case, the use of Eq. (3), as done in Refs. 4–6, 15, and 18, leads to a definition of the impedance $\langle Z_s\rangle$ averaged over the surface, i.e.,

$$\Delta f=\frac{i}{2}G\langle Z_s\rangle. \tag{5}$$

In accordance with Eq. (2), we have the average impedance

$$\langle Z_s\rangle=\frac{\int_S H^2 Z_s(H) d^2 r}{\int_S H^2 d^2 r}, \tag{6}$$

which is now determined not only by the properties of the material but also by the nonuniformity of H, which in turn depends on the type of resonator and the excited mode. The latter circumstance decreases the practical value of the results obtained using Eq. (3) and makes it difficult to compare them when the measurements are performed in different types of resonators.

THE INTEGRAL EQUATION

We shall assume that the resonator has one HTSC film as a conducting wall, whose impedance $Z_s(H)$ is to be determined. The structure of the magnetic field on the surface of the HTSC near the resonance frequency is determined, provided that Q of the nonlinear resonator remains high, by the characteristic function $\Phi(\mathbf{r})$ of the corresponding mode

$$H(\mathbf{r})=H_m\Phi(\mathbf{r}), \tag{7}$$

where $\mathbf{r}$ is a coordinate vector on the surface of the HTSC and H_m is the maximum value of the field H (Φ is normalized so that $\Phi_{\max}=1$).

The field H_m of the nonlinear resonator is calculated for each value of the power P_{in} according to the measured value of the loaded Q. For example, the method for a microstrip resonator is described in Ref. 19. Switching in Eq. (6) from integration over the coordinates to integration over H and substituting variables in accordance with Eq. (7) we obtain the equation

$$\langle R_s(H_m)\rangle=\int_0^{H_m} K(H_m,H)R_s(H)dH, \tag{8}$$

where $K(H_m,H)$ is the kernel of the integral equation. The form of the kernel depends on the type of resonator and the excited mode, and in addition $\int_0^{H_m}K(H_m,H)dH=1$.

The equation (8) was written for the surface resistance $R_s=\text{Re}(Z_s)$, the possibility of determining which we shall investigate below. Obviously, all subsequent conclusions can be transferred to $X_s=\text{Im}(Z_s)$ since the corresponding equations are identical. We shall present expressions for the kernel $K(H_m,H)$ and the quantity (8) for certain types of resonators.

a) *Quasioptical confocal resonator.* In resonators of this type[17] the HTSC film is installed as a flat reflector in the focal plane of a spherical mirror, consisting of a normal metal or a low-temperature superconductor. The field (7) on the HTSC is ordinarily approximated by a radially symmetric Gaussian function

$$\Phi(r)=\exp(-r^2/a^2). \tag{9}$$

Changing in Eq. (6) the variable of integration as $r\to H$, we obtain

$$K(H_m,H)=2H/H_m^2. \tag{10}$$

b) *Cavity resonator.* Let us consider a square resonator, one wall of which is a HTSC plate. For the fundamental mode the field H (7) in a rectangular coordinate system with origin at the center of the plate has the structure

$$\Phi(\mathbf{r})=\Phi(x)=\cos\left(\frac{\pi}{l}x\right), \tag{11}$$

where l is the length of the plate in the direction x.

Making the substitution $x\to H$ in Eq. (6) yields

$$K(H_m,H)=\frac{4}{\pi H_m}\frac{(H/H_m)^2}{\sqrt{1-(H/H_m)^2}}. \tag{12}$$

We note that the cavity resonators as well as dielectric resonators with modes having close structures have been used most often in recent years for diagnostics of the nonlinear properties of the HTSCs.[11,14–16] However, the works indicated all used cylindrical resonators, in which the TE_{011} mode with a radially symmetrical field structure

$$\Phi(r)=J_1\left(\frac{A_{01}}{\rho}r\right)/J_1^{\max}, \tag{13}$$

where ρ is the radius of the cylinder, $A_{01}\approx 3.832$, and $J_1(x)$ is a Bessel function of the first kind, was excited.

Clearly, the kernel of Eq. (8) for such resonators is different from expression (12), but as calculations show the normalized integrands of Eq. (6) for $\Phi(\mathbf{r})$ of the types (11) and (13) in the coordinate variables x and r, respectively, are close to one another, if $l=\rho$. On this basis the results obtained below for the kernel (12) can be taken as valid for all of the resonators listed here.

c) *Microstrip resonator*. In resonators of this type the HTSC film is a thin, usually curved, strip that serves as one of the conductors of a segment of a microstrip line.[6,12,13] The field H has a quite complicated two-dimensional structure, which for the TEM mode can be represented in the form

$$\Phi(x,y)=\varphi(y)\cos\left(\frac{\pi}{l}x\right), \tag{14}$$

where the coordinate x is measured in the direction of the line of symmetry of the strip. The transverse dependence $\varphi(y)$ is calculated using quite unwieldy numerical algorithms.[19] The approach considered here simplifies the analytical approximation of the function $\Phi(\mathbf{r})$. To this end, we employ the solutions of Ref. 20, which were obtained in the approximation of zero strip thickness $t=0$. These solutions hold for all y, excluding narrow regions near the strip edges $w/2-\zeta\leqslant|y|\leqslant w/2$, where w is the strip width and $\zeta\approx 2\Lambda^2/t$ (Λ — London penetration depth of the field in the superconductor).[21] In the limit $|y|\to w/2$ these solutions have a singularity $\varphi\to\infty$. We shall use the approximation proposed in Ref. 21, which reduces to limiting the region of integration in Eq. (6) to $0\leqslant|y|\leqslant w/2-\zeta$. Specifically, for a resonator based on a symmetric microstrip line, investigated in Refs. 5 and 12, we obtain

$$\varphi(y)=\sqrt{\frac{1-\exp(-2\pi\zeta/d)}{1-\exp(-2\pi(w/2-|y|)/d)}}, \tag{15}$$

where d is the thickness of the line.

Formula (15) is applicable for $w>0.3d$.[20] For purposes of estimation, we shall determine the parameter ζ in Eq. (15) using the computational results for the transverse structure of the current density on the surface of the strip.[19] For $t=0.2$ μm, $\Lambda=0.3$ μm, and resonator configuration identical to that of Ref. 19 with $w=150$ μm and $d=500$ μm, we find $\zeta=0.15$ μm.

We shall now present an expression that is obtained for the kernel $K(H_m,H)$ by making the substitution of variables of integration in Eq. (6) $x,y\to H=H(x,y)$, $y'=w/2-y$ taking account of the evenness of the function (15):

$$K(H_m,H)=\frac{\tilde{K}(H_m,H)}{\int_0^{H_m}\tilde{K}(H_m,H)\,dH}, \tag{16a}$$

$$\tilde{K}(H_m,H)=\begin{cases} H^2\int_\zeta^{w/2}dy'\,J(H,y'), & 0\leqslant H\leqslant H_w,\\ H^2\int_\zeta^{\psi(H)}dy'\,J(H,y'), & H_w\leqslant H\leqslant H_m,\end{cases} \tag{16b}$$

$$\psi(H)=-\frac{d}{2\pi}\ln\left[1-\left(\frac{H_m}{H}\right)^2(1-\exp(-2\pi\zeta/d))\right], \tag{16c}$$

$$J(H,y')=\sqrt{\frac{[1-\exp(-2\pi y'/d)]/[1-\exp(-2\pi\zeta/d)]}{\{1-(H/H_m)^2[1-\exp(-2\pi y'/d)]/[1-\exp(-2\pi\zeta/d)]\}}}, \tag{16d}$$

$$H_w=H_m\sqrt{[1-\exp(-2\pi\zeta/d)]/[1-\exp(-\pi w/d)]}. \tag{16e}$$

Using Eq. (8), the expressions for $K(H_m,H)$, and the algebraic equation (3) we calculated the relative error Δ in determining $R_s(H)$

$$\Delta=(R_s(H)-\langle R_s(H)\rangle)/R_s(H) \tag{17}$$

for a nonlinearity of the HTSC in the form

$$R_s(H)=R_0(1+bH^n). \tag{18}$$

For sufficiently strong fields, such that $bH^n\gg 1$, the computational results for $\Delta(n)$ are presented in Fig. 1 for different types of resonators. For $H<100$ Oe HTSC films usually have $n=2-4$,[5–7] and R_s exceeds $\langle R_s\rangle$ by a large amount: by a factor of 1.3–1.7 for cavity resonators, by a factor of 2–3 for quasioptical resonators, and by a factor of 4.5–7 for microstrip resonators. In the latter case the large values of Δ are due to the sharp growth of H at the edges of the strip (edge effect). Figure 1 also illustrates the degree to which the results obtained for the same HTSC film using different resonators, where using Eq. (3) $\langle R_s\rangle$ is determined instead of R_s, differ. The use of the algebraic equations (3) leads to the fact that the results with respect to the degree of nonlinearity of the HTSC are underestimated. An example of the correction of the experimental results which were obtained in Ref. 5 for a microstrip resonator is presented in Fig. 2, which displays curves of the coefficient b for the nonlinearity (18) for $n=2$.

METHOD FOR SOLVING THE INVERSE PROBLEM

The method proposed in the present work consists in using a HTSC resonator to measure $\langle R_s(H_m)\rangle$ for several values of P_{in} and determining $R_s(H)$ by inverting Eq. (8). The equation (8) is a nonlinear Volterra integral equation of

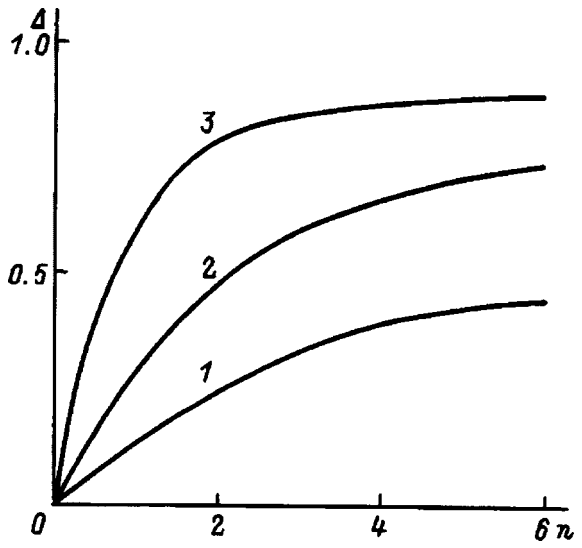

FIG. 1. Relative deviation Δ versus the exponent n in Eq. (18) for different types of resonators. *1* — Cavity and dielectric, *2* — confocal, *3* — microstrip.

the first kind. It is known[21,22] that the problem of obtaining the solution of an integral equation is, as a rule, ill-posed, in solving which without using additional (*a priori*) information small errors correspond to arbitrarily large reconstruction errors. This is why existing additional information about the properties of the exact solution must be invoked in order to solve the inverse problem. The choice of the specific solution algorithm also depends on the form of the information used. If the desired distribution can be described explicitly as a known function with definite parameters, a very effective method is to reduce the problem to a system of equations for these unknown parameters and to solve this system, for example, by the least-squares method, as done for the problem at hand in Ref. 23. However, if the desired distribution is not described in reality by the prescribed function, the errors can be arbitrarily large.

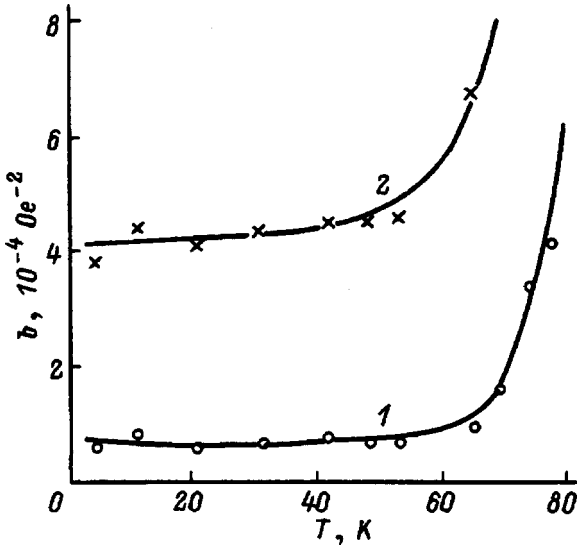

FIG. 2. Correction of the temperature dependence of the coefficient b in Eq. (18). *1* — Data of Ref. 5, *2* — corrected results.

Rigorous mathematical approaches to obtaining a solution are based on the application of diverse algorithms for regularizing the problem. In the present paper the solution is based on an application of the Tikhonov principle of the generalized residual (Ref. 22, p. 101). This principle employs very general information about the quadratic integrability of the exact solution and its derivative that solves the problem in the class of continuous functions and agrees well with the specific nature of the problem at hand.

Let us rewrite Eq. (8) in operator form (the index s is dropped here and below)

$$KR = R_m^{\delta}, \tag{19}$$

where R_m^{δ} is the vector of data obtained with a certain error, a measure of which in the Tikhonov method is the integral error δR_m defined as

$$\delta R_m^2 = \|KR - R_m^{\delta}\|_{L_2}^2 = \frac{1}{H_m^{\max}} \int_0^{H_m^{\max}} [R_m(H_m) - R_m^{\delta}(H_m)]^2 dH_m, \tag{20}$$

where R_m is the right-hand side of Eq. (19), which corresponds to the exact solution $R(H)$; $\|f\|_{L_2}^2$ is the norm of the function f in the space L_2 (space of quadratically integrable functions[22]); $H_m^{\max}$ is the maximum value of H_m that corresponds to the largest value of P_{in}.

In the Tikhonov method[22] the approximate solution minimizes the smoothing functional

$$M^{\alpha}(R) = \|K_h R - R_m^{\delta}\|_{L_2}^2 + \alpha \|R\|_{W_2^1}^2, \tag{21}$$

i.e., the desired solution can be found by searching for a function $R(H)$ which gives the functional (21) a minimum. In the relations presented above

$$\|R\|_{W_2^1}^2 = \frac{1}{H^{\max}} \int_0^{H^{\max}} \left[R(H)^2 + \left(H^{\max} \frac{dR(H)}{dH} \right)^2 \right] dH$$

denotes the norm of the function $R(H)$ as an element of the functional space W_2^1 (space of quadratically integrable functions which possess quadratically integrable derivatives[22]), $H^{\max} = H_m^{\max}$. The problem of minimizing a convex functional, which Eq. (21) is, reduces, after the appropriate discretization, to its finite-dimensional analog consisting of a computationally well-studied problem of quadratic programming and can be solved by standard gradient methods. In the present work we used the method of conjugate gradients, a description of which can be found in, for example, Ref. 24 (p. 273) (a Fortran algorithm for it is presented in Ref. 22). The numerical implementation of the method in the present work (using the Borland Pascal 7.0) solves the problem in several seconds on a IBM 486 PC.

A fundamental element of Eq. (21) is the regularization parameter α, which determines the degree of smoothing of the approximate solution. The second (stabilizing) term in Eq. (21) gives convexity and therefore ensures the very possibility of minimizing the functional and solving the problem. The solution obtained singles out from the set of functions satisfying the initial ill-posed equation the function that

is minimal in the sense of the W_2^1 norm, containing both the function itself and its derivative, of the stabilizing term, i.e., it realizes the condition of a certain compromise of minimality with respect to both the absolute value and the smoothness for the desired function. As shown in Ref. 22, the regularization parameter is uniquely related with the integral measure of the error in the data (number), which decreases monotonically as the error level decreases, but more slowly. As the error increases, the role of the second stabilizing term in Eq. (21) gradually decreases. The parameter α is found as the root of the one-dimensional nonlinear equation of the generalized residual

$$\rho(\alpha)=\|K_hR^\alpha-R_m^\delta\|_{L_2}^2-\delta^2=0, \tag{22}$$

where R^α is a function which minimizes the functional (21), i.e., the algebraic equation (22) is solved simultaneously with the functional equation (21).

The meaning of Eq. (22) is that the norm of the residual of the solution obtained should equal exactly the norm of the error, since there are no grounds for minimizing the deviation from the measurement data outside the error limits . The equation (22) contains the parameter δ of the effective error, which must be determined *a priori* on the basis of the specific conditions of the solution of the problem. This parameter must include all components of the measurement and interpretation errors. Specifically, δ must include both the random and systematic error δR_m of the measurements. The method can also include the error in the kernel $K(H_m,H)$, which includes the discretization error in the numerical solution and the possible uncertainty of its approximation by the corresponding functions (for example, expression (10), (12), and (16)),

$$\delta_h^2=\|K_hR-KR\|_{L_2}^2, \tag{23}$$

where K_h is the approximate kernel prescribed in solving Eq. (19).

The indicated errors can also cause the data vector to be incompatible with the equation being solved, since the smoothing action of the kernel (8) limits the class of possible realizations $R_m(H_m)$ and in the presence of a random error the function R_m^δ can drop out of the admissible class, i.e., it is impossible to obtain the measured distribution $R_m^\delta(H_m)$ for any distribution $R(H)$. A measure of incompatibility δ_μ, which is one of the parameters of the Tikhonov generalized-residual method, obviously, cannot exceed the total error of the kernel and the measurements:

$$\delta_\mu^2=\|K_hR-R_m^\delta\|_{L_2}^2\leqslant(\delta R_m+\delta_h)^2. \tag{24}$$

Thus, the quantity

$$\delta^2=(\delta R_m+\delta_h)^2+\delta_\mu^2, \tag{25}$$

which takes account of the measurement and discretization errors and other inaccuracies in the description of the kernel and also the measure, which depends on these factors, of the incompatibility of the equation with its right-hand side, is taken as the parameter of the effective error. In the Tikhonov method the values of the parameters appearing in Eq. (25) must represent the corresponding estimates with respect to the maximum in the class of possible realizations for the desired function. The value of the regularization parameter α and therefore the degree of smoothing of the solution are associated with the value of the parameter δ characterizing the effective error. The latter circumstance is a very important advantage of the method, since now the investigator's subjectivity is transferred from interpretation of the experimental data to estimation of the measurement error. Since an estimate of the error always contains an uncertainty, there exists a possibility of choosing a solution strategy. Thus, if the problem is to exclude nonexisting details in the solution, then it is better to adopt an estimate of the error with a certain excess, which, of course, can result in smoothing of some real details of the fine structure. If in solving the problem it is more important not to lose these details, then it is necessary to adopt the smallest error from the region of its actually possible values. In this case, however, spurious details, which do not exist in reality, can appear in the solution. A correct estimate of the error makes it possible to obtain an optimal solution by the Tikhonov method. After δ is determined, the procedure for obtaining the final result becomes formal.

A very important advantage of the generalized-residual method over other well-known methods is that as δ approaches zero in the integral metric, the approximate solution converges to the exact solution uniformly, i.e., in a metric where the norm is the maximum of the modulus, though, as a rule, in contrast to properly-posed problems the rate of convergence is not proportional to the decrease in δ but rather it is slower. Uniform convergence makes it possible to use the method of single numerical experiments with typical or extremal initial distributions to estimate the error, which is impossible to do in the cases of integral or rms convergence, since for convergence of these types there can exist regions of values of the argument where the desired function does not converge to the exact solution or even diverges.

We note that for some specific forms of the kernels $K(H_m,H)$ Eq. (8) possesses a solution in an explicit form. For example, for $K(H_m,H)$ of the simplest form (10), differentiating expression (8) with respect to the upper limit making the substitution $H_m\to H$ yields

$$R(H)=\langle R(H)\rangle-\frac{1}{2}H\frac{d\langle R(H)\rangle}{dH}. \tag{26}$$

Despite the simple form of the solution (26), we have obtained a well-known ill-posed problem of calculating a derivative from the experimental data.[22]

For $K(H_m,H)$ of the form (12) the equation (8) possesses a kernel with a weak singularity. This equation can be solved by the method of iterated kernels, similarly to the Abel equation,[25]

$$R(H)=\frac{1}{2H}\int_0^H\frac{d}{dH_m}[\langle R(H_m)\rangle H_m^2]\frac{dH_m}{\sqrt{H^2-H_m^2}}. \tag{27}$$

The solution (27) contains in the integrand a derivative of the initial data, which are obtained with some error. On account of the singularity of the kernel at the point $H_m=H$, the smoothing action of the integral does not eliminate

the error arising in the solution as a result of the errors in calculating the derivative. Indeed, in the numerical integration in Eq. (27) with step ΔH_m, the contribution of the integral near the singularity from $H-\Delta H_m$ to H is proportional to $\sqrt{\Delta H_m}$, while the error of the derivative in the integrand, in the case of an uncorrelated error, increases as $1/\Delta H_m$, i.e., the error in calculating the integral will grow without bound as $1/\sqrt{\Delta H_m}$, which attests to the ill-posed nature of the problem under study.

In summary, a systematic approach to solving both problems studied above must consist in solving the initial integral equation (8) on the basis of the rigorous methods of regularization, among which the Tikhonov generalized-residual method[22] is preferred for the class of problems under study, if the desired solutions belong to the class of continuous and differentiable functions.

RESULTS OF NUMERICAL MODELING

The specific nature of ill-posed problems lies in the fact that there does not exist a definite relation between the error of the right-hand side and the accuracy of the reconstruction, since the latter also depends strongly on the form of the desired function itself. For this reason, the reconstruction possibilities can be investigated only on the basis of a numerical experiment for typical distributions $R(H)$ and realizable levels of error.

Let us consider the solution of Eq. (8) for the example of $K(H_m,H)$ in the form (12), corresponding to cavity resonators. Since the problem is ill-posed, we shall use the Tikhonov generalized-residual method described above to solve it. Here it should be underscored that for a prescribed discretization the error in the solution of the problem, obtained using the explicit solution (27) or by inverting the numerical analog of the initial integral equation (8) (which consists of a linear system of algebraic equations with a triangular matrix of the kernel), will be a completely determined and finite quantity, which in principle can satisfy practical requirements. Thus, the correct procedure is to make an analysis by comparing the results of the direct-inversion and Tihkonov methods.

The numerical modeling of the process of reconstructing $R(H)$ was performed using the following closed scheme. The exact dependence $R_m(H_m)=\langle R(H_m)\rangle$, on which a random error, simulating the measurement error, was imposed at discrete points $m=1,2,\ldots,M$, was calculated for a specific initial function $R(H)$ using Eq. (8). A random number generator, which produced a sequence of M numbers with a normal distribution with a prescribed mean value ΔR_m and standard deviation σR_m, was used for the modeling. The ''measurement data'' obtained in this manner were used to solve the inverse problem by two methods and the reconstructed distributions $R(H)$ were compared with the initial distributions. The variance $\sigma_R^2(H)$ of the reconstruction error was determined by accumulating a statistical sample for a set of independent realizations of the random error.

For practical applications of the method under study it is necessary to prescribe the parameter δ, characterizing the effect of error and determined by the relation (25) in terms of the parameters δR_m, δ_μ, and δ_h, starting from the known conditions of the solution of the problem taking account of the circumstance that the experimental errors are random. Since the efficiency of the method is determined only by means of a numerical experiment and the optimality of the choice made can be checked, we shall take as an estimate of the error δR_M for Eq. (25) not the maximum but rather the average value of the integral in Eq. (20). Then we obtain

$$\delta R_m^2=\frac{1}{H_m^{\max}}\int_0^{H_m^{\max}}\langle[R_m(H_m)-R_m^{\delta}(H_m)]^2\rangle dH_m$$

$$=\frac{1}{H_m^{\max}}\int_0^{H_m^{\max}}[\sigma R_m^2(H_m)+\Delta R_m^2(H_m)]dH_m. \quad (28)$$

For constant parameters ΔR_m and σR_m

$$\delta R_m=\sqrt{\sigma R_m^2+\Delta R_m^2}. \quad (29)$$

One can see that for a zero systematic error $\delta R_m=\sigma R_m$. Conversely, for a zero random error $\delta R_m=\Delta R_m$. For a fixed discretization M, the component δ_h (23) of the total error was calculated by comparing with the result of exact integration of Eq. (8). Since the estimate (28) is, evidently, somewhat smaller than the error estimated according to the maximum of Eq. (25), it is natural to choose as the estimate of the measure δ_μ (22) of incompatibility its maximum possible value $\delta_\mu=\delta R_m+\delta_h$, which will bring the estimate of the total error δ closer to its true value. As a result, from Eq. (25) the following expression is obtained for the total error δ:

$$\delta=\sqrt{2}(\delta R_m+\delta_h), \quad (30)$$

where δR_m is determined from Eq. (28).

One can see from Eq. (8) that for a constant systematic error ΔR_m the error of the solution is $\delta R(H)=\Delta R_m$ (this follows from the linearity of Eq. (8) and the unit normalization of its kernel), which makes it possible to investigate the effect of only the random component of the error.

The equation (8) with the kernel (12) was solved numerically for a function $R(H)$ of the form (18) with the parameters $R_0=0.02$ mΩ, $n=4$, and $b=2.0\times10^{-7}$ Oe^{-4} in the range of magnetic fields $0\leqslant H\leqslant 100$ Oe. The indicated values of the parameters are close to the corresponding values obtained in Ref. 5 for frequency $f_0=1.5$ GHz at temperature $T=77$ K. The accuracies of the measurements were varied in the range $0.001\leqslant\delta R_m\leqslant 0.05$ mΩ.

The results of the numerical analysis with the reconstruction errors averaged over the realizations showed that the Tikhonov method makes it possible to obtain a solution of much higher quality with a fixed level of error and discretization. The functions $\sigma_R(H)$ for the Tikhonov method and the method of direct inversion are shown in Figs. 3a and b, respectively, in comparison with the difference $R(H)-R_m(H=H_m)$, which is the natural level of informativeness of the solution. One can see that the Tikhonov method gives good reconstruction for $\delta R_m\leqslant 0.02$ mΩ, while direct inversion gives under the same conditions appreciably worse results, limiting the admissible level of error in the measurements by the quantity $\delta R_m\leqslant 0.005$ mΩ close to the highest

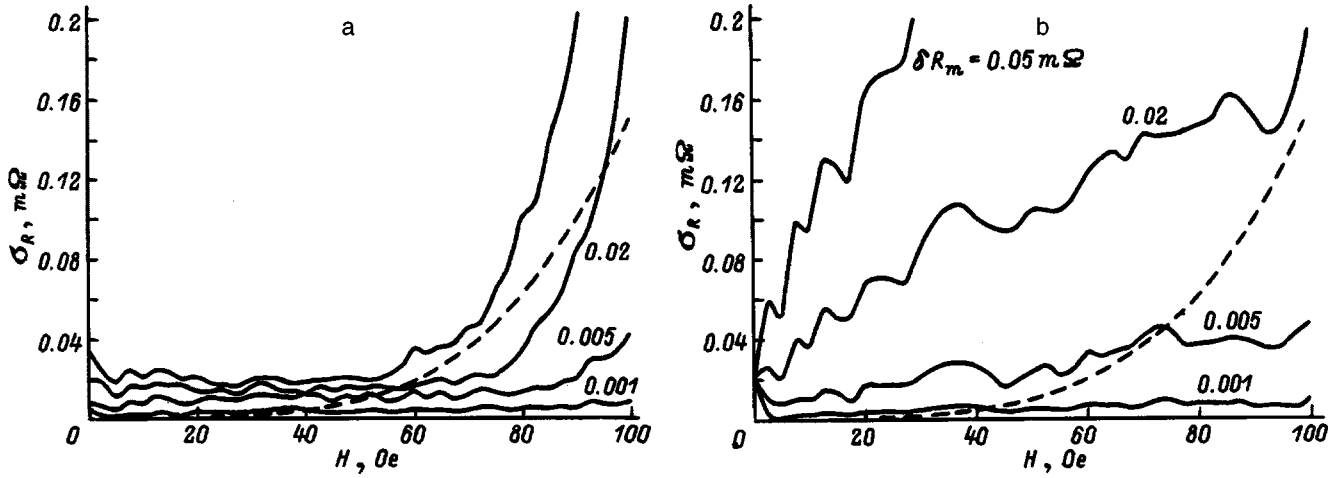

FIG. 3. Accuracy of solution of Eq. (8) for the Tikhonov method (a) and for the direct-inversion method (b) with different levels of measurement errors in comparison with the reconstructed deviation $R(H)-R_m(H=H_m)$ (dashed curve). The number of experimental points $M=20$.

accuracies that are attainable at present. It is also evident from the figures that for a function $R(H)$ of the type (18) the region of informative reconstruction lies in the range $H>50$ Oe.

By analogy to the measure of error δR_m (28), it is possible to introduce an integral reconstruction error δR that characterizes the quality of reconstruction on the entire interval H as a whole in the metric L_2

$$\delta R^2=\frac{1}{H^{\max}}\int_0^{H^{\max}}\langle[R_{\rm ret}(H)-R(H)]^2\rangle dH, \tag{31}$$

where $R(H)$ is the initial distribution and $R_{\rm ret}(H)$ is its reconstruction.

Figure 4 displays the relative reconstruction error $\delta R/\delta R_m$ versus the discretization parameter M (number of experimental points). Note the presence of an optimal value of M, for which the reconstruction error is minimal, and in addition this the minimum for the Tikhonov method corresponds to a much smaller error than the error for the direct-inversion method and it is reached at a smaller value of M. The existence of a minimum is explained as follows. As the value of M decreases, the discretization error increases, specifically, for the case presented in the figure the discretization error δ_h, which decreases as M increases, equals δR_m for $M=10$. The increase in the reconstruction error as M increases (which at first glance seems paradoxical) is due to the specific effect of the ill-posed nature of the inverse problem, which intensifies as M increases. The same effect, evidently, occurs in the calculation of the derivative from the experimental data.

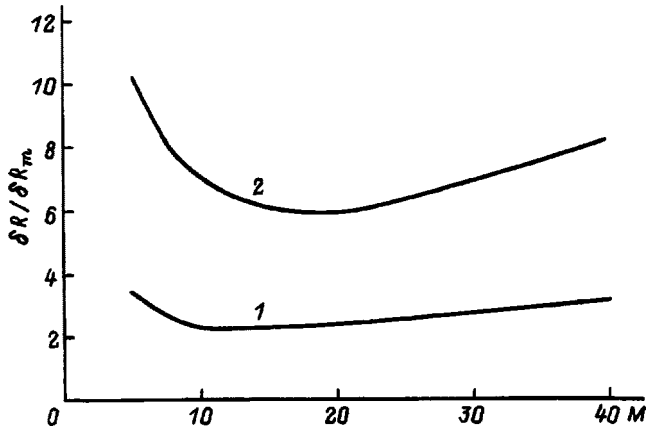

FIG. 4. Interval averaged relative reconstruction error as a function of the number of experimental points M. *1* — Tikhonov method, *2* — direct-inversion method.

A very important and convenient, for practical applications, feature of the Tikhonov method is the possibility of reconstructing $R(H)$ on the entire interval $0\leqslant H\leqslant H^{\max}$ using measurements of $R_m(H_m)$ in a narrower interval $H_m^{\min}\leqslant H_m\leqslant H_m^{\max}$ ($H_m^{\max}=H^{\max}$) or on a grid H_{mi} that is different from the grid H_i on which the initial distribution $R(H)$ is prescribed and reconstructed, and in addition this grid can be much more sparse. The dependence $R(H)$ can be effectively reconstructed outside the limits of the measurement interval of $R_m(H_m)$ in the direction of weaker fields. The corresponding inverse problem for the interval $0\leqslant H\leqslant H_m^{\min}$ is now described not by a Volterra equation but rather by a Fredholm equation of the first kind, which is a strongly ill-posed problem.[22] In this more general formulation the solution, in principle, cannot be obtained by direct inversion of the initial equation but only on the basis of the regularization method. Figure 5 presents an example of the reconstruction of the initial distribution of the type (18) with the above-indicated values of the parameters in the interval $0\leqslant H\leqslant 100$ Oe according to the ''experimental data'' obtained for $75\leqslant H_m\leqslant 100$ Oe (i.e., using only 1/4 of the reconstruction interval) and with the simulated error $\delta R_m=0.005$ mΩ. One can see that a reconstruction of quite good quality is obtained for $40\leqslant H_m\leqslant 100$ Oe, i.e., the Tikhonov method makes it possible to reconstruct $R(H)$ outside the interval of the measurements. In this region, where the problem is strongly ill-posed, the solution is much smoother.

In conclusion, we note that the approach developed above makes it possible to reconstruct arbitrary functions $R(H)$, including functions of a form much more complicated

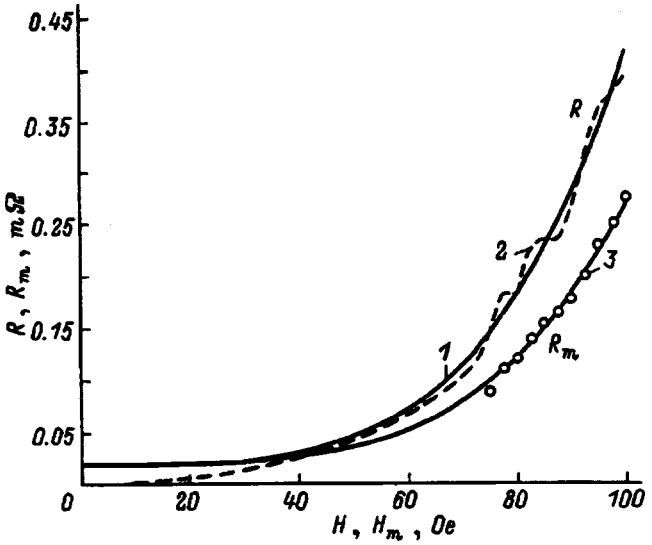

FIG. 5. Reconstruction of $R_s(H)$ from the experimental data given on 1/4 of the interval $0 \leqslant H \leqslant H_m$. *1* — Initial distribution, *2* — distribution reconstructed by the Tikhonov method, *3* — the function $\langle R_s(H_m)\rangle$. The circles mark the experimental data for $\sigma R_m = 0.005$ mΩ.

than Eq. (18). This fact is illustrated in Fig. 6, which shows an example of the reconstruction of a function $R(H)$ exhibiting saturation, which was observed in Ref. 14 (granular HTSC films, $f_0 = 18$ GHz, $T = 4$ K). The results are presented for one of the random realizations of the measurement error with $\delta R_m = 0.005$ mΩ and $M = 13$. We note that the error in the solution of the ill-posed problem considered here depends on the degree of complexity of the function $R(H)$ and should be investigated for each type of solution desired by numerical modeling similar to that described above.

CONCLUSIONS

In this paper, a method was developed for performing diagnostics of the nonlinear electromagnetic properties of HTSCs using microwave resonators. The traditional algorithms underestimate the nonlinearity of the HTSCs, making it possible to obtain an average surface resistance $\langle R_s(H)\rangle$.

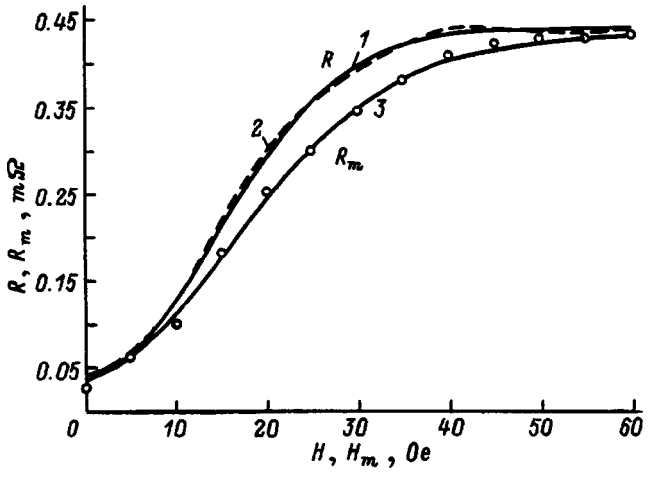

FIG. 6. $R_s(H)$ with saturation reconstructed by the Tikhonov method (the labels are the same as in Fig. 5).

The value of R_s can exceed $\langle R_s\rangle$ by a factor of ~ 1.5 in cavity and dielectric resonators, by a factor of $\sim 2-3$ in quasioptical resonators, and by a factor $\sim 5-7$ in microstrip resonators. The method proposed here is based on solving integral equations relating $\langle R_s(H)\rangle$ with $R_s(H)$, which makes it possible to take account of the nonuniformity of the structure of the field H in microwave resonators. The most suitable algorithms for solving the equations obtained employ a regularization technique developed in the theory of ill-posed inverse problems. Such methods give high-quality reconstructions of the functions $R_s(H)$ with currently attainable accuracies of radiophysical measurements ($\delta R \propto 5-10$ $\mu\Omega$) and minimal volumes of experimental data (the number of points $M \propto 10-15$).

We thank A. A. Zharov for helpful remarks.

This work was supported by the Scientific-Technical Program ''High-Temperature Superconductors,'' Project No. 96130, and the Russian Fund for Fundamental Research, Project No. 96-02-16997.

[1] N. Newman and W. G. Lyons, J. Supercond. **6**, 119 (1993).
[2] H. J. Chaloupka, M. A. Hein, and G. Muller, ''High-Tc microwave superconductors and applications,'' Proc. Soc. Photo-Opt. Instrum. Eng. (SPIE) **2156**, 36 (1994).
[3] D. E. Oates, A. C. Anderson, D. M. Sheen *et al.*, IEEE Trans. Microwave Theory Tech. **MTT-39**, 1522 (1991).
[4] A. M. Portis, J. Supercond. **5**, 319 (1992).
[5] P. P. Nguyen, D. E. Oates, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **48**, 6400 (1993).
[6] M. A. Golosovsky, H. J. Snortland, and M. R. Beasley, Phys. Rev. B **51**, 6462 (1950).
[7] B. L. Willemsen, J. S. Derov, J. H. Silva, and S. Sridhar, IEEE Trans. Appl. Supercond. **5**, 1753 (1995).
[8] J. Halbritter, J. Appl. Phys. **71**, 339 (1992).
[9] A. Porch and A. M. Portis, Physica C **235-240**, 3381 (1994).
[10] M. Hein, W. Diete, M. Getta *et al.*, IEEE Trans. Appl. Supercond. **AS-7**, 1264 (1997).
[11] J. Wosik, L. M. Xie, K. Nesteruk *et al.*, Preprint N 97 : 005, Texas Center Supercond., Univ. of Houston, Texas 77204–5932 (1997).
[12] D. E. Oates, A. C. Anderson, and P. M. Mankiewich, J. Supercond. **3**, 251 (1990).
[13] S. J. Hedges, M. J. Adams, B. F. Nicholson, and N. G. Chew, Electron. Lett. . **26**, 977 (1990).
[14] D. W. Cooke, E. R. Gray, and R. N. Arendt, J. Supercond. **3**, 261 (1990).
[15] W. Diete, B. Aschermann, H. Chaloupka *et al.*, in *Applied Superconductivity*, Vol. 148, edited by D. Dew-Hughes, IOP Publishing, Bristol, 1995, pp. 1107–1110.
[16] Z.-Y. Shen, C. Wilker, and P. Pang, IEEE Trans. Microwave Theory Tech. **MTT-40**, 2424 (1992).
[17] T. E. Harrington, J. Wosik, and S. A. Long, Preprint N 96:106, Texas Center Supercond. Univ. of Houston, Texas 77204–5932 (1996).
[18] D. W. Cooke, P. N. Arendt, E. R. Gray, and A. M. Portis, IEEE Trans. Microwave Theory Tech. **MTT-39**, 1539 (1991).
[19] D. M. Sheen, S. M. Ali, D. E. Oates *et al.*, IEEE Trans. Appl. Supercond. **AS-1**, 108 (1991).
[20] V. I. Vol'man, *Handbook on Calculations and Design of Microwave Strip Devices* [in Russian], Radio i Svyaz', Moscow (1982), 328 pp.
[21] O. G. Vendik and A. Yu. Popov, Zh. Tekh. Fiz. **63**, 1 (1993) [Sov. Phys. Tech. Phys. **38**, 535 (1993)].
[22] A. N. Tikhonov, A. V. Goncharskiĭ, V. V. Stepanov, and A. G. Yagola, *Regularizing Algorithms and A Priori Information* [in Russian], Nauka, Moscow (1983), 200 pp.
[23] A. N. Reznik, IEEE Trans. Appl. Supercond. **AS-7**, 1474 (1997).
[24] F. P. Vasil'ev, *Numerical Methods for Solving Extremal Problems* [in Russian], Nauka, Moscow (1981), 520 pp.
[25] M. L. Krasnov, *Integral Equations* [in Russian], Nauka, Moscow (1975), 304 pp.

Translated by M. E. Alferieff

Scintillation characteristics of thin-film calcium iodide crystal x-ray detectors

S. S. Novosad

I. Franko L'vov State University, 290005 L'vov, Ukraine
(Submitted April 29, 1997; resubmitted August 26, 1997)
Zh. Tekh. Fiz. **68**, 87–90 (August 1998)

A container design for fabricating low-energy x-ray scintillation detectors is proposed. CaI_2 and CaI_2:Eu crystal detectors are fabricated and their characteristics are investigated. It is shown that on account of their layered structure, perfect cleavage, and high light output, calcium iodide scintillators can be used to fabricate thin-film detectors for long-wavelength x rays.
 [S1063-7842(98)01408-1]

Special applications of NaI and CsI scintillators are soft γ- and x-ray spectrometry and α-particle detection in the presence of a γ-ray background.[1–5] One of the basic problems that must be solved to detect ionizing radiations with low penetrating power is to effect selective detection of this radiation against a background consisting of radiation with high penetrating power. In many cases thin-film scintillation detectors are used to solve this problem.[2,5]

The use of thin-film detectors for low-energy ionizing radiations also makes it possible to decrease the scintillation light losses at the side surface and on crystal defects and to decrease the influence of reabsorption.

The fabrication of detectors for low-energy ionizing radiations runs into the difficulty of obtaining and encapsulating wafers of crystalline scintillators. In the case at hand the quality of the crystal surface plays an especially important role because soft ionizing radiation is absorbed in a layer near the surface.

The production of thin-film scintillation detectors with a large working area on the basis of NaI:Tl, CsI:Tl, and CsI:Na single crystals by cleaving, grinding, polishing, etching, or partial dissolution is a difficult technological process.[2]

CaI_2 and CaI_2:Eu crystals have a high light output and a better energy resolution than NaI:Tl scintillators.[3,4,6–9] They concede essentially nothing to CsI:Na with respect to the de-excitation time and effective number.[10] In this connection it is of practical interest to investigate the production of thin-film x-ray detectors based on calcium iodide single crystals and to investigate their properties.

The first attempt to use CaI_2 crystals to detect soft x rays was made in Ref. 3. It was shown there on the basis of calculations that 5.5–20 keV x rays, which are most widely used in x-ray crystallographic analysis, are actually completely absorbed by a 0.1 mm thick crystal. However, in Ref. 3 the calcium iodide crystal x-ray detectors were fabricated using 1–1.5 mm thick single-crystal scintillator wafers with an area of 6×16 mm.

Layered calcium iodide crystals have a hexagonal structure with pronounced cleavage. They are plastic and have low hardness. Such properties of CaI_2 make it possible to fabricate from a crystal very thin wafers (up to 0.05 mm thick) by splitting off and cutting into any desired shape. However, thin wafers can be fabricated from a large, strongly hygroscopic, layered scintillator only in a ''glued-on'' state (for example, glued to the exit window of a detector), because otherwise the single-crystal wafers deform during the fabrication and packaging process. For this reason, in the past[3,4,7,8] CaI_2-based x-ray detectors were fabricated with thickness greater than 1 mm. Moreover, the standard container structure employed (Fig. 1) is unsuitable for packaging thin-film CaI_2 wafers with a large working area because it does not allow free access to the scintillator.

To fabricate CaI_2-based x-ray detectors, we improved the technology of growing single crystals,[7,8] the fabrication of crystal wafers, and the structural implementation of the containers themselves.[11] Certain results of the investigations performed in this direction are also reported in this paper.

In the present work we employed containers with a different design to package thin scintillator wafers (Fig. 2). A characteristic feature of this design is that the optical exit window *2*, fabricated in the form of a truncated glass cone, was placed in a recess made in the inner part of the case *1*. The proposed container design, just as the one presented in Ref. 11, makes it possible to change the sequence of operations in the fabrication and packaging of thin scintillators, since the final adjustment of the crystal wafer *4* over the height can be made by splitting off (using a razor blade edge) in a state glued to (with optical glue *3*) the exit window *2*; separate metal templates can be used. The scintillator wafer,

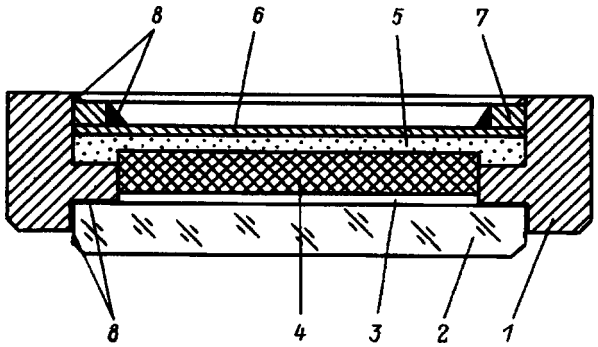

FIG. 1. Design of a scintillation x-ray detector. *1* — Duralumin case, *2* — K-8 glass exit window, *3* — OK-50 optical glue, *4* — *scintillator*, *5* — MgO reflector, *6* — entrance window (beryllium, 0.2 mm thick), *7* — clamping ring, *8* — sealing material (epoxy glue).

prepared in this manner, together with the exit window are placed into a recess in the case, and are mechanically strongly and hermetically secured, using sealing material *9*, for example, epoxy glue. Next, a fluoroplastic separating ring *5* and a reflector *6*, consisting of a metallized Mylar film, are placed into the case.

For additional sealing, the Mylar film reflector can be placed directly on the beryllium entrance window *7* using epoxy glue *9*. The last parts of the container are secured in the case with a clamping device *8* and simultaneously the scintillator is sealed with the gluing material *9*.

The proposed container is especially advantageous for packaging thin-film scintillators based on the hygroscopic compound calcium iodide, since the degree to which atmospheric moisture can influence the characteristics of the detector can be decreased at the time the scintillator wafers are prepared and packaged. This is accomplished because the worked surface of the scintillator is in contact with the atmosphere where the packaging is performed for a shorter period of time, i.e., it is possible to obtain detectors with a thinner ''dead layer'' of the scintillator. This container design can also be used successfully for packaging deposited scintillator layers.

CaI_2 and CaI_2:Eu crystals, grown by the Stockbarger method,[7,8] with high spectrometric characteristics were used to fabricate thin-film x-ray detectors. The x-ray detectors, fabricated from these crystals, with 2.0–2.5 mm thick and 16 mm in diameter scintillation wafers with the standard design (Fig. 1) had a 1.6–1.9 times higher light output when detecting ^{137}Cs γ rays than NaI:Tl x-ray detectors, and they were characterized by an energy resolution of 4.5–5.5%. At room temperature the crystals obtained were also characterized by a 1.8–2.2 times higher stationary x-ray luminescence output than a NaI:Tl scintillator.

The results of measurements of the luminescence properties with x-ray excitation and the scintillation characteristics with γ-ray excitation of our CaI_2 and CaI_2:Eu crystals agree with the data presented in Ref. 6.

Single crystal wafers (0.2–0.8 mm thick, 16 and 20 mm in diameter) of the CaI_2 and CaI_2:Eu scintillators were packaged, by the method described in Ref. 11, in containers with

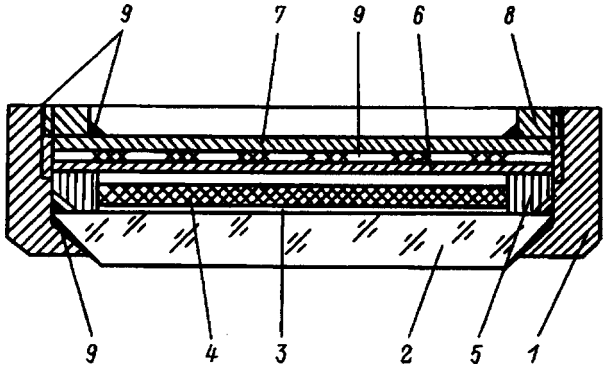

FIG. 2. Construction of a scintillation detector for long-wavelength x rays. *1* — Duralumin case, *2* — K-8 glass exit window, *3* — OK-50 optical glue, *4* — scintillator, *5* — separating ring (fluoroplastic lateral reflector), *6* — reflector consisting of an aluminized Mylar film, *7* — 0.2 mm thick beryllium entrance window, *8* — clamping ring, *9* — sealing material (epoxy glue).

the above-described design (Fig. 2) in hermetically sealed 7B-0S type boxes in a dry-air atmosphere. Measurements, performed by the method described in Ref. 12, of the scintillation properties of the thin x-ray detectors obtained showed that calcium iodide crystal the detectors had a 1.5–1.75 times higher light output when detecting Cu $K\alpha$ x rays ($E_{ex}=8.05$ keV) than NaI:Tl detectors, and they were characterized by an energy resolution of 39–45%. The energy resolution of the best calcium iodide x-ray detectors when detecting Mo $K\alpha$ radiation ($E_{ex}=17.4$ keV) reached 26%.

The curves of the amplitude distribution of the pulses of CaI_2 and CaI_2:Eu detectors were gaussian in the region of the photopeak, irrespective of the photon energy of the detected x rays.

In the course of the basic technological and investigatory work it was established that calcium iodide crystals doped with lead, indium, gadolinium, and sodium are also effective scintillators. The light output of $CaI_2:GdCl_3$ and $CaI_2:InCl_3$ crystals with the optimal impurity concentration is virtually identical to that of the $CaI_2:EuCl_3$ scintillator,[12] and the $CaI_2:PbI_2$ and CaI_2:NaI crystals are characterized by a

TABLE I. Luminescence and scintillation characteristics of x-ray detectors ($E_{ex}=17.4$ keV, FÉU-35A).

Crystal No.	Crystal	Luminescence band maximum, nm	Light output arb. units	Energy resolution, %
1	CaI_2	410–415	100	30
2	CaI_2: 0.005 mol % $EuCl_3$	465–470	105	29
3	CaI_2: 0.01 mol % $EuCl_3$	465–470	105	30
4	CaI_2: 0.005 mol % PbI_2	390–410	95	32
5	CaI_2: 0.01 mol % PbI_2	390–415	95	33
6	CaI_2: 0.50 mol % PbI_2	430–440	70	38
7	CaI_2: 0.01 mol % $InCl_3$	420–430	95	33
8	CaI_2: 0.10 mol % $InCl_3$	420–430	105	30
9	CaI_2: 1.00 mol % $InCl_3$	425–435	75	36
10	CaI_2: 0.02 mol % $GdCl_3$	420–430	95	32
11	CaI_2: 0.05 mol % $GdCl_3$	420–440	105	30
12	CaI_2: 0.50 mol % $GdCl_3$	465–470	73	34
13	CaI_2: 0.04 mol % NaI	420–430	90	32
14	CaI_2: 0.20 mol % NaI	420–430	85	33

somewhat lower light output than the highly efficient calcium iodide scintillators.

The results of measurements of the luminescence and scintillation characteristics of x-ray detectors, based on pure and activated calcium iodide crystals, with 1.5–2.5 mm thick and 16 mm in diameter scintillation wafers, packaged in containers with the design shown in Fig. 1, are summarized in Table I. The table gives the average light output and energy resolution, obtained while detecting Mo $K\alpha$ x rays with at least three detectors fabricated from the same single crystal. It follows from the data in the table that besides CaI_2 and CaI_2:Eu crystals thin x-ray detectors can also be fabricated using CaI_2:Gd, CaI_2:Pb, and CaI_2:Na crystals. The spectral composition of the x-ray luminescence of these scintillators matches well with the spectral sensitivity of the photomultipliers most widely used in scintillation technology.

Analysis of the published data and the results of investigations of the optical-luminescence properties of CaI_2:$EuCl_3$ and CaI_2:$GdCl_3$ crystals established that the europium and gadolinium ions are incorporated into the calcium iodide lattice in a divalent state. The activator bands observed in the absorption spectra of these crystals are associated with $f-d$ transitions in the Eu^{2+} and Gd^{2+} ions. The luminescence of europium- and gadolinium-based calcium iodide crystals with a maximum in the range 465–470 nm was attributed to a complex of centers formed by the rare-earth impurity and interstitial calcium ions. Investigations of the curves and spectra of thermally-stimulated luminescence, the temperature dependence of the intensity and spectral composition of the x-ray luminescence, and the spectra for producing and stimulating the photostimulated luminescence of these crystals established that the Gd^{2+} and Eu^{2+} ions in CaI_2 are effective electron and hole trapping centers. The electron–hole mechanism of luminescence dominates in the CaI_2:Gd and CaI_2:Eu scintillators at low temperatures. In heavily europium-doped calcium iodide crystals the radiative-reabsorption mechanism of energy transfer to the luminescence centers is also observed with x-ray excitation. The decrease in the light output and degradation of the energy resolution in heavily doped CaI_2:Gd and CaI_2:Eu crystals are due mainly to the effect of reabsorption and the persistent losses of excitation energy on trapping centers.

The wide complex bands observed in the x-ray luminescence spectra of CaI_2:Pb, CaI_2:In, and CaI_2:Na crystals at room temperature are due to centers of nonactivated origin. In heavily doped CaI_2:Pb crystals there is a substantial accumulation of the light sum both on trapping centers characteristic for CaI_2 and trapping centers associated with Pb^{2+} ions. Indium- and sodium-doped calcium iodide crystals are characterized by weak thermally stimulated luminescence. The decrease in the light output in these scintillators with increasing impurity content is caused mainly by losses of excitation energy and emission on defects associated with degradation of the structure of the crystals.

It should be noted that the container designs proposed in Ref. 11 and in the present work can also be used for packaging thick (2–10 mm) layered scintillators, since they permit solving successfully the question of removing excess optical glue and packaging more perfect crystal wafers.

In summary, the results obtained show that on account of their layered structure with perfect cleavage and high light output calcium iodide scintillators can be used to fabricate thin detectors for long-wavelength x rays.

I wish to thank M. R. Panasyuk for measuring the scintillation properties of the thin x-ray detectors.

[1] Yu. A. Tsirlin, A. R. Daĭch, and A. M. Radivanyuk, *Scintillation Detection Units* [in Russian], Atomizdat, Moscow (1973), 264 pp.
[2] A. B. Lyskovich, *Author's Abstract of Doctoral Dissertation* [in Russian], L'vov (1971), 30 pp.
[3] V. R. Al'perovich, I. I. Komyak, A. B. Lyskovich, and G. M. Pentsak, Prib. Tekh. Éksp., No. 4, 223 (1970).
[4] V. R. Al'perovich, I. I. Komyak, A. B. Lyskovich, and G. M. Pentsak, Opt. Spektrosk. **33**, 499 (1972) [Opt. Spectrosc. **33**, 269 (1972)].
[5] É. F. Chaĭovskiĭ, Yu. A. Tsirlin, Yu. T. Vidaĭ, and G. M. Ivanova, in *Single Crystals and Technology* [in Russian], All-Union Scientific-Research Institute of Single Crystals, Kharkov, 1973, No. 2(9), pp. 141–145.
[6] R. Hofstadter, E. W. O'Dell, and C. T. Schmidt, IEEE Trans. Nucl. Sci. **NS-11**, N3, 12 (1964).
[7] A. B. Lyskovich, S. S. Novosad, and M. R. Panasyuk, in *Abstracts of the 6th All-Union Conference on the Synthesis and Investigation of Scintillators* [in Russian], All-Union Scientific-Research Institute of Single Crystals, Kharkov, 1971, pp. 18–21.
[8] A. B. Lyskovich, S. S. Novosad, M. R. Panasyuk, and T. I. Triska, in *Single-Crystal Scintillators and Organic Luminophors* [in Russian], All-Union Scientific-Research Institute of Single Crystals, Cherkassy, 1972, pp. 41–43.
[9] S. S. Novosad, Vestnik L'vov. Universiteta, Ser. Fiz., No. 11, Éksp. Teor. Fiz., Vishcha Shkola, L'vov, 1976, pp. 44–57.
[10] L. V. Viktorov, A. V. Kruzhanov, V. K. Petrov, V. M. Skorikov, and B. V. Shul'gin, Deposited in VINITI, Moscow, 1990.
[11] S. S. Novosad and N. K. Gloskovskaya, ''Method of fabricating a scintillation detector based on layered scintillators'' [in Russian], Inventor's Certificate No. 784535 (USSR).
[12] A. B. Lyskovich, S. S. Novosad, T. I. Triska *et al.*, in *Proceedings of the 7th All-Union Conference on Prospects for the Development of a Technology of Methods of Monitoring and Producing Scintillators and Scintillation Devices During the 10th Five-Year Plan* [in Russian], All-Union Scientific-Research Institute of Single Crystals, Kharkov, 1976, Part I, pp. 223–228 (Manuscript No. 1175/77 DSP, deposited in the Reference–Information Foundation of the Division of the Scientific-Research Institute of Technical-Economic Research of the Soviet Ministry of the Chemical Industry).

Translated by M. E. Alferieff

Electromagnetic wave generation in a cylindrical resonator by electrons orbiting in a radial electrostatic field

V. V. Dolgopolov, Yu. V. Kirichenko, Yu. F. Lonin, and I. F. Kharchenko

National Science Center, Kharkov Physicotechnical Institute, Ukrainian Academy of Sciences, 310108 Kharkov, Ukraine

(Submitted August 12, 1996; resubmitted August 4, 1997)
Zh. Tekh. Fiz. **68**, 91–94 (August 1998)

The mechanisms leading to the generation of microwaves by electrons orbiting in a radial electrostatic field produced by a positively charged filament on the axis of a cylindrical resonator are investigated theoretically. The dispersion relations describing the interaction of the waves with the electrons are obtained. It is shown that the generation of electromagnetic fields is possible on account of both Cherenkov and plasma resonances. The frequencies and growth rates of waves under Cherenkov resonance conditions and also plasma resonance conditions in uniform and nonuniform electron layers are found. The advantages and disadvantages of different generation mechanisms are determined. © *1998 American Institute of Physics.* [S1063-7842(98)01508-6]

INTRODUCTION

Investigations of the dynamics of charged-particle motion in a cylindrically symmetric electrostatic field of a thin conducting filament are of interest because such systems have an entire series of practical applications: Geiger–Müller counters, electrostatic filters for contaminated gases, ion-plasma pumps, gas-discharge high-vacuum meters, and others. Systems with centrifugal-electrostatic focusing of the electron stream (spiratrons) were studied in Ref. 1. In the last few years a novel, unconventional millimeter-wave generator, which the authors of the device term an orbitron, was proposed.[2] This generator has the advantage that there is no external magnetic field. It was shown experimentally that in an orbitron it is possible to obtain radiation power of the order of 10 W at 40 GHz with a voltage of 2 kV on the filament. The authors assert that further elaboration of this device will make it possible to obtain submillimeter waves right down to 0.1 mm.

The orbitron generator is a cylindrical resonator. The resonator axis is also the axis of a thin metal rod (filament) to which a positive voltage is applied. A cylindrically symmetric layer of electrons orbits around the rod. The electrons are held in a circular orbit by the electrostatic field of the rod, which compensates the centrifugal force acting on the electrons. Under certain conditions this system is unstable with respect to disturbances of the electron orbits and densities and the appearance of electromagnetic field disturbances associated with these disturbances. The radius of the cylindrical layer of electrons decreases with radiation generation and emission from the resonator.

A theoretical description of the instability of the electron layer in an orbitron is given in Ref. 2. However, it is unsatisfactory in many respects. This was the reason the present theoretical investigation was undertaken.

BASIC EQUATIONS

Let the resonator be a cavity $a \leqslant r \leqslant b$ which is unbounded along the z axis (a cylindrical coordinate system r, φ, z is used) and bounded at the surfaces $r=a$ and $r=b$ ($a \ll b$) by an ideally conducting metal. The metal rod ($r<a$) possesses a positive charge Q per unit length, corresponding to the potential difference $\Psi = 2Q\ln(b/a)$ between the rod and the outer shell of the resonator. The density $n_{e0}(r)$ of the cylindrically symmetric electron layer is different from zero only in a narrow region between the surfaces $r=r_-$ and $r=r_+$

$$n_{e0}(r) \neq 0 \quad \text{for} \quad r_- < r < r_+,$$

$$n_{e0}(r) = 0 \quad \text{for} \quad r \leqslant r_- \quad \text{and} \quad r \geqslant r_+,$$

$$\delta r = r_+ - r_- \ll r_-, \quad \frac{\partial n_{e0}}{\partial \varphi} = 0,$$

$$\frac{\partial n_{e0}}{\partial z} = 0, \quad a < r_- < r_+ < b. \tag{1}$$

Assuming the field and the disturbances of the density and velocity of the electrons to be independent of the axial coordinate z, the linearized nonrelativistic equations of motion of the electrons can be written in the form

$$\frac{\partial V_r}{\partial t} + \frac{V_0}{r}\frac{\partial V_r}{\partial \varphi} - 2\frac{V_0}{r}V_\varphi = -\frac{e}{m_e}E_r, \tag{2}$$

$$\frac{\partial V_\varphi}{\partial t} + \frac{V_0}{r}\frac{\partial V_\varphi}{\partial \varphi} + \frac{V_0}{r}V_r = -\frac{e}{m_e}E_\varphi, \tag{3}$$

where $V_0 = (2eQ/m_e)^{1/2}$, $\mathbf{V}_0 = \mathbf{e}_\varphi V_0$ is the unperturbed velocity of the electrons, V_r and V_φ are, respectively, the radial and azimuthal components of the perturbation of the electron velocity, $\mathbf{e}_\varphi$ is a unit vector in the direction of the azimuth φ, E_r and E_φ are the components of the electric field of the wave, and $-e$ and m_e are the electron charge and mass.

The components of the current density **j** of the disturbances can be expressed simply in terms of the perturbations of the velocity **V** and density n_e of the electrons as

$$j_r=-en_{e0}V_r, \quad j_\varphi=-e(n_{e0}V_\varphi+V_0n_e). \tag{4}$$

We shall determine the dependence of the variable quantities on the azimuthal coordinate φ and the time t by the factor $\exp\{i(m\varphi-\omega t)\}$, where m is an integer. For definiteness, we assume $\mathrm{Re}(\omega)>0$, $m>0$. Then the following equations for the components of the electromagnetic field in the region $r_-\leqslant r\leqslant r_+$ can be obtained from the continuity equation, Maxwell's equations, and Eqs. (2)–(4):

$$E_r=\frac{i}{W_1}\left(\frac{mc^2}{r^2}\frac{\partial}{\partial r}(rE_\varphi)+\frac{2\omega V_0\Omega^2}{rW}E_\varphi\right), \tag{5}$$

$$H_z=-i\frac{c}{W_1}\left(\left(\omega-\frac{\Omega^2}{W}\omega_m\right)\frac{1}{r}\frac{\partial}{\partial r}(rE_\varphi)+\frac{2mV_0\Omega^2}{r^2W}E_\varphi\right), \tag{6}$$

$$\frac{\partial}{\partial r}\left\{\frac{1}{W_1}\left(\omega_m\left(1-\frac{\Omega^2}{W}\right)\frac{1}{r}\frac{\partial}{\partial r}(rE_\varphi)+\frac{2V_0\Omega^2}{rW}\times\left(\frac{m}{r}-\frac{\omega V_0}{c^2}\right)E_\varphi\right)\right\}=\frac{mV_0}{W_1r^3}\left(1+\frac{\Omega^2}{W}\right)\frac{\partial}{\partial r}(rE_\varphi)+\frac{1}{c^2}\times\left\{\frac{\Omega^2}{W}\left(\omega_m+\frac{2\omega V_0^2}{W_1r^2}\left(1+\frac{\Omega^2}{W}\right)\right)-\omega_m\right\}E_\varphi, \tag{7}$$

where $\omega_m=\omega-mV_0/r$, $W=\omega_m^2-2V_0^2/r^2$, $W_1=\omega(\omega-\omega_m\Omega^2/W)-m^2c^2/r^2$, $\Omega(r)=\{4\pi e^2n_{e0}/m_e\}^{1/2}$ is the electron plasma frequency, H_z is the magnetic field of the wave, and c is the velocity of light. The equations (5)–(7) hold when the electrostatic field of the electron layer is weak compared with the electrostatic field of the rod. This condition has the form

$$\int_{r_-}^{r_+}r\Omega^2dr\ll V_0^2. \tag{8}$$

It is impossible to solve Eq. (7) analytically in the general case. For this reason, it was solved approximately up to terms of zeroth and first orders, inclusive, in the small parameter $\delta r/r_-$ (a similar method was used in Refs. 3–5). Matching the expressions obtained in this manner for the fields at the boundaries of the layer with the expressions for the fields in vacuum and assuming the field component E_φ at the boundaries with the metal to be zero, we obtain the dispersion relation. On account of the complexity of this equation, we shall analyze it only in a case of practical importance where

$$\left|\frac{\omega}{c}r_-\right|\ll 1, \quad \left|\frac{\omega}{c}b\right|\gg m^2/2, \tag{9}$$

and under conditions such that wave generation is possible. These are the conditions for Cherenkov and plasma resonances.

CHERENKOV RESONANCE

Cherenkov resonance arises when the angular phase velocity of the wave is close to the angular velocity of the electrons $\omega_m(r_-)\simeq 0$. Then the dispersion relation can be put into the form

$$\omega_m^3+(i\nu-\Delta)\omega_m^2-\Delta_2^2\omega_m-\Delta_3^3=0, \tag{10}$$

where

$$\Delta=\omega^{(0)}-\frac{mV_0}{r_-}+\Delta_\nu, \tag{11}$$

$$\Delta_2^2=\frac{\pi x_1^{2m}(\eta^{4m}-1)\omega^{(0)}cr_-}{2^{2m-1}m!(m-1)!\,\eta^{2m}b}\int_{r_-}^{r_+}dr\,\frac{\Omega^2}{2V_0^2+r^2\Omega^2}, \tag{12}$$

$$\Delta_3^3=\frac{\pi x_1^{2m}(\eta^{2m}-1)^2c}{2^{2m+1}[(m-1)!]^2\eta^{2m}br_-}\int_{r_-}^{r_+}dr\,\frac{r^2\Omega^2-2V_0^2}{2V_0^2+r^2\Omega^2}\Omega^2, \tag{13}$$

$$\eta=\frac{r_-}{a}>1, \quad x_1=\frac{\omega^{(0)}}{c}a\omega^{(0)}\simeq\frac{c}{b}\left(n\pi+\frac{\pi}{2}m+\frac{\pi}{4}\right)$$

is the characteristic frequency of the resonator in the absence of the electron layer and neglecting losses; n is an integer; here and below, the term $i\nu$ takes into account phenomenologically the losses due to radiation emission from the resonator and the absorption of energy by the resonator walls, while the term $-\Delta_\nu$ takes account of the frequency shift associated with these losses.

The equation (10) is a cubic equation for ω_m. The imaginary part of ω_m equals the imaginary part of ω and therefore it is the growth rate (damping rate) of the wave. The last term on the left-hand side of Eq. (10), as follows from the relation (13), can vanish. In this case the cubic equation transforms into a quadratic equation, all of whose solutions correspond to damped oscillations.

If $|i\nu-\Delta|\ll|\Delta_3|$, oscillations growing in time arise if

$$|\Delta_3^3|>2|\Delta_2^3|/3^{3/2}. \tag{14}$$

For $\Delta_3^3\neq 0$ a more stringent condition than the condition (14) is satisfied:

$$|\Delta_3^3|\gg|\Delta_2^3|. \tag{15}$$

Then the dispersion relation assumes the form

$$\omega_m^3\simeq\Delta_3^3. \tag{16}$$

One of the three roots of Eq. (16) always corresponds to oscillations growing in time. The growth rate is maximum in this case. If $\Delta_3^3>0$, which corresponds to high electron densities and large r_-, then the phase velocity of the growing wave near r_- is less than the velocity of the electrons, whereas if $\Delta_3^3<0$, which corresponds to high electron velocities, then the phase velocity of the growing wave near r_- is greater than the velocity of the electrons.

For large losses or detuning from resonance, such that the conditions

$$|\Delta_2^2|\ll|\Delta_3^{3/2}(i\nu-\Delta)^{1/2}|, \quad |\Delta_3|\ll|i\nu-\Delta|, \tag{17}$$

are satisfied, the solution of Eq. (10) has the form

$$\omega_m \simeq \pm\left(\frac{\Delta_3^3}{i\nu-\Delta}\right)^{1/2}. \tag{18}$$

One root of Eq. (18) corresponds to oscillations growing in time. One can see that as the losses due to radiation emission from the resonator or to dissipation increase, the growth rate of the wave does not increase, as happens in the results presented in Ref. 2, but rather decreases. The frequencies of waves amplified under Cherenkov resonance conditions $\omega \simeq mV_0/r_-$ grow as the radius r_- of the electron layer decreases.

PLASMA RESONANCE IN A UNIFORM ELECTRON LAYER

Plasma resonance occurs when the frequency of the electromagnetic field is close to the frequency of the characteristic local longitudinal oscillations of the electrons in the laboratory coordinate system. This frequency ω_p is determined by the equation

$$W-\Omega^2=0 \tag{19}$$

and has the form

$$\omega_p(r)=\frac{mV_0}{r}\pm\left(2\frac{V_0^2}{r^2}+\Omega^2(r)\right)^{1/2}. \tag{20}$$

Let us consider the case where the electron density grows rapidly from zero up to its maximum value near $r=r_-$, remains practically constant in the region $r_-<r<r_+$, and then decays rapidly to zero near $r=r_+$, i.e., $\partial\Omega^2/\partial r\simeq 0$ for $r_-<r<r_+$. Then the solution of the dispersion relation assumes the form

$$\delta\omega=\frac{1}{2}\{\Delta_p-i\nu\pm((\Delta_p-i\nu)^2-4q)^{1/2}\}, \tag{21}$$

where

$$\delta\omega=\omega-\omega_p, \quad \Delta_p\simeq\omega^{(0)}-\omega_p+\Delta_\nu, \tag{22}$$

$$q=-\frac{\pi x_1^{2m-1}(\eta^{2m}-1)^2 W\omega\delta r}{2^{2m}[(m-1)!]^2\eta^{2m+1}b\omega_m}\left(\frac{V_0}{r_-\omega_m}-\frac{1}{2}\frac{\eta^{2m}+1}{\eta^{2m}-1}\right)^2. \tag{23}$$

In the latter relations the r-dependence of ω_p can be neglected. The imaginary part of $\delta\omega$ is the growth rate (damping rate) of the wave. As follows from expression (21), instability occurs when $q>0$. This inequality, according to expression (23), corresponds to the condition $\omega/\omega_m<0$, i.e., oscillations whose phase velocity near the electron layer is less than the velocity of the electrons can grow in time. For large detuning Δ_p or large losses due to radiation emission from the resonator or to dissipation the growth rate of the wave is determined by the expression

$$\mathrm{Im}(\omega)\simeq\frac{q\nu}{\Delta_p^2+\nu^2}, \tag{24}$$

whence it follows that an increase in the detuning or the losses decreases the growth rate. However, for $\Delta_p^2>4q$ the losses ($\nu\neq 0$) determine the instability.

PLASMA RESONANCE IN A NONUNIFORM ELECTRON LAYER

It is obvious that in an experiment it is very difficult to achieve $\partial\Omega^2/\partial r^2\simeq 0$ in the region $r_-<r<r_+$. In a real electron layer the electron density will depend on the radius r. For this reason, it is desirable to investigate the possibility of electromagnetic wave generation under plasma resonance conditions in a nonuniform electron layer. Let the electron density increase gradually from zero at the point $r=r_-$ up to its maximum value and then fall off continuously to zero at the point $r=r_+$. Plasma resonance will occur near the points $r=r_1$ and $r=r_2$ ($r_-<r_1<r_2<r_+$), satisfying the equation

$$\left(\omega^{(0)}-\frac{mV_0}{r}\right)^2-2\frac{V_0^2}{r^2}-\Omega^2(r)=0. \tag{25}$$

In this case the dispersion relation assumes the form

$$\omega^{(1)}=\Delta_\nu-i\left\{\nu+\frac{\pi^2 x_1^{2m-1}(\eta^{2m}-1)^2 h\omega^{(0)}\,\mathrm{sign}(\omega_m)}{2^{m-1}[(m-1)!]^2\eta^{2m+1}b}\times\left(\frac{V_0}{r\omega_m}-\frac{1}{2}\frac{\eta^{2m}+1}{\eta^{2m}-1}\right)^2\right\}, \tag{26}$$

where

$$\omega^{(1)}=\omega-\omega^{(0)},$$

$$h=\left|\Omega^2/\frac{\partial\Omega^2}{\partial r}\right|_{r=r_1}+\left|\Omega^2/\frac{\partial\Omega^2}{\partial r}\right|_{r=r_2}. \tag{27}$$

The expression (26) determines the frequency shift and growth rate (damping rate) of the wave. According to this relation, instability occurs when the second term in braces in Eq. (26) is negative ($\omega^{(0)}\mathrm{sign}(\omega_m)<0$) and larger in modulus than the first term, i.e., when the phase velocity of the wave near the electron layer is less than the velocity of the electrons, while the losses due to radiation emission from the resonator and to dissipation are negligible.

We note that wave generation by a nonuniform electron layer under plasma resonance conditions does not require that resonance conditions of the type $\omega^{(0)}=mV_0/r_-$ (Cherenkov resonance) or $\omega^{(0)}\simeq mV_0/r_- -\{2V_0^2/r_-^2+\Omega^2\}^{1/2}>0$ (plasma resonance in an uniform layer ($\partial\Omega^2/\partial r\simeq 0$)) be satisfied. In our case a nonuniform electron layer is unstable with respect to excitation of waves whose characteristic frequencies $\omega^{(0)}$ satisfy the inequalities

$$\frac{mV_0}{r_-}-\left(2\frac{V_0^2}{r_-^2}+\Omega_M^2\right)^{1/2}<\omega^{(0)}<\frac{mV_0}{r_-}-\left(2\frac{V_0^2}{r_-^2}\right)^{1/2}>0,$$

$$\omega^{(0)}>0, \tag{28}$$

where Ω_M is the maximum of the function $\Omega(r)$.

At plasma resonance in a nonuniform layer the energy of the electromagnetic field grows as a result of the interaction of the wave not with all electrons but only with the electrons located near the points $r=r_1$ and $r=r_2$. For this reason, the growth rates are lower here than in the case of a plasma resonance in an uniform layer.

CONCLUSIONS

It follows from the foregoing analysis that wave generation in an orbitron is possible as a result of Cherenkov or plasma resonances. Under Cherenkov resonance conditions the growth rate of a wave can reach its maximum value, where, according to the relation (16), it is proportional to the cube root of the small parameter $\delta r/b$. In the case of plasma resonance in an uniform electron layer, according to expression (21), the maximum growth rate is proportional to the square root of $\delta r/b$. Under plasma resonance conditions in a nonuniform layer, according to (26), the growth rate is proportional to the first power of the small parameter h/b. However, wave generation due to plasma resonance in a nonuniform electron layer has a number of advantages. Transferring energy to the wave, the electron layer approaches the axis of the resonator (r_- decreases), and in an uniform layer this destroys the Cherenkov and plasma resonance conditions. As a result, a different wave with a higher frequency can be fall into resonance. The conditions for plasma resonance in a nonuniform layer do not break down as the radius r_- changes. The position of the plasma resonance points r_1 and r_2 changes. They approach the point where the electron density is maximum, when the radius r_- decreases. In the cases of Cherenkov and plasma resonances in an uniform layer the frequencies of the growing and damped waves are close or equal to one another. Therefore weak nonlinear effects can transform a growing wave into a damped wave. Under plasma resonance conditions in a nonuniform layer the dispersion relation has one solution (26). For this reason, here weak nonlinear effects apparently cannot prevent wave intensification. We also note that in a nonuniform layer many waves can be amplified simultaneously on account of plasma resonance.

[1] Z. S. Chernov, Radiotekh. Elektron. **1**, 1428 (1956).
[2] J. Alexeff and F. Dyer, Phys. Rev. Lett. **45**, 351 (1980).
[3] K. N. Stepanov, Zh. Tekh. Fiz. **35**, 1002 (1965) [Sov. Phys. Tech. Phys. **10**, 773 (1965)].
[4] K. N. Stepanov, Zh. Tekh. Fiz. **35**, 1349 (1965) [Sov. Phys. Tech. Phys. **10**, 1048 (1965)].
[5] V. V. Dolgopolov and A. Ya. Omel'chenko, Zh. Éksp. Teor. Fiz. **58**, 1384 (1970) [Sov. Phys. JETP **31**, 741 (1970)].

Translated by M. E. Alferieff

Self-consistent problem of the excitation of beams of magnetostatic waves

S. V. Zagryadskiĭ and A. G. Rezvanov

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
(Submitted December 27, 1996; resubmitted April 16, 1997)
Zh. Tekh. Fiz. **68**, 95–101 (August 1998)

An analytical solution is obtained for the self-consistent problem of the excitation of beams of magnetostatic waves by a converter of arbitrary type. Radiation patterns are calculated numerically for stripline converters at the frequencies where magnetostatic surface waves exist.
 [S1063-7842(98)01608-0]

INTRODUCTION

Beams of magnetostatic waves are usually taken to be a superposition of these waves with equal frequencies, but with wave vectors that differ both in magnitude and direction. Knowledge of the relationships governing the formation of wave packets of magnetostatic waves produced by different kinds of converters is required in order to improve the methods for doing calculations for specific devices and for integrated microwave circuits based on thin ferrite films. The problems of exciting beams of magnetostatic surface and volume waves by a point current element and by a filament with a constant current along its length have been examined in a number of papers.[1–4] Focusing converters, which emit beams of waves, have been studied.[5,6] These problems, however, have all been solved in the approximation of a given converter current.

In this paper, for the first time we solve the self-consistent problem of the excitation of beams of magnetostatic surface and volume waves, including the effect on the converter of the magnetization which it excites in the ferrite film. An analytical solution is obtained in closed form for converters of arbitrary type, and numerical calculations are done for stripline converters at frequencies where magnetostatic surface waves exist. A comparison of this solution with the approximation of a fixed current (field) shows that it is necessary, in principle, to take the back reaction of the excited waves on the converter into account.

EQUATIONS FOR THE SELF-CONSISTENT PROBLEM

Let us consider a ferrite film of thickness d which is unbounded in the yz plane, has been magnetized to saturation in the η direction, and is excited by a converter of arbitrary type (Fig. 1, where *1* denotes the ferrite film). In the transmission line which forms the converter, we assume that (in the absence of a ferrite film) only a wave having the lowest mode ν with propagation constant γ and electromagnetic field $\mathbf{E}_{\pm\nu}=\mathbf{E}_{\pm\nu 0}(x,y)\exp(\mu i\gamma z)$, $\mathbf{H}_{\pm\nu}=\mathbf{H}_{\pm\nu 0}(x,y)\times\exp(\mu i\gamma z)$ can propagate. The characteristic vector functions $\mathbf{E}_\nu$ and $\mathbf{H}_\nu$ are known for most types of transmission lines used as magnetostatic wave converters. Without loss of generality in the discussion, we shall assume that the magnetic field in the transmission line is transverse. In addition, assuming that the line is lossless, we take the functions $\mathbf{E}_{\nu 0}$ and $\mathbf{H}_{\nu 0}$ to be real. In this case,[7] we can set $\mathbf{E}_{\nu 0}=\mathbf{E}_{-\nu 0}$ and $\mathbf{H}_{\nu 0}=\mathbf{H}_{-\nu 0}$. The magnetic field of the incident electromagnetic wave of the fundamental ν mode equals $c\mathbf{H}_\nu$ in the $z=0$ cross section. The converter is loaded by an arbitrary load which gives a reflection coefficient of Γ with respect to the transverse component of the magnetic field at the $z=L$ cross section. The eigenfunctions for the magnetization $\mathbf{m}_n=\mathbf{m}_n^0(x)\exp(-\mathbf{k}_n\cdot\mathbf{r})$ and the dispersion characteristics of the magnetostatic waves are also assumed known. (Here $\mathbf{k}_n$ and $\mathbf{r}$ are the wave vector of the wave and the radius vector in the plane of the film.)

The magnetization excited in a ferrite film by an external rf magnetic field $\mathbf{h}_\nu$ at frequency ω can be found[8] in the form of an integral expansion over the magnetization eigenfunction in a two-dimensional wave number space in the plane of the film:

$$\mathbf{M}=\sum_n \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(c_n\mathbf{m}_n)\,dk_{ny}dk_{nz}, \tag{1}$$

where

$$c_n=\varphi_n\int_V(\mathbf{h}_\nu\mathbf{m}_n^*)\,dV,$$

$$\varphi_n=\frac{i\omega_M}{\Phi_n}\,\frac{1}{\omega_n-\omega},\quad \Phi_n=(2\pi)^2\int_0^d[\mathbf{m}_n^0\times\mathbf{m}_n^{0*}]_\eta\,dx, \tag{2}$$

and V is the volume of the film.

The sum is taken over all branches of the wave spectrum. The excited magnetization is determined from Eq. (1) in the electrodynamic, magnetostatic, or dipole-exchange approximation, depending on the approximation in which the magnetization eigenfunctions $\mathbf{m}_n$ have been found.

We shall take $\mathbf{h}_\nu$ to be the magnetic field of the converter at the site of the ferrite film. In order to account for the back reaction on the converter of the magnetization excited in the film, this field must be represented in terms of a given magnetization and we must obtain a system of equations for the self-consistent problem. A representation of this sort can be found based on waveguide excitation theory,[7] which is also applicable in this case because the magnetic field of the transmission line which forms the converter goes to zero in the transverse cross section with distance from the longitu-

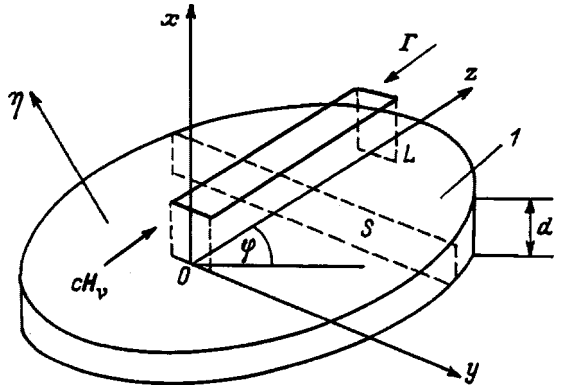

FIG. 1.

dinal axis of the line. The field $\mathbf{h}_\nu$ is made up of the field of the wave incident on the entrance plus $\mathbf{H}_\nu$ and the field of the waves excited by the ferrite film (including reflection from the load at the other end of the line). The last is defined as the field excited in the transmission line by an external equivalent magnetic current, in the source region when the source is locally removed. If the ferrite occupies a small part of the transverse cross section of the line (as happens for magnetostatic wave structures), then the single mode approximation[9] is applicable:

$$\mathbf{h}_\nu=[c+c_\nu(z)]\mathbf{H}_\nu+\{c_{-\nu}(z)+[c+c_\nu(L)]\Gamma\exp(-2i\gamma L)\}\mathbf{H}_{-\nu}, \tag{3}$$

where

$$c_\nu(z)=-\frac{i\omega\mu_0}{N_\nu}\int_0^z dz\int_S(\mathbf{M}\cdot\mathbf{H}_{-\nu})dS,$$

$$c_{-\nu}(z)=-\frac{i\omega\mu_0}{N_\nu}\int_z^L dz\int_S(\mathbf{M}\cdot\mathbf{H}_{\nu})dS,$$

$$N_\nu=\int_S\{[\mathbf{E}_{-\nu}\times\mathbf{H}_\nu]-[\mathbf{E}_\nu\times\mathbf{H}_{-\nu}]\}_z dS, \tag{4}$$

S is the transverse cross section of the film, and N_ν is the norm of the electromagnetic wave in mode ν.

Note that the total field in a transmission line with a ferrite film is made up of the field $\mathbf{h}_\nu$ (3) and the field induced by the magnetization. The system of singular integral Eqs. (1) and (3) for $\mathbf{M}$ and $\mathbf{h}_\nu$ is also a mathematical formulation for the self-consistent problem of the excitation of magnetization in a ferrite film.

SOLUTION OF THE SYSTEM OF EQUATIONS FOR THE SELF-CONSISTENT PROBLEM

The amplitudes $c_\nu(z)$ and $c_{-\nu}(z)$ of the electromagnetic waves can be taken as new independent unknown functions in the system of Eqs. (1) and (3). Precisely in this case the system is relatively easy to solve, owing to the integral character of these functions. On introducing the functional

$$\Phi(k_{nz})=\int_0^L[c_\nu(z)\exp(-i\gamma z)+c_{-\nu}(z)\exp(i\gamma z)]\exp(ik_{nz}z)dz \tag{5}$$

the system of Eqs. (1) and (3) reduces to a Fredholm integral equation of the second kind,

$$\Phi(k_{nz})=\int_{-\infty}^{\infty}\Phi(k'_{nz})K(k_{nz},k'_{nz})dk'_{nz}+\Psi(k_{nz}). \tag{6}$$

The functions which show up in the kernel are given in the Appendix. They depend on the coupling parameter

$$F_\nu(k_{nz})=\frac{\omega\mu_0}{N_\nu}\int_{-\infty}^{\infty}\varphi_n(k_{ny},k_{nz})|I_\nu(k_{ny},k_{nz})|^2dk_{ny} \tag{7}$$

which characterizes the interaction of the converter with the ferrite film in the self-consistent problem. The excitation integral in Eq. (7) is given by

$$I_\nu(k_{ny},k_{nz})=\int_S\mathbf{H}_{\nu 0}(x,y)\mathbf{m}_n^{0*}(x,k_{ny},k_{nz})\exp(ik_{ny}y)dS. \tag{8}$$

In the approximation of a given field in the converter, $F_\nu(k_{nz})\equiv 0$ and the kernel goes to zero and the functional $\Phi(k_{nz})$ equals the free term in the equation

$$\Psi(k_{nz})=c\left\{\frac{(g-1)\Gamma\exp(-2i\gamma L)}{1+a\Gamma\exp(-2i\gamma L)}[b(k_{nz})+h(k_{nz})]-[d(k_{nz})+f(k_{nz})]\right\}. \tag{9}$$

We solve the Fredholm equation of the second kind by successive approximations,

$$\Phi^{[j+1]}(k_{nz})=\int_{-\infty}^{\infty}\Phi^{[j]}(k_{nz})K(k_{nz},k'_{nz})dk'_{nz}+\Psi(k_{nz}), \tag{10}$$

taking the zeroth ($j=0$) approximation $\Phi^{[0]}(k_{nz})=\Psi(k_{nz})$ to be the value of the free term in the equation, calculated for the specified (given) field in the transmission line

$$\mathbf{h}_\nu=c\mathbf{H}_\nu+c\Gamma\exp(-2i\gamma L)\mathbf{H}_{-\nu} \tag{11}$$

which corresponds to $c_\nu(z)\equiv 0$ and $c_{-\nu}(z)\equiv 0$ in Eq. (3). The first approximation ($j=1$) yields a solution to the problem which takes into account the back reaction of the propagating magnetostatic waves excited by the converter with the local magnetization. As a result, we find the amplitudes c_n of the waves in the beam,

$$c_n=\varphi_n I_\nu(k_{ny},k_{nz})\left\{c\,\frac{\exp[iL(k_{nz}-\gamma)]-1}{i(k_{nz}-\gamma)}+\Phi(k_{nz})+[c+c_\nu(L)]\Gamma\exp(-2i\gamma L)\times\frac{\exp[iL(k_{nz}+\gamma)]-1}{i(k_{nz}+\gamma)}\right\}, \tag{12}$$

where $c_\nu(L)$ is determined in terms of the functional $\Phi(k_{nz})$ from an algebraic equation obtained by substituting Eq. (12) in Eq. (4) with $z=L$,

$$c_\nu(L)=\frac{\int_{-\infty}^{+\infty}H(k'_{nz})\Phi(k'_{nz})dk'_{nz}+c[g+a\Gamma\ \exp(-2i\gamma L)]}{1+a\Gamma\ \exp(-2i\gamma L)}. \tag{13}$$

From the known c_n we determine the rf magnetization excited by the converter in the far zone by calculating the double integral (1) in the polar coordinates $k_{ny}=k_n\ \sin(\Theta)$, $k_{nz}=k_n\ \cos(\Theta)$, $y=r\ \sin(\varphi)$, $z=r\ \cos(\varphi)$. In the inner integral with respect to k_n, we exclude the residues corresponding to traveling magnetostatic waves (we consider radiation only in the direction of positive y), and calculate the outer integral with respect to Θ (a superposition of traveling waves) by the method of stationary phase. Finally, we obtain the solution for the magnetization in the form

$$\mathbf{M}(r,\varphi)\approx c'_n(k_{n0}(\Theta_S(\varphi)),\Theta_S(\varphi))\mathbf{m}_n^0(x,k_{n0}(\Theta_S(\varphi)),\Theta_S(\varphi))r^{-1/2}\ \exp[-ik_{n0}(\Theta_S(\varphi))r\ \cos(\Theta_S(\varphi)-\varphi)]\times\exp\left[\frac{-\gamma\Delta H}{V_n(\Theta_S(\varphi))}\frac{y}{\sin(\Theta_S(\varphi))}\right], \tag{14}$$

where

$$c'_n(k_{n0}(\Theta_S(\varphi)),\Theta_S(\varphi))=-2\pi i\ \frac{c\omega_M}{V_n(\Theta_S(\varphi))}\times k_{n0}(\Theta_S(\varphi))\ \frac{I_\nu(k_{n0}(\Theta_S(\varphi)),\Theta_S(\varphi))}{\Phi_n(k_{n0}(\Theta_S(\varphi)),\Theta_S(\varphi))}\times\ \xi(k_{n0}(\Theta_S(\varphi)),\Theta_S(\varphi))\sqrt{\frac{2\pi}{\Psi''(\Theta_S(\varphi))}}\ \exp(i\pi/4), \tag{15}$$

φ is the polar angle of the observation point, reckoned from the z axis of the converter, $k_{n0}(\Theta_S(\varphi))$ is the root of the dispersion relation $\omega_n(\mathbf{k}_n)=\omega$, and $V_n(\Theta_S(\varphi))=\partial\omega_n/\partial k_n$ is the corresponding group velocity of the magnetostatic wave.

The phase $\Psi(\Theta)$ is a function of the angle Θ, $\Psi(\Theta)=k_{n0}(\Theta)\cos(\Theta-\varphi)$, determined for a given φ from the equation for the fixed points,

$$k'_{n0}(\Theta)\cos(\Theta-\varphi)-k_{n0}(\Theta)\sin(\Theta-\varphi)=0. \tag{A}$$

The fixed points found from this equation are denoted above by $\Theta_S(\varphi)$. The second derivative $\Phi''(\Theta_S(\varphi))$ in Eq. (15) is calculated for the same value, $\Theta_S(\varphi)$. The function $\xi(k_{n0}(\Theta_S(\varphi)),\Theta_S(\varphi))$ in Eq. (15) is given by the formula

$$\xi(k_{n0},\Theta_S)=\frac{\exp[iL(k_{nz}-\gamma)]-1}{i(k_{nz}-\gamma)}+\left[\frac{\pi iF_\nu(\gamma)}{\gamma}\left(\frac{\exp[iL(k_{nz}-\gamma)]}{k_{nz}-\gamma}-\frac{\exp[iL(k_{nz}+\gamma)]}{k_{nz}+\gamma}\right)-\frac{4\pi iF_\nu(k_{nz})\gamma}{(k_{nz}+\gamma)^2(k_{nz}-\gamma)}-\frac{2\pi iF_\nu(\gamma)}{k_{nz}+\gamma}\left(\frac{\exp[iL(k_{nz}+\gamma)]}{k_{nz}+\gamma}-iL\right)\right]\times\frac{(g-1)\Gamma\ \exp(-2i\gamma L)}{1+a\Gamma\ \exp(-2i\gamma L)}$$
$$-\left[-\frac{4\pi iF_\nu(k_{nz})\gamma}{(k_{nz}-\gamma)^2(k_{nz}+\gamma)}+2\pi iF_\nu(\gamma)\frac{\exp[iL(k_{nz}-\gamma)]}{(\gamma-k_{nz})}\times\left(\frac{1}{(\gamma-k_{nz})}+\frac{1}{2\gamma}+iL\right)+\pi iF_\nu(\gamma)\times\frac{\exp[iL(k_{nz}-\gamma)]-\exp(-2i\gamma L)+1}{\gamma(k_{nz}+\gamma)}\right]+\frac{(1-g)\Gamma\ \exp(-2i\gamma L)}{1+a\Gamma\ \exp(-2i\gamma L)}\times\frac{\exp[iL(k_{nz}+\gamma)]-1}{i(k_{nz}+\gamma)}+\frac{\Gamma\ \exp(-2i\gamma L)}{1+a\Gamma\ \exp(-2i\gamma L)}\times\frac{\exp[iL(k_{nz}+\gamma)]-1}{i(k_{nz}+\gamma)}\ \frac{\pi^2}{\gamma^2}\times F_\nu^2(\gamma)\left[2(g-1)\Gamma\ \frac{2i\gamma L-\exp(-2i\gamma L)+1}{1+a\Gamma\ \exp(-2i\gamma L)}\times\exp(-2i\gamma L)+\exp(-2i\gamma L)+2i\gamma L\right]. \tag{16}$$

The resulting analytical solution (14) of the self-consistent problem for the excitation of beams of magnetostatic surface and volume waves is extremely general in character, since all the features of the electrodynamic system are taken into account in the eigenfunctions and dispersion characteristics of the waves. This solution is valid for any type of converter and for arbitrary waveguide structures, in particular, those containing dielectric layers, multilayer ferrite films, and metallic shields.

THE EMISSION OF WAVE BEAMS OF MAGNETOSTATIC SURFACE WAVES BY STRIPLINE CONVERTERS

We shall consider the development of the directional diagrams using the example of a ferrite film magnetized along its longitudinal z axis. Figures 2–6 show the normalized amplitude radiation patterns ARP(φ) of a converter short-circuited at the end ($\Gamma=1$) for different frequency bands within which surface magnetostatic waves exist. The amplitude directional diagram represents the angular dependence of the modulus of the amplitude factor c'_n for the magnetization eigenfunction $\mathbf{m}_n^0$ in Eq. (14). Eliminating all terms from Eq. (14) which do not contain an angular dependence and using Eq. (15), for the amplitude directional diagram we can write

$$\mathrm{ARP}(\varphi)=\left|\frac{I_\nu(k_{n0}(\varphi),\Theta_S(\varphi))k_{n0}(\varphi)\xi(k_{n0}(\varphi),\Theta_S(\varphi))}{V_n(\varphi)\Phi_n(k_{n0}(\varphi),\Theta_S(\varphi))\sqrt{\Psi''(\Theta_S(\varphi))}}\right|. \tag{17}$$

Here the square of ARP(φ) yields the radiation pattern of the converter with respect to power.

The amplitude radiation pattern was calculated numerically using Eq. (17) for a converter based on a symmetric stripline (see the inset to Fig. 2) with a dielectric filling (ε

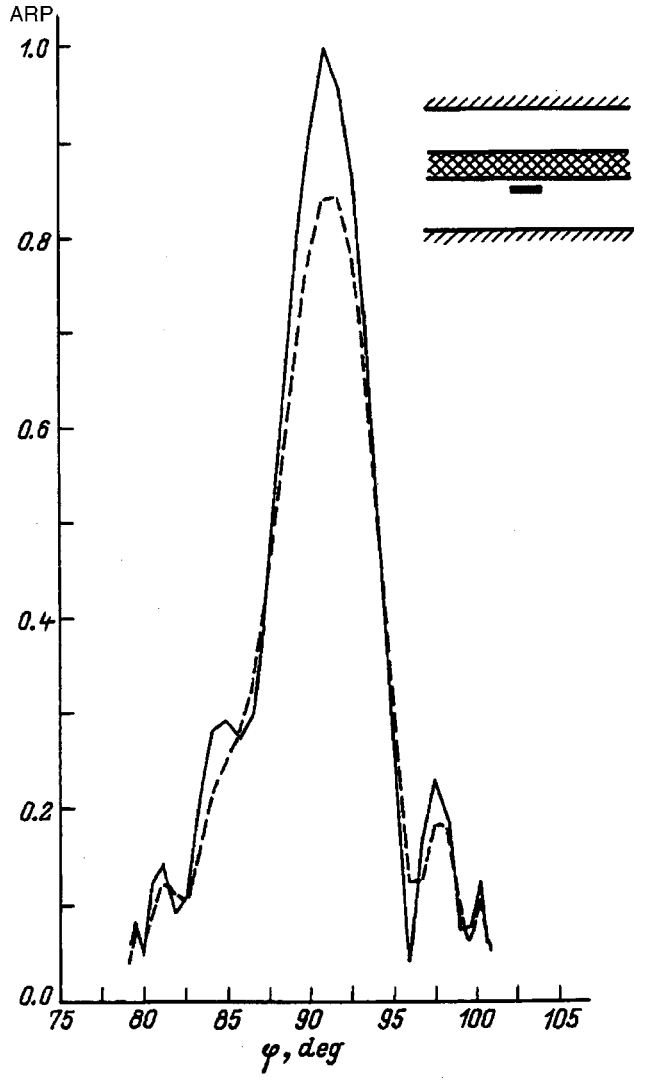

FIG. 2. Amplitude directional diagrams. The smooth curve is the self-consistent solution and the dashed curve is for the approximation of a given field; $f=5.2$ GHz.

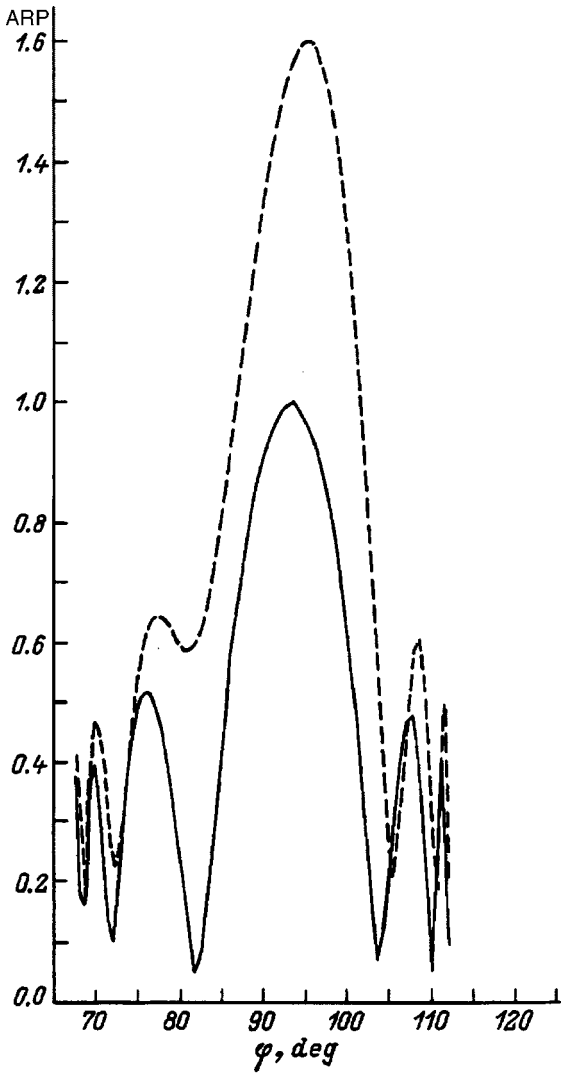

FIG. 3. As in Fig. 2, but for $f=5.0$ GHz.

$=9.8$), a strip width of 200 μm, and a separation between the metal shields of 1 mm. A ferrite film with a saturation magnetization of 1750 G and thickness $d=50$ μm was separated from the strip by a distance of 10 μm. The length of the converter is $L=4$ mm. The magnetizing field equals 1030 Oe. In Figs. 2–5, the dashed curves were calculated in the approximation of a given field ($F_\nu(k_{nz})\equiv 0$), while the other curves were calculated taking the back reaction of the excited magnetostatic waves on the converter into account ($F_\nu(k_{nz})\neq 0$). The radiation patterns for each family of curves were normalized to the maximum of the amplitude radiation pattern corresponding to the solution of the self-consistent problem.

As the frequency is lowered, the width of the radiation patterns increases owing to a widening of the sector of polar angles φ within which surface magnetostatic waves exist. As opposed to the case of a uniform distribution of current over the length of the converter,[4] a nonuniform distribution leads to an asymmetry in the radiation patterns relative to $\varphi=90°$ (the y axis), even for relatively small converter lengths ($L=4$ mm). The preferred propagation direction of the waves is shifted to larger (90°) polar angles for the observation point.

At higher and intermediate magnetostatic surface wave frequencies (in this case, from 4738 to 5334 MHz), the shapes of the radiation patterns calculated in the approximation of a given field and taking the back reaction of the excited waves into account (Figs. 2 and 3) differ little and are determined, as noted previously,[4] by the angular dependence of the group velocity. At intermediate frequencies, however, the given field approximation yields a higher result for the amplitude of the magnetization in the far field of the converter (especially for $\varphi\approx 90°$), which may cause the power emitted by the converter to exceed the input power and, therefore, violate the conservation of energy. At the lower frequencies in this range (Fig. 4), two preferred directions for the wave emission show up, which are symmetric with respect to the y axis ($\varphi=90°$). This is the well-known[4] doubling of beams of surface magnetostatic waves in the far zone. However, compared to the given field approximation,[4] the radiation patterns for each of these directions turn out to be much narrower and the peaks are shifted by several de-

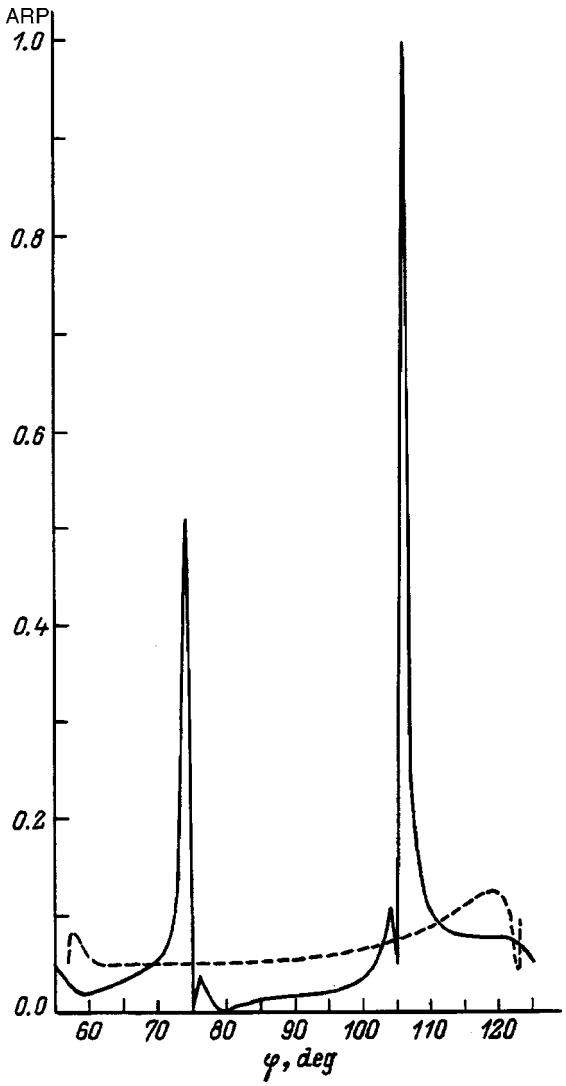

FIG. 4. As in Fig. 2, but for $f=4.8$ GHz.

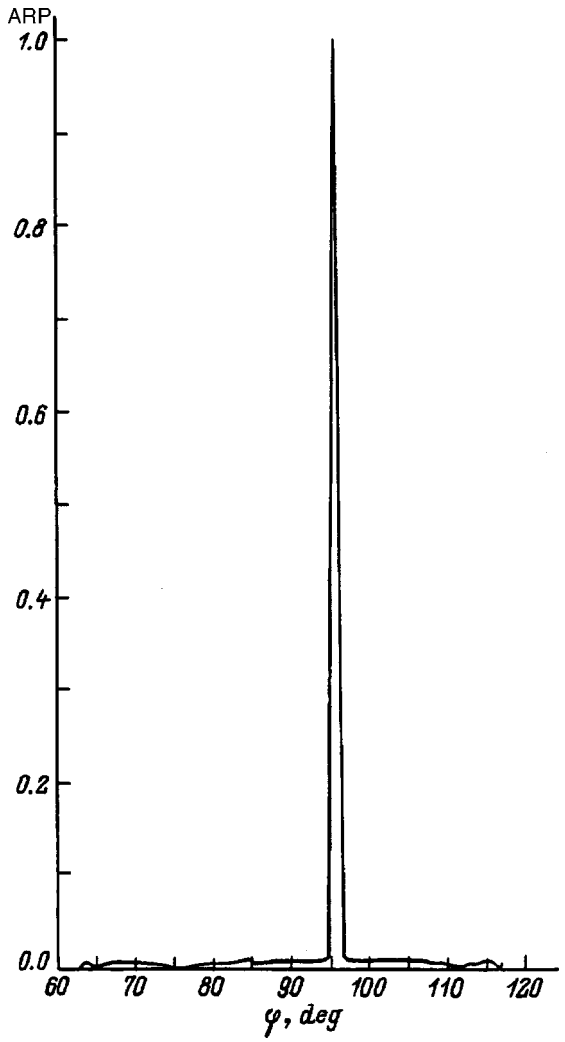

FIG. 5. As in Fig. 2, but for $f=4.91$ GHz.

grees from the edges to the center of the sector of angles within which surface waves can propagate. Sharp directional diagrams of this sort have been observed experimentally (see Fig. 13.20 of Vashkovskiĭ *et al.*[10]). The angular position of the peaks in the directional diagram in the self-consistent problem is determined, in accordance with Eqs. (16) and (17), by the condition of phase synchronism

$$\gamma=\pm k_{nz}, \tag{18}$$

corresponding to equality of the phase velocity of the electromagnetic wave in the transmission line forming the converter to the component of the phase velocity of the magnetostatic wave along the z axis. Magnetostatic waves satisfying the condition (18) interact most efficiently with the converter and have the largest amplitude in a beam. Owing to the symmetry of the dispersion characteristics of the magnetostatic waves, the angular positions of these peaks are also symmetric with respect to the $\varphi=90°$ direction, but the level of the amplitude radiation pattern in these peaks differs because of the nonuniform distribution of the field along the converter. Within a narrow frequency band (on the order of 10–15 MHz) this difference leads to relative suppression of one of the peaks at $\varphi<90°$ and the converter becomes a ''superdirectional'' antenna with a directional diagram of width less than 1° at a level of 0.707, which emits a wave beam only in a direction $\varphi>90°$ (Fig. 5).

At intermediate and higher magnetostatic wave frequencies, the phase synchronism condition (18) is satisfied only by waves with large components of the wave vectors k_{nz}, for which the excitation integral (8) is small. As they have a small amplitude, these waves have no influence on the formation of the radiation pattern (Figs. 3 and 4).

The finite transverse dimensions of the converter have not been taken into account in the earlier models,[1–6] but they do affect the radiation pattern. Thus, at the higher and intermediate magnetostatic wave frequencies, a change in the strip width leads to a change in the angular position of the minima in the amplitude radiation pattern, especially near the principal maximum, whose shape then changes little. On the other hand, at the lower frequencies, where the angular position of the peaks in the amplitude radiation pattern is determined by the phase synchronism condition (18), widening the strip causes an enlargement of the spatial region within which the electromagnetic and magnetostatic waves interact; this leads to a relative enhancement in these peaks (Fig. 6).

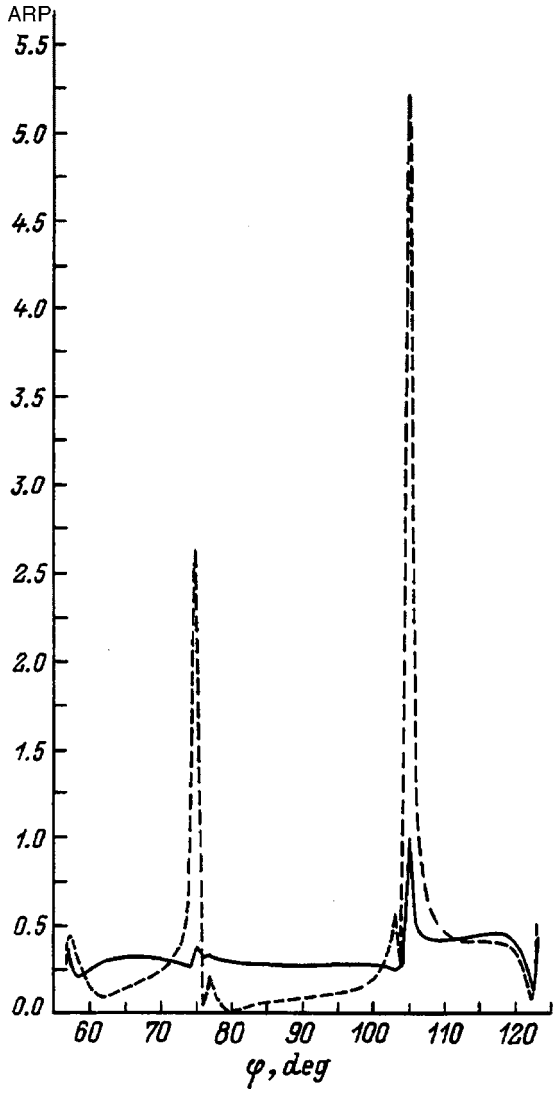

FIG. 6. The same as in Fig. 2, but for $f=4.8$ GHz. Strip width (μm): 1 (smoothed curve), 100 (dashed curve).

CONCLUSION

In this paper we have demonstrated the fundamental need to solve the problem of magnetostatic wave excitation in ferrite films in a self-consistent formulation. A general analytical solution of the self-consistent problem has been obtained for a converter of arbitrary type which is valid in the electrodynamic, magnetostatic, and dipole–exchange approximations for an arbitrary direction of the magnetizing field and any sort of waveguide structure, especially those containing dielectric layers, multilayer ferrite films, and metallic shields. It has been shown that at the intermediate and higher frequencies where magnetostatic waves exist, the radiation pattern of a converter is asymmetric with respect to a direction perpendicular to the longitudinal axis of the converter, even when the ferrite film is magnetized along this axis. At lower frequencies, the shape of the directional diagram is determined by the phase synchronism of the excited magnetostatic waves and the electromagnetic wave in the transmission line which forms the converter, while the level of the peaks in the radiation pattern depends significantly on the transverse dimensions of the converter.

APPENDIX

The kernel of the integral equation (6) has the form

$$K(k_{nz},k'_{nz})=\frac{[b(k_{nz})+h(k_{nz})]\Gamma\ \exp(-2i\gamma L)}{1+a\Gamma\ \exp(-2i\gamma L)}\times H(k'_{nz})-P(k_{nz},k'_{nz})-S(k_{nz},k'_{nz}),$$

where

$$H(k'_{nz})=F_\nu(k'_{nz})\ \frac{\exp[iL(\gamma-k'_{nz})]-1}{\gamma-k'_{nz}},$$

$$P(k_{nz},k'_{nz})=\frac{F_\nu(k'_{nz})}{i(\gamma-k'_{nz})}\left\{\frac{\exp[iL(k_{nz}-k'_{nz})]-1}{k_{nz}-k'_{nz}}-\frac{\exp[iL(k_{nz}-\gamma)]-1}{k_{nz}-\gamma}\right\},$$

$$S(k_{nz},k'_{nz})=\frac{F_\nu(k'_{nz})}{i(\gamma+k'_{nz})}\left\{\frac{\exp[iL(k_{nz}-k'_{nz})]-1}{k_{nz}-k'_{nz}}-\frac{\exp[iL(k_{nz}-k'_{nz})]-\exp[iL(k'_{nz}+\gamma)]}{k_{nz}+\gamma}\right\},$$

$$b(k_{nz})=\int_{-\infty}^{+\infty}P(k_{nz},k'_{nz})\ \frac{\exp[iL(k'_{nz}+\gamma)]-1}{i(k'_{nz}+\gamma)}\,dk'_{nz},$$

$$d(k_{nz})=\int_{-\infty}^{+\infty}P(k_{nz},k'_{nz})\ \frac{\exp[iL(k'_{nz}-\gamma)]-1}{i(k'_{nz}-\gamma)}\,dk'_{nz},$$

$$f(k_{nz})=\int_{-\infty}^{+\infty}S(k_{nz},k'_{nz})\ \frac{\exp[iL(k'_{nz}-\gamma)]-1}{i(k'_{nz}-\gamma)}\,dk'_{nz},$$

$$h(k_{nz})=\int_{-\infty}^{+\infty}S(k_{nz},k'_{nz})\ \frac{\exp[iL(k'_{nz}+\gamma)]-1}{i(k'_{nz}+\gamma)}\,dk'_{nz},$$

$$a=\int_{-\infty}^{+\infty}H(k'_{nz})\ \frac{\exp[iL(k'_{nz}+\gamma)]-1}{i(k'_{nz}+\gamma)}\,dk'_{nz},$$

$$g=\int_{-\infty}^{+\infty}H(k'_{nz})\ \frac{\exp[iL(k'_{nz}-\gamma)]-1}{i(k'_{nz}-\gamma)}\,dk'_{nz}.$$

[1] L. B. Gol'dberg and V. V. Penzyakov, Zh. Tekh. Fiz. **56**, 1049 (1986) [Sov. Phys. Tech. Phys. **31**, 614 (1986)].
[2] L. B. Gol'dberg, Zh. Tekh. Fiz. **56**, 1893 (1986) [Sov. Phys. Tech. Phys. **31**, 1132 (1986)].
[3] S. V. Vlaskin and G. M. Novikov, Izv. Vyssh. Ucheb. Zaved. Fiz., No. 4, 32 (1989).
[4] A. V. Vashkovskiĭ, A. V. Stal'makhov, and D. G. Shakhnazaryan, Izv. Vyssh. Ucheb. Zaved. Fiz., No. 11, 67 (1988).

[5] A. V. Vashkovskiĭ, K. V. Grechushkin, A. V. Stal'makhov, and V. A. Tyulyukin, Radiotekh. Elektron. **31**, 838 (1986).
[6] A. V. Vashkovskiĭ, K. V. Grechushkin, A. V. Stal'makhov, and V. A. Tyulyukin, Radiotekh. Elektron. **32**, 1176 (1987).
[7] L. A. Vaĭnshteĭn, *Electromagnetic Waves* [in Russian], Radio i Svyaz', Moscow (1988), 440 pp.
[8] S. V. Zagryadskiĭ, Radiotekhnika, No. 3, 29 (1991).
[9] S. V. Zagryadskiĭ, Radiotekhnika, No. 6, 31 (1990).
[10] A. V. Vashkovskiĭ, V. S. Stal'makhov, and Yu. P. Sharaevskiĭ, *Magnetostatic Waves in Microwave Electronics* [in Russian], Izd-vo Saratovskogo Universiteta (1993), 311 pp.

Translated by D. H. McNeill

Interaction of H^- ions with foil targets in the charge exchange system of a beam transport channel

A. S. Artemov, Yu. K. Baĭgachev, A. K. Gevorkov,† and A. O. Sidorin

Joint Institute for Nuclear Research, 141980 Dubna, Moscow District, Russia
(Submitted March 3, 1997)
Zh. Tekh. Fiz. **68**, 102–105 (August 1998)

The possibility of experimentally modeling the interaction of high energy H^- ions with foil targets using beams with more accessible lower energies that have the same dimensionless interaction parameter and similar current characteristics is pointed out. Results are presented from the first stage of a study of the beam–foil stripping of 2 and 7 MeV H^- ions. An analysis of the charge composition of the beam after a carbon foil serves as a basis for determining the corresponding cross sections for stripping of the ions and ionization of the product hydrogen atoms. The data from these and other beam–foil experiments are in good agreement with theoretical cross sections on carbon at different energies, as well with calculated values based on the superposition of experimental cross sections for gaseous carbon-containing targets.
 [S1063-7842(98)01708-5]

INTRODUCTION

A peculiarity of the atomic structure of H^- ions makes it possible to use a charge exchange technique for controlling the particle fluxes in modern accelerator-storage ring systems.[1] Various internal foil targets are used to realize this technique at energies above ten MeV (see, e.g., Refs. 2–4). The target material and thickness are chosen taking into account the required working life of the target and the charge composition of the beam after the interaction. Assuming pairwise collisions and given the cross sections for the major channels of the stripping process in a target,[5,6] the relative numbers of H^- ions (η_-), H^0 atoms (η_0), and protons (η_+) in the beam are given by

$$\eta_- = \exp(-\sigma_{-1d}\cdot\hat{t}),$$

$$\eta_0 = \frac{\sigma_{-1d}}{\sigma_{-1d}-\sigma_{01}}\left[\exp(-\sigma_{01}\cdot\hat{t})-\exp(-\sigma_{-1d}\cdot\hat{t})\right],$$

$$\eta_+ = 1-\eta_- -\eta_0, \qquad (1)$$

where σ_{-1d} is the cross section (cm^2) for stripping of an H^- ion into the various states of the H^0 atom, including the continuum, σ_{01} is the average ionization cross section (cm^2) of the hydrogen atoms created by stripping, $\hat{t}=6\cdot10^{23}t/A_0$, and t and A_0 are, respectively, the thickness of the target ($\mathrm{g/cm}^2$) and its effective atomic number.

Characteristic plots of the charge composition of a beam as a function of the dimensionless interaction parameter $\sigma_{-1d}\cdot\hat{t}$ are shown in Fig. 1 for a ratio $\sigma_{-1d}/\sigma_{01}=2.5$. For the same functional dependence of these cross sections on energy, these curves are universal for the interaction of a beam of H^- ions with a target of a chosen material. In this case, a number of features of this interaction can be modeled experimentally at low (more accessible) energies and predictions can be made for high energy beams. The information of interest includes the optimum thickness and type of target, as well as its operating lifetime for a given intensity, duration, and repetition rate of the beam pulse. Here it is possible to test methods for monitoring the target integrity, e.g., from its electron or photon emission. The possibility of modeling the operating lifetime of a foil target is based on the fact that the dependence of the energy release in the target is analogous to that of the cross sections σ_{-1d} and σ_{01} ($\propto 1/\beta^2$, where β is the particle velocity) for a given effective charge of the beam particles over a wide range of energies from a few MeV to relativistic energies[7–9] with radiative heat transfer from the target surface. Thus, by creating the required charge composition of a beam for one energy and target thickness $\hat{t}$, it is possible to study the thermal loads expected from an analogous, in terms of intensity, flux of ions at another energy and target thickness corresponding to the same dimensionless parameter $\sigma_{-1d}\cdot\hat{t}$ (Fig. 1). This is especially important for choosing the optimum material, structure, and fabrication technique for foil targets in the charge exchange portion of a high energy beam transport channel.[4,10] At the same time, it has been shown[11] that the behavior of relativistic H^- ions differs from that of low-energy positive ions in the beam–

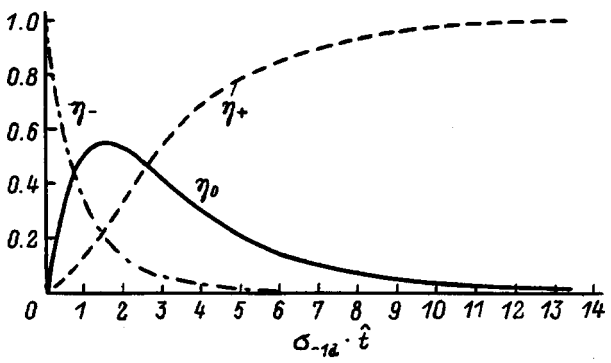

FIG. 1. The charge composition of a beam as a function of the beam–foil interaction parameter $\sigma_{-1d}\cdot\hat{t}$.

foil interaction processes responsible for the formation of fluxes of highly excited neutral atoms. No experimental information on these processes involving low-energy H^- ions is currently available. In this regard, despite the small relative number of these excited particles in a beam, a more detailed study of their production and loss at various energies for the same interaction time of the H^- ion beam with the ''field'' of the target material is of some interest.

In this paper we present data from the first stage of experiments performed in 1987–1988 to test the beam–foil interaction at low H^- ion energies corresponding to the expected dimensionless parameter $\sigma_{-1d}\cdot\hat{t}$ in the charge exchange section ($E_{H^-}=600$ MeV) of the MEGAN linear accelerator (Troitsk).[12]

MEASUREMENT TECHNIQUE

The setup of the experiments with 2 and 7 MeV ion beams is shown in Fig. 2. A beam of H^- ions was formed at the outlet of section *1* with spatially uniform ($E_{H^-}=2$ MeV) and spatially periodic ($E_{H^-}=7$ MeV) rf quadrupole focusing of the linear accelerator. Two types of foils were used as charge exchange targets: polymer films of colloxylin $\{C_6H_7O_2(OH)(ONO_2)_2\}_n$, with a high durability under the pressure drops, and a carbon foil with a thickness of 2 μg/cm^2 ($\approx$100 Å), which was deposited on a Ni grid with an optical transparency of 84% at the Kurchatov Institute of Atomic Energy. The thickness of the carbon film was determined to within $\pm$15% from the energy lost by α particles from an ^{241}Am source ($E_\alpha=5.486|_{85\%}$, $5.443|_{13\%}$ MeV) using a method, described elsewhere,[13] that takes into account the distortions in the measured spectra (with and without the target) induced by the analyzer. The polymer films were prepared by floatation on water followed by deposition on a metal grid with an optical transparency of 98%. Films of different thicknesses were obtained by varying the concentration of the initial material in the solvent (acetone). The thicknesses of the polymer films were measured interferometrically and lay within the range 500–2000 Å. Rings with different targets *2*, as well as with the grids (to determine their transparency relative to the H^- ions) were mounted in a rotatable drum *3*. Experiments showed that the grids without targets ensure complete passage of the H^- ion beam without loss, but with some change in its charge composition and energy spectrum. By rotating the drum to fixed angles, it is possible to change the target during the course of an experiment. The beam current prior to the interaction was monitored by an induction probe *4*. After the target, the flux of H^-, H^0, and H^+ particles was resolved spatially in terms of charge by a separator magnet *5*. Depending on the polarity of the magnetic field, a Faraday cup *6* detected H^- ions or protons. After 100% stripping on a 10-μm-thick grounded Al foil *7*, the H^0 component was detected by a Faraday cup *8*. Secondary emission from the surfaces of the Faraday cups *6* and *8*, along with accompanying electrons from the aluminum foil, were efficiently suppressed by the edge magnetic field of the separator. The amplified current signals from probes *4*, *6*, and *8* were fed to an analog-to-digital converter and on to a computer for processing. The particle flux measurements were monitored periodically through their overall balance and from the equality of the signals from the probes in the absence of a target with the separator field turned on (probes *4* and *6*) and off (*4* and *8*). The fidelity of the current signals from probe *6* when the polarity of the magnetic field was changed was verified during 100% stripping of H^- ions into protons on a control target *2* consisting of a 10-μm-thick Al foil.

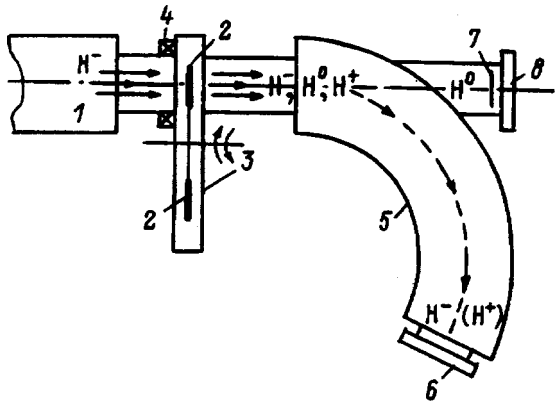

FIG. 2. The setup for the beam–foil interaction experiments.

EXPERIMENTAL DATA AND DISCUSSION

The experiments at 7 MeV showed that the stripping coefficient for H^- ions into neutral atoms on the thinnest polymer film was $\approx$0.2%. The rest of the beam was converted into protons. The operating lifetime of the film with respect to destruction by a pulsed beam current at the surface of $j_i\approx 20$ mA/cm^2, a pulse duration of $\tau_j\approx 20$ μs, and a repetition rate of $f=0.5$ Hz was $\approx$15 min. As the thickness of the films was increased, the coefficient η_0 decreased in accord with the exponential dependence of Eq. (1). In the experiments with a beam energy of 2 MeV, almost complete stripping of H^- ions into protons was achieved, while the lifetime of the films under the same current conditions was somewhat shorter. The fact that the lifetime of the polymer test films in beams with fixed characteristics increased as their thickness was raised was somewhat unexpected. An analysis of the heat loading and properties of colloxylin, as well as an examination of used targets under a microscope, showed that the reason for this (and for the rather short operating life of the films) is apparently their thermal decomposition at the places where they come into contact with the grid, which have been heated to $T_c>90°$ as a result of energy loss by the ions striking it. In this case, we might expect a substantial increase in the lifetime of these targets in relativistic beams with similar current characteristics.

The carbon foil was capable of working for several days under these experimental conditions. No structural changes were observed during an examination of the surface state of the foil under a microscope. The experiments yielded the following interaction coefficients without breakup of the H^- ions, as well as with their striping into H^0 atoms and protons on the target with a carbon foil: $\eta_-|_t=(3.8\pm0.2)\times10^{-5}$, $\eta_0|_t=(2.55\pm0.13)\times10^{-2}$, $\eta_+|_t=(9.75\pm0.50)\times10^{-1}$

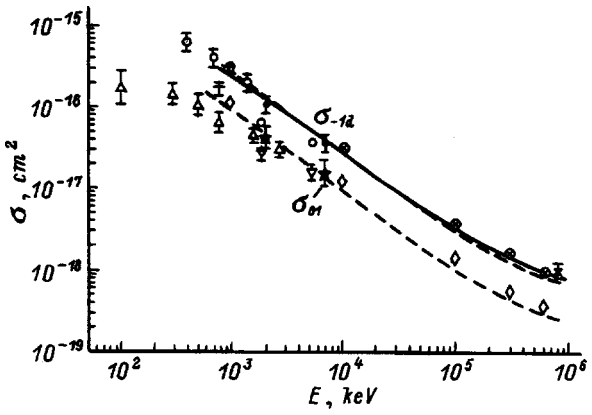

FIG. 3. Cross sections for stripping of H^- ions ($\sigma_{-1d}\rightarrow\bigcirc,\otimes,\bullet,+,*$, continuous curve) and hydrogen atoms ($\sigma_{01}\rightarrow\triangle,\triangledown,\bigstar,\diamond$) on carbon. $\bigcirc$, $\triangledown$ are the results obtained by superposition of cross sections and experimental data on the stripping of H^- ions and H(1*s*) atoms on CO_2 and O_2; $\otimes$ and $\diamond$ are theoretical results[15] for H^- ions and H(1*s*) atoms; $+$ and $*$ are cross sections obtained from beam-foil experiments[14,11]; $\bullet$ and $\bigstar$ are data from the present experiments; $\triangle$ are results obtained[18] by superposing the measured cross sections for stripping of ground state hydrogen atoms on carbon-containing gases (the errors correspond to the scatter in the data on combining different molecules); the smooth curve is constructed in accordance with Ref. 16; and, the dashed curve is an approximation of the data shown here by a curve $\propto 1/\beta^2$.

(2 MeV), and $\eta_-|_t=(1.7\pm0.1)\times10^{-2}$, $\eta_0|_t=0.28\pm0.01$, $\eta_+|_t=0.70\pm0.04$ (7 MeV). The measured transparency of the grid for unstripped ions is 0.84 ± 0.04. Given this, we obtain the relative number of particles in the different charge states following the interaction of the beam directly with the carbon foil. For beam energies of 2 and 7 MeV, these coefficients are $\eta_-=(4.50\pm0.45)\times10^{-5}$, $\eta_0=(3.0\pm0.3)\times10^{-2}$, $\eta_+=0.97\pm0.10$, and $\eta_-=(2.0\pm0.2)\times10^{-2}$, $\eta_0=0.33\pm0.03$, $\eta_+=0.65\pm0.07$, respectively. It has been shown[14] that the functional dependences of the detected charge states of the incident particles having an atomic structure on the thickness of a foil target are analogous to the characteristic dependences for gas targets up to thicknesses corresponding to the transition to the equilibrium charges. For H^- ions and H^0 atoms interacting with a carbon foil, this limiting thickness corresponds to $\eta_-\approx5\times10^{-8}$ and $\eta_0\approx10^{-3}$, respectively. In this regard, using the concept of pairwise collisions and Eq. (1), we can obtain the cross sections σ_{-1d} and σ_{01} for a carbon foil from the measured coefficients η_- and η_0. For energies of 2 and 7 MeV, these cross sections are, respectively, $\sigma_{-1d}=(1.00\pm0.16)\times10^{-16}$ cm^2, $\sigma_{01}=(4.0\pm^{1.4}_{0.8})\times10^{-17}$ cm^2, and $\sigma_{-1d}=(3.5\pm0.6)\times10^{-17}$ cm^2, $\sigma_{01}=(1.5\pm^{0.6}_{0.4})\times10^{-17}$ cm^2; these are shown in Fig. 3. Also shown there for comparison are cross sections obtained from beam–foil interaction experiments at energies of 0.8 MeV[14] and 800 MeV,[11] as well as theoretical data[15,16] and the results of calculations based on the superposition of cross sections and known experimental data[5,17,18] for gaseous carbon-containing targets. The good agreement of these data among themselves, including their approximate $\propto1/\beta^2$ dependence, indicates that it is appropriate to regard fast H^- ions in solids as independent atomic systems subject to pairwise collisions. As opposed to an earlier theoretical prediction,[19] this agreement among the data also shows that the contribution of cascade processes involving intermediate excited hydrogen atoms during beam–foil stripping of H^- ions is small.

In concluding, let us analyze the systematic errors in these results owing to neglecting the partial contribution of absorbed gases to the measured thickness of the carbon foil. The amount of absorption in various methods of fabricating carbon foils has been measured.[20] In the experiments the cross sections for elastic scattering of fast α particles ($E_\alpha=25$ MeV) at a given angle and the dependence of their final energy on the target nuclear mass were used. The results show that the main components in the absorbed gases are hydrogen ($\approx13-60\%$ in terms of numbers of atoms) and oxygen ($\approx2-5\%$), where the majority of the hydrogen atoms are chemically bound to carbon. Neglecting the contribution of absorbed gases to the energy loss by the α particles leads to a higher value for the number of carbon atoms in a determination of the target thickness. At the same time, additional stripping of H^- ions and H^0 atoms takes place on the atoms of these gases. The partial contributions of these two channels, with their different signs, determine the systematic error in the measured cross sections owing to absorption in the foil. Since the specific energy losses by α particles on atomic hydrogen are $\approx23\%$ of the losses on oxygen,[21,22] we obtain an imaginary additional thickness of $\approx3-14\%$ for a carbon target owing to the hydrogen contained in it. The particle stripping cross sections on atoms of this gas[6] are roughly a factor of ten smaller than on carbon. As a result, the systematic error owing to absorption of hydrogen is estimated to be $\approx1.5-7\%$ on the side of lowering the cross sections. Similar considerations for absorbed oxygen lead to almost complete compensation of the errors through the energy loss and stripping channels. Thus, when the systematic errors owing to the expected absorption of hydrogen and oxygen in the carbon foil are taken into account, the resulting stripping cross sections should be increased by $\leqslant7\%$. This is substantially smaller than the measurement error and has almost no effect on the results presented above.

We thank L. V. Arinin, Yu. L. Vengerov, and other colleagues for help in preparing and conducting the experiments.

†Deceased.

[1] G. I. Dimov and V. G. Dudnikov, Fiz. Plazmy **4**, 692 (1978) [Sov. J. Plasma Phys. **4**, 388 (1978)].
[2] W. S. Aaron, M. Petek, L. A. Zevenbergen, and J. R. Gibson, Nucl. Instrum. Methods Phys. Res. A **282**, 147 (1989).
[3] Isao Yamane, in *ICANS-11. International Collaboration on Advanced Neutron Sources* (KEK, Tsukuba, 1990), pp. 224–234.
[4] G. E. Adamson, M. J. Borden, R. N. Johnson, and W. F. Nicaise, Nucl. Instrum. Methods Phys. Res. A **33**, 63 (1991).
[5] Y. Nakai, T. Shirai, T. Tabata, and R. Ito, At. Data Nucl. Data Tables **37**, 69 (1987).
[6] L. W. Anderson, C. J. Anderson, L. Durand, and K. Riesselmann, Phys. Rev. A **43**, 5934 (1991).
[7] H. H. Andersen and J. F. Ziegler, *Hydrogen Stopping Powers and Ranges in All Elements* (Pergamon Press, New York, 1977).
[8] J. I. Janni, At. Data Nucl. Data Tables **27**, 147 (1982).
[9] V. P. Kovalev, *Effective Ion Charge* [in Russian], Énergoatomizdat (1991), 168 pp.

[10] Lu Hao-Lin, W. F. Sommer, and M. J. Borden, Nucl. Instrum. Methods Phys. Res. A **362**, 239 (1995).
[11] A. H. Mohagheghi, H. C. Bryant *et al.*, Phys. Rev. A **43**, 1345 (1991).
[12] M. V. Kazarnovskiĭ, N. M. Kurbangaliev *et al.*, in *Proceedings of the Tenth All-Union Conference on Charged Particle Accelerators*, Vol. 2 [in Russian], Dubna (1987), pp. 253–256.
[13] H. K. Abele, P. Glassel *et al.*, Nucl. Instrum. Methods **137**, 157 (1976).
[14] N. Cue, N. V. de Castro Faria *et al.*, Nucl. Instrum. Methods **170**, 67 (1980).
[15] A. K. Kaminskiĭ, R. A. Meshcherov, and V. S. Nikolaev, Tr. Radiotekhnicheskogo Instituta, Moscow (1973), No. 16, pp. 330–335.
[16] G. H. Gillespie, Phys. Rev. A **16**, 943 (1977).
[17] V. I. Radchenko, Zh. Éksp. Teor. Fiz. **105**, 834 (1994) [JETP **78**, 445 (1994)].
[18] L. H. Toburen, M. Y. Nakai, and R. A. Langley, Phys. Rev. **171**, 114 (1968).
[19] A. K. Kaminsky, R. A. Mescherov, and M. I. Popova, Nucl. Instrum. Methods **137**, 183 (1976).
[20] N. R. S. Tait and D. W. L. Tolfree, Nucl. Instrum. Methods **176**, 433 (1980).
[21] N. N. Pucherov, S. V. Romanovskiĭ, and T. D. Chesnokova, *Tables of the Mass Stopping Power and Mean Free Paths of Charged Particles with Energies of 1–100 MeV* [in Russian], Naukova Dumka, Kiev (1975), 295 pp.
[22] C. Hanke and J. Laursen, Nucl. Instrum. Methods **151**, 253 (1978).

Translated by D. H. McNeill

Dynamics of neutralized charged particle beams in external magnetic and self fields

S. Yu. Udovichenko

D. V. Efremov Scientific-Research Institute of Electrophysical Apparatus, 189631 St. Petersburg, Russia
(Submitted May 12, 1997; resubmitted October 13, 1997)
Zh. Tekh. Fiz. **68**, 106–109 (August 1998)

The dynamics of charged particle beams under the influence of their self-magnetic field and an external magnetic field is examined on the basis of equations for the trajectory of a boundary particle. A study is made of the change in the dynamics of fast particles due to the influence of the electric field of the partially neutralized space charge of the beam, the stationary electric field, and the field of the oscillations in the quasineutral beam plasma. Changes in the total beam energy caused by the self-electric field and in the longitudinal velocity owing to the self-magnetic field are taken into account.
[S1063-7842(98)01808-X]

INTRODUCTION

Ion and electron beams obtained from plasma sources are transported in a residual gaseous medium. Residual gas leaks into the drift space of the beam from the source itself. During shaping of high-current charged particle beams, gas is injected forcibly into the transport channels in order to reduce the effect of space charge on the dynamics of the fast particles. As a result of the ionization of the gas atoms by the beam particles, secondary charged particles accumulate in the transport channel. At low gas pressures such that $A=\nu_l r/2v_s \ll 1$, the density of the plasma particles is of the order of the density of the beam particles.[1,2] The quantity $\nu_l r$ determines the rate of ionization of the medium, where $\nu_l = n_g \sigma_i \dot{z}$, with n_g being the density of gas atoms, σ_i the cross section for ionization of a gas atom by a beam particle, and $\dot{z}=dz/dt$ the velocity of the fast particles, and r is the instantaneous radius of the beam; $v_s=(T_e/m_i)^{1/2}$, which equals the ion sound speed (T_e is the plasma electron temperature and m_i is the plasma ion mass), determines the rate at which the collisionless plasma is lost to the walls. In the opposite limit of $A \gg 1$, the plasma density can be much higher than the beam particle density. In both cases the space charge of the beam is fully neutralized. The stationary electric field of a quasineutral beam plasma can have a significant effect on the dynamics of charged particle beams.[3] This conclusion relies on the results of a numerical simulation and has not been substantiated by analytical calculations.

An electric field develops in a quasineutral beam plasma if the time for neutralizing the beam charge and the time for ambipolar drift of the plasma components perpendicular to the beam to the vessel walls are shorter than the beam pulse duration, $\tau_k + \tau_a < \tau_{\text{puls}}$. High current beams of charged particles have short pulse lengths. If the characteristic lifetime of the secondary particles exceeds the beam pulse duration ($\tau_a > \tau_{\text{puls}}$), then the plasma component which neutralizes the charge of the beam can be regarded as motionless. Then the resulting electric field of the beam–plasma system equals zero. If, however, the beam neutralization time $\tau_k = 1/n_g \sigma_i \dot{z} > \tau_{\text{puls}}$, then the beam will be partially charge-neutralized. Thus, the self-electric field of a beam propagating in a gaseous medium may be the field of the incompletely neutralized space charge or the field of the quasineutral beam plasma.

Along with the self-electric field, the self-magnetic field has a significant effect on the beam dynamics, causing the beam to pinch and ultimately limiting transport. To prevent pinching of the beam, an external longitudinal magnetic field is used.[4] Transport of high-current, short-pulse beams in an external magnetic field and the self-fields has been examined elsewhere.[5,6] References 4 and 6, however, deal with the special case of a completely charge-neutralized beam, when there is no self-electric field, and in Ref. 5 the change in the total beam energy owing to the action of the self-electric field of partially neutralized charges is neglected, along with the change in the longitudinal velocity of the beam owing to the self-magnetic field.

It is known that in long pulse beams ($\tau_{\text{puls}} > \tau_a, \tau_{\text{osc}}$, where τ_{osc} is the period of the oscillations), deneutralization of the space charge is caused by developed low- and high-frequency plasma oscillations.[7] The possibilities for improving the transport of beams with increasing space charge along the propagation direction remain unexamined.

The purpose of this paper is to study the dynamics of fully and partially neutralized charged particle beams in external and self-magnetic fields. Here for short-pulse beams the changes in the total energy owing to the self-electric field of partially neutralized space charge and in the longitudinal velocity under the influence of an external magnetic field are taken into account, while for long pulse beams the effect of the self-electric field produced by the directed and oscillatory motion of the plasma components is studied.

BEAM DYNAMICS IN THE ABSENCE OF DIRECTED AND OSCILLATORY MOTION OF THE PLASMA

The equations which determine the trajectory of a boundary particle in an axially symmetric beam in an external magnetic field and the self-fields are[4,5]

$$\frac{d(\gamma\dot{r})}{dt}=\frac{k\dot{z}_0c^2}{r}\left(\frac{1-\alpha}{\dot{z}}-\frac{\dot{z}}{c^2}\right)-\frac{1}{4}\frac{r\omega_H^2}{\gamma}\left(1-\frac{r_0^4}{r^4}\right), \quad (1)$$

and

$$\frac{d(\gamma\dot{z})}{dt}=k\dot{z}_0\frac{\dot{z}}{r}, \quad (2)$$

where $\dot{r}=dr/dt$, $k=2e_\alpha I_b/m_{b\alpha}c^2\dot{z}_0$, e_α and $m_{b\alpha}$ are the charge and mass of the beam ions or electrons, I_b is the beam current, $\gamma=[1-(\dot{z}^2+\dot{r}^2+(r\dot{\theta})^2)/c^2]^{-1/2}$ is the relativistic factor, $(r\dot{\theta})^2=\omega_H^2r_0^2(R^2-1)^2/4\gamma^2R^2$, $\omega_H=e_\alpha H/m_{b\alpha}c$, $R=r/r_0$, H is the strength of the longitudinal external magnetic field, $\dot{z}$ and r_0 are the initial longitudinal velocity and radius of the beam, and α is the degree of neutralization of the space charge.

The system of Eqs. (1) and (2) has been studied[4] for the case of a completely charge-neutralized beam ($\alpha=1$), when the radial self-electric field E_r equals zero. Then $\dot{\gamma}=e_\alpha E_r\dot{r}/m_{b\alpha}c^2=0$ or $\gamma=\gamma_0$, where γ_0 is the relativistic factor for the initial beam conditions. On the right-hand side of Eq. (1) the second term is related to the force produced by the self-magnetic field $H_\theta=2I_b/cr$, and the third, to the rotational force in the external magnetic field. In Ref. 2, Eq. (1) includes the value of the self-electric field at the beam radius, $E_r=2I_b(1-\alpha)/\dot{z}$, and is solved without including Eq. (2). A model of a uniformly charged cylindrical beam, in which the secondary particles that neutralize its charge are at rest, was used.

In the general case, for $0\leqslant\alpha\leqslant1$ the solution of the system of Eqs. (1) and (2) is given by

$$\dot{z}=\dot{z}_0\frac{\gamma_0}{\gamma}\left(1+\frac{k}{\gamma_0}\ln R\right),$$

$$\gamma=\gamma_0\left(1+\frac{k}{\gamma_0}\ln R\right)^{1-\alpha}. \quad (3)$$

The solution (3) can be used to find the minimum external magnetic field for which $\dot{r}=0$ and $\dot{z}$ is minimal. The function H/H_θ of the parameter $\dot{z}$ has a minimum for $\dot{z}=\dot{z}_0\beta_0^{-1}(1-\alpha)^{1/2}$, where $\beta_0=\dot{z}_0/c$. In a fully neutralized beam ($\alpha=1$), $\dot{z}$ and $\dot{r}$ go simultaneously to zero when[4]

$$\left(\frac{H}{H_\theta}\right)_{E_r=0}=\frac{2^{1/2}\gamma_0}{k[\cosh(2\gamma_0/k)-1]^{1/2}}. \quad (4)$$

Under these conditions there are still no return motions of the beam particles and it does not pinch. Under the condition $\alpha<1$, the minimum external magnetic field, when $\dot{r}=0$, is attained for nonzero minimum $\dot{z}$. The dependence of the minimum external magnetic field on the self-field of a partially neutralized beam is given by

$$\left(\frac{H}{H_\theta}\right)_{E_r\neq0}=\frac{2^{1/2}\gamma_0}{\beta_0k[\cosh(2p)-1]^{1/2}}\left[\alpha\left(\frac{kp}{\gamma_0}+1\right)^{2(\alpha-1)}-\frac{1}{\gamma_0^2}\right],$$

$$p=\frac{\gamma_0}{k}\left[\left(\frac{1-\alpha}{\beta_0^2}\right)^{1/2\alpha}-1\right]. \quad (5)$$

In the Introduction we mentioned that in Ref. 5 the change in the longitudinal velocity of the beam owing to the self-magnetic field was neglected in a discussion of beam pinching in the absence of an external magnetic field. We shall demonstrate the necessity of including it using the example of a fully neutralized beam, where there is no self-electric field and the total beam energy is constant ($\gamma=\gamma_0$).

Finding the value of $\dot{r}(t)$ from Eq. (3) for $\alpha=1$ and taking $z=\dot{z}t$, we obtain the instantaneous radius of the focused beam as a function of path length z,

$$z=\frac{\gamma_0r_0}{k}\exp(-\gamma_0/k)\int_1^{\dot{z}/\dot{z}_0}\frac{x\exp(x/k)dx}{(1-x^2)^{1/2}}. \quad (6)$$

The distance over which returning particles arise and a helical beam current can develop is given by Eq. (6) for $\dot{z}=0$ or $R=\exp(-\gamma_0/k)$. In this case, the integral in Eq. (6) is a tabulated function of the variable k, equal to $-{}_1F_2(1;1/2,3/2;\gamma_0^2/4k^2)-(3\pi\gamma_0/4k){}_1F_2(3/2;3/2,2;\gamma_0^2/4k^2)$, where ${}_1F_2$ is the generalized hypergeometric function.

The distance at which an initially parallel beam ($\dot{r}=0$) constricts sharply to a ''point'' and the beam is ''blocked'' by pinching is given by[5] $z=r_0(\pi\gamma_0/k)^{1/2}/2$. A comparison of this expression with Eq. (6) shows that the returning-particle regime and the beam ''blocking'' associated with it sets in earlier.

In terms of this model, all the particles inside the beam move similarly to a boundary particle. This model does not take into account the redistribution of the current density over the beam cross section and, therefore, of the forces acting on the particles owing to the intersection of their trajectories.

THE INFLUENCE OF THE STATIONARY ELECTRIC FIELD OF A BEAM PLASMA ON THE FAST PARTICLE DYNAMICS

Let us consider the dynamics of long-pulse neutralized charged particle beams in the absence of an external magnetic field. Negative ion and electron beams are focused, but positive ion beams spread out under the influence of the self-electric field of the quasineutral beam plasma. The magnitude of this field at the beam radius for low gas pressures ($A\ll1$) for these types of beams are, respectively, $E_r=m_i(\sigma_in_g\dot{z})^2r/e$ (Refs. 2 and 5) and $E_r=2T_e/er$ (Ref. 2). We shall use the equation for the beam envelope[5,8]

$$\frac{d^2r}{dz^2}=\frac{\varepsilon^2}{r^3}-\frac{k}{\gamma_0r}-\frac{r\omega_H^2}{4\dot{z}^2\gamma_0^2}\left(1-\frac{r_0^4}{r^4}\right)\pm\begin{cases}N_+/r\gamma_0,\\ rN_-/\gamma_0,\end{cases} \quad (7)$$

where ε is the beam emittance, the parameter k associated with the self-magnetic field is defined in Eq. (1), $N_+=2T_e/m_{b\alpha}\dot{z}^2$ for a positive ion beam, and $N_-=m_i(\sigma_in_g)^2/m_{b\alpha}$ for negative ion and electron beams.

In order to simplify Eq. (7), we neglect the variation in the energy of the beam particles owing to the self-electric field ($\gamma=\gamma_0$).

First we shall study the dynamics of low-current beams, where the self-magnetic field can be neglected,

$$\varepsilon^2/r^2 \gg k/\gamma_0. \tag{8}$$

In the case of negative ion and electron beams in the absence of an external magnetic field, the first integral in Eq. (7) has the form

$$\left(\frac{dR}{dz_1}\right)^2 = (R^2-1)\left(\frac{a}{R^2}+1\right), \tag{9}$$

where $a=\varepsilon^2\gamma_0/r_0^4N_-$, $z_1=z(N_-/\gamma_0)^{1/2}$, $R=r/r_0$, and we have used the initial condition $(dR/dz_1)=0$ for $z_1=0$.

The second integral in Eq. (7) gives the dependence of the instantaneous radius of the focused beam on the path length z_1,

$$z_1=\frac{\pi}{2}-\arcsin\frac{1+a-2R^2}{|a-1|}. \tag{10}$$

After determining the minimum beam radius $R_{\min}=a$. We use Eq. (9) for $dR/dz_1=0$ to find the site where the beam crossover is localized,

$$z_b=\frac{\pi}{\sigma_i n_g}\left(\frac{m_{b\alpha}}{m_i}\,\gamma_0\right)^{1/2}. \tag{11}$$

The limitation on the current density in focused beams of negative ions and electrons due to the presence of the self-electric beam–plasma field is determined by the current density at the crossover,

$$j_b=\frac{I_b r_0^2 N_-}{\pi\varepsilon^2\gamma_0}. \tag{12}$$

The spreading of a neutralized positive ion beam takes place mainly through the action of the ambipolar electric field of the plasma, since for $T_e/\gamma_0 \gg T_b$, where T_b is the beam temperature, the inequality $\varepsilon^2 \ll r^2 N_-/\gamma_0$ holds. The well-known expression $\varepsilon=2^{3/2}(T_b/m_{b\alpha})^{1/2}r_0/\dot{z}$ is used for the emittance.[9] We introduce the dependence of the instantaneous radius of the spreading beam on the longitudinal coordinate in the case of a zero initial condition $(dR/dz)_0=0$,

$$z=(\pi\gamma_0/2N_+)^{1/2}r_0\,\mathrm{erfi}(\ln^{1/2}R), \tag{13}$$

where $\mathrm{erfi}(x)$ denotes the probability integral with an imaginary argument.

In studies of the dynamics of high current, long pulse beams, the self-magnetic field cannot be neglected in Eq. (7). As the inequality (8) breaks down in negative ion and electron beams, pinching takes place. In positive ion beams pinching does not set in until $N_+ \geqslant k$.

The minimum external longitudinal magnetic field necessary to prevent pinching is found in the same way as for short-pulse beams. Given that $\gamma=\gamma_0(1\pm N_\pm \ln R/\gamma_0)$ and $\dot{z}=\dot{z}_0(1+k\ln R/\gamma_0)(1+N_\pm \ln R/\gamma_0)^{-1}$, we obtain

$$\frac{H}{H_\theta}=\frac{1}{\beta_0}\left[\left(1\pm\frac{N_\pm}{k}\right)^2-\frac{1}{\gamma_0^2}\right]\left(\frac{H}{H_\theta}\right)_{E_r=0}, \tag{14}$$

where the quantity $(H/H_\theta)_{E_r=0}$ has been defined in Eq. (4).

BEAM TRANSPORT UNDER CONDITIONS OF DEVELOPED PLASMA OSCILLATIONS

In unstable beams subject to developed beam–plasma instabilities, dynamic deneutralization of the space charge occurs. A positive ion beam becomes deneutralized with respect to charge if the amplitude of the electron plasma oscillations increases to a level sufficient for capture and removal of secondary electrons with the beam. The stationary electric field rises along this kind of beam from a value corresponding to the ambipolar field of the quasineutral beam plasma to the field of a completely charge-unneutralized beam.[7] Its value at the beam radius is given by $E_r=2I_b^+\eta\sigma_i n_g \times(z-z_0^e)/\dot{z}_0 r$, where $\eta\approx 1/2$ is the fraction of electrons trapped in the oscillations, n_g is the constant gas density along the beam, and z_0^e is the distance over which the high-frequency oscillations are suppressed nonlinearly by the capture of plasma electrons.

A negative ion or electron beam becomes deneutralized as a result of the trapping of slow neutralizing ions in ion plasma oscillations and their removal from the beam. There is a known analytical expression for the distribution of the stationary electric field in partially neutralized beams of this sort, where the plasma ions are removed along the beam.[7] The magnitude of this field at the beam radius is $E_r=-2I_b^-\sigma_i n_g\gamma_i(z-z_0^+)/\omega_{pi}r$, where γ_i and ω_{pi} are the growth rate and frequency of the ion plasma oscillations and z_0^+ is the distance over which the slow ions are trapped in the oscillations. In both cases the electric force is defocusing.

If the beam is placed in an external magnetic field $H=H_0(z/z_0^{e,+}-1)^{1/2}$, the square of which increases along the z axis in the same way as the self-electric field, then it is possible to obtain the equilibrium size of a so-called Brillouin beam. An equilibrium magnetic field of this type can be created in a cylindrical solenoid with a variable density of windings carrying a current such that $n(z)=H(z)/I$, where $n(z)$ is the number of turns per unit length and I is the current in the solenoid coil. For example, in the case of a positive ion beam, the envelope equation (7), including the self-electric field under conditions of deneutralization, yields an equation for the equilibrium radius:

$$\frac{\varepsilon^2}{r_0^3}-g(z)r_0+\frac{k}{\gamma_0 r_0}\left[\frac{\eta\sigma_i n_g}{\beta_0^2}(z-z_0^e)-1\right]=0, \tag{15}$$

where $g(z)=\omega_H^2(z)/4\dot{z}^2\gamma_0^2$ is a parameter of the channel with variable focusing.

The solution (15) gives an equilibrium radius of the beam of $r_0=\varepsilon(\gamma_0/k)^{1/2}$ and an equilibrium magnetic field of $H_0=4(\eta\sigma_i n_g z_0^e)^{1/2}I_b^+/\varepsilon\dot{z}_0$.

In order to obtain an unperturbed Brillouin flux it is necessary to match the beam with a magnetic focusing lens, i.e., at the entrance to the solenoid, to ensure simultaneously the required beam size and the correct orientation of the phase ellipse.

CONCLUSION

In this paper the equations for the trajectory of boundary particles have been used to study analytically the influence of

the electric field of the partially neutralized space charge and the field of the quasineutral beam plasma on the dynamics of negative ion, positive ion, and electron beams. Here we have included the effect of the self- and external magnetic fields on the beam. The magnitude of the stationary electric field of a quasineutral plasma was taken from Ref. 2, and its magnitude in an unstable deneutralized beam, from Ref. 6.

We have investigated the dynamics of short-pulse, neutralized beams in which there is no directed or oscillatory motion of the plasma. We have determined the magnitude of the minimum external magnetic field required to prevent beam pinching with allowance for the change in the beam energy owing to the electric field of the partially neutralized charge.

We have shown that in a fully neutralized beam in the absence of self-electric and external magnetic fields, the change in the longitudinal velocity owing to the self-magnetic field of the beam must be taken into account. The distance over which return motion of the fast particles sets in and ''blocking'' of the beam becomes possible is shorter than the distance over which the beam is pinched.

We have investigated the dynamics of long-pulse beams, for which it is necessary to include the directed and oscillatory motion of the plasma. We have found the location of the crossover of low-current precision beams of negative ions and electrons in the absence of an external magnetic field. We have determined the maximum current density for beams of this type in the presence of the self-electric field of the quasineutral beam plasma. We have ascertained the way in which a neutralized positive ion beam spreads under the influence of the self-electric field.

We have found the minimum external magnetic field required to prevent pinching of a long-pulse, neutralized charged particle beam.

We have proposed that the transport of unstable neutralized beams can be improved. An equilibrium state of a beam with an increasing space charge along the propagation direction can be realized in the rising magnetic field of a suitable solenoid. This prevents spreading of the beam by the developed oscillations of the neutralizing plasma.

The results obtained here can be used to solve problems in the transport of charged particle beams in gaseous media associated with obtaining a stationary state in beam-plasma systems, and with equilibrium configurations of beams in external and self-intrinsic fields and their stability. Problems of this sort arise during the shaping of beams in various injector systems.

[1] E. B. Hooper, O. A. Andersen, and P. A. Willman, Phys. Fluids **22**, 2334 (1979).
[2] S. Yu. Udovichenko, Zh. Tekh. Fiz. **65**, 31 (1995) [Tech. Phys. **40**, 307 (1995)].
[3] V. P. Sidorov, S. Yu. Udovichenko, A. M. Astapkovich *et al.*, in *Proceedings of the Symposium on the Production and Neutralization of Negative Ions and Beams*, Brookhaven, USA (1989), pp. 614–628.
[4] M. I. Avramenko, V. S. Kuznetsov, and R. P. Fidel'skaya, Vopr. At. Nauk Tekh. Ser. Élektrofiz. Apparat., No. 26, 66 (1993).
[5] I. N. Meshkov, *Transport of Charged Particle Beams* [in Russian], Nauka, Novosibirsk, 1991.
[6] R. R. Kikvidze *et al.*, Fiz. Plazmy **10**, 976 (1984) [Sov. J. Plasma Phys. **10**, 559 (1984)].
[7] S. Yu. Udovichenko, Zh. Tekh. Fiz. **64**, 104 (1994) [Tech. Phys. **39**, 802 (1994)].
[8] I. M. Kapchinksiĭ, *Particle Dynamics in Linear Resonance Accelerators* [in Russian], Atomizdat, Moscow, 1966.
[9] V. A. Teplyakov and V. I. Derbilov, Prib. Éksp. Tekh. No. 1, 21 (1969).

Translated by D. H. McNeill

Surface magnetic anisotropy effects in the elastoexchange spin dynamics of thin magnetic films

S. V. Tarasenko

Donetsk Physicotechnical Institute, Academy of Sciences of Ukraine, 340114 Donetsk, Ukraine
(Submitted August 2, 1996; resubmitted July 21, 1997)
Zh. Tekh. Fiz. **68**, 110–115 (August 1998)

Anomalies arising in the elastoexchange spin dynamics of a bounded magnet when exchange boundary conditions are imposed are studied for the example of a uniformly magnetized film of a rhombic antiferromagnetic. © *1998 American Institute of Physics.* [S1063-7842(98)01908-4]

It has recently been shown[1] that if the frequency ω and wave vector $k_\parallel$ of spin waves propagating in the plane of a magnetic film satisfy the elastostatic criterion

$$\omega^2 \ll s^2 k_\parallel^2, \tag{1}$$

then for these ω and $k_\parallel$ the resonance properties of a bounded magnetically ordered crystal are determined by a coupled system of dynamic equations consisting of the Landau–Lifshitz equations for the sublattice magnetization vectors and the equations of electrostatics[2] for the lattice displacement vector **u**. This sort of ''elastostatic'' approach to analyzing the magnetoelastic boundary value problem has made it possible to study analytically a wide range of questions involving the influence of the lattice on the low-frequency spin dynamics in bounded magnetoelastic crystals. The physical mechanism which creates the anomalies in the resonance properties of a bounded magnetic sample when conditions (1) are satisfied is indirect spin–spin exchange through the long-range field of quasistatic magnetoelastic strains. Even when inhomogeneous exchange is neglected (the approximation of zero exchange), this type of spin–spin interaction in bounded magnetic materials leads to the formation of a new class of zero-exchange spin-wave excitations, namely elastostatic spin waves. This class of spin oscillations forms part of the overall spectrum of magnetoelastic waves in a bounded magnetic material, just as magnetostatic waves are part of the spectrum of coupled spin–electromagnetic oscillations in magnets of finite size. For a given magnitude and orientation of the wave vector **k**, when this phonon mechanism for spin–spin exchange is taken into account simultaneously with the inhomogeneous exchange interaction and conditions (1) are satisfied, the result is a major readjustment of the spectrum of both the exchange and elastostatic waves, so that one can speak of elastoexchange spin-wave dynamics in thin magnetic films. In the case of a thin film of rhombic antiferromagnet[1] these elastoexchange anomalies in the spectrum of propagating spin waves include: 1) the possibility of forming a minimum in the dispersion curve for a traveling exchange spin wave; 2) for certain $k_\parallel$ the existence of conditions for a inhomogeneous spin–spin resonance at points where the spectra of propagating exchange and elastostatic volume spin waves are degenerate; and, 3) the possibility of forming, in a model of a semibounded rhombic antiferromagnet for $k_\parallel \neq 0$, a bipartite elastoexchange bound spin-wave state, which for $k \ll k_c$ represents a quasisurface[3] (generalized surface[4]) spin wave which transforms smoothly at $k = k_c$ into a purely surface bipartite spin wave. However, all the analyses of the elastoexchange spin dynamics of thin antiferromagnetic films done up to now have been for completely free magnetic spins on the crystal surface (Amenta–Rado boundary conditions[5]). At the same time, it follows from a study of the spectra of magnetic-dipole spin waves[3] that a surface magnetic anisotropy can have a significant effect on the structure of the magnetic excitations in a thin magnetic film as compared to the case of free spins.

In this regard, the purpose of the present paper is to study theoretically the influence of surface magnetic anisotropy on the elastoexchange dynamics of thin antiferromagnetic films, since it is known that exchange enhancement of magnetoelastic effects and exchange attenuation of magnetic-dipole effects occur simultaneously in the magnetic spectra of antiferromagnetic crystals. As a result, when this condition is satisfied, the spin dynamics of antiferromagnetic materials can be described on the basis of a simultaneously inclusion the Heisenberg and magnetoelastic interactions, with the dipole mechanism of spin–spin exchange neglected. As an example, we take a uniformly magnetized film of a rhombic antiferromagnet whose elastoexchange dynamics have been studied previously[1] with Rado–Amenta boundary conditions. If we consider a two-sublattice (where $\mathbf{M}_{1,2}$ are the magnetizations of the sublattices, with $|\mathbf{M}_1| = |\mathbf{M}_2| = M_0$) model of an exchange-collinear antiferromagnetic crystal and introduce the ferro- and antiferromagnetism vectors (**m** and **l**), then in the case of sufficiently weak (compared to the exchange) magnetic fields H, the following approximation will be satisfied with good accuracy:

$$|\mathbf{m}| \ll |\mathbf{l}|, \quad \mathbf{m} = \frac{\mathbf{M}_1 + \mathbf{M}_2}{2M_0}, \quad \mathbf{l} = \frac{\mathbf{M}_1 - \mathbf{M}_2}{2M_0}. \tag{2}$$

Given this circumstance, the energy density W of a rhombic antiferromagnet, which describes the interaction of the magnetic and elastic subsystems of the crystal, can be written in the form[1]

$$W=W_m+W_{me}+W_e,$$

$$W_m=M_0^2\left\{\frac{\alpha}{2}(\nabla\mathbf{l})^2+\frac{\beta_z}{2}l_z^2+\frac{\beta_y}{2}l_y^2+\frac{\delta}{2}\mathbf{m}^2\right\},$$

$$W_e=\frac{1}{2}(c_{11}u_{xx}^2+c_{22}u_{yy}^2+c_{33}u_{zz}^2)+(c_{12}u_{xx}u_{yy}+c_{12}u_{xx}u_{zz}+c_{23}u_{yy}u_{zz})+2(c_{44}u_{yz}^2+c_{55}u_{zz}^2+c_{66}u_{xy}^2),$$

$$W_{me}=(p_{11}l_x^2+p_{12}l_y^2+p_{13}l_z^2)u_{xx}+(p_{21}l_x^2+p_{22}l_y^2+p_{23}l_z^2)u_{yy}+(p_{31}l_x^2+p_{32}l_y^2+p_{33}l_z^2)u_{zz}+2(p_{44}l_yl_zu_{yz}+p_{55}l_xl_zu_{xz}+p_{66}l_xl_yu_{xy}), \quad (3)$$

where

$$u_{ik}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_k}+\frac{\partial u_k}{\partial x_i}\right)$$

is the strain tensor, $c_{\alpha\beta}$ and $p_{\alpha\beta}$ are the elastic constants and magnetoelastic interaction constants, respectively, $\beta_{y,z}$ are the magnetic anisotropy constants (including renormalization owing to the magnetoelastic interaction[1]), δ and α are the homogeneous and inhomogeneous exchange constants, respectively, $c^2=g^2\alpha\delta M_0^2/2$ is the propagation speed of the spin waves in the unbounded magnetic material, and g is the gyromagnetic ratio.

As in Ref. 1, in the following we shall assume, without loss of generality, that the following relations exist among the magnetic anisotropy constants ($\tilde{\beta}_{y,z}$) calculated with the magnetostriction deformations in the ground state taken into account:

$$\tilde{\beta}_z>\tilde{\beta}_y>0. \quad (4)$$

This corresponds to an equilibrium orientation of the antiferromagnetism vector along the X axis. If both surfaces of the given thin magnetic film are free of other stresses, then the system of boundary conditions which determines the spin dynamics in this model of a magnet with a normal to the surface $\mathbf{n}\perp\mathbf{l}$ is conveniently represented in the form

$$\left.\begin{aligned}\frac{\partial\tilde{l}_{y,z}}{\partial\xi}+b_{\pm}\tilde{l}_{y,z}=0\\ \sigma_{ik}n_k=0\end{aligned}\right\}\ \xi=0,d, \quad (5)$$

where $\tilde{l}_y$ ($\tilde{l}_z$) describes the amplitude of small oscillations of the antiferromagnetism vector $\mathbf{l}$ about the equilibrium $\mathbf{l}\|X$ along the Y (Z) axis; ξ is the coordinate along the direction of $\mathbf{n}$; $b_{\pm}$ is the surface magnetic anisotropy parameter, which characterizes the degree of pinning of the mode of magnetic oscillations with polarization $\tilde{l}_y$ ($\tilde{l}_z$) for $\xi=d(b_+)$ and for $\xi=0(b_-)$, and σ_{ik} is the elastic stress tensor.

For convenience in comparing our results with Ref. 1, in the following we restrict ourselves to analyzing the elastoexchange spin dynamics of a bounded antiferromagnetic material caused by indirect spin–spin exchange through the field of ''elastostatic'' phonons for $\mathbf{u}\perp\mathbf{k}$. Here we shall assume that the direction of the normal to the film surface $\mathbf{n}$ coincides with one of the Cartesian axes of the antiferromagnetic crystal. In this case, it can be shown that for $k\in XZ$ ($\mathbf{u}\|0Y$) when $\mathbf{n}\|0X$ or $k\in XY$ ($\mathbf{u}\|0Z$) when $\mathbf{n}\|0X$, the standard method for solving the Sturm–Liouville problem for this system of boundary conditions reduces to the following dispersion relation which governs the elastoexchange dynamics of a thin film of a rhombic antiferromagnet with an arbitrary degree of pinning of the magnetic moments on the film surface ($t_i\equiv\tanh q_id$; $i=1,2$; $R_i\equiv1/(k_\|^2-\gamma q_i^2)$):

$$Ab_+b_-+B(b_++b_-)+C=0,$$

$$A=2(R_2-R_1)^2t_1t_2,$$

$$B=(R_2-R_1)[q_2R_1t_1-q_1R_2t_2+t_1t_2(q_2R_1t_2-q_1R_2t_1)], \quad (6)$$

$$C=2[R_2q_1t_1-q_2R_1t_2][-q_2R_1t_1+q_1R_2t_2],$$

$$\omega^2=\omega_0^2+c^2(k_\|^2-q^2)+\frac{\omega_{me}^2k_\|^2}{k_\|^2-\gamma q^2}. \quad (7)$$

For $k\in XZ$ and $\mathbf{u}\|0Y$ we have $\omega_0^2=c^2\beta_y/\alpha$, $\gamma\equiv c_{66}/c_{44}$, $\omega_{me}^2\equiv g^2c^2p_{66}^2M_0^2/(c_{66}\alpha)$; for $k\in XY$ and $\mathbf{u}\|0Z$ we have $\omega_0^2=c^2\beta_z/\alpha$, $\gamma\equiv c_{55}/c_{44}$, $\omega_{me}^2\equiv g^2c^2p_{55}^2M_0^2/(c_{55}\alpha)$; here $q_{1,2}(\omega,k_\|)$ are the roots of the characteristic equation (7), which is biquadratic in q ($q^2=-(kn)^2$) and in the elastostatic approximation determines the structure of the spin waves in this model of a magnet without taking the boundary conditions into account. It should be noted that in the limiting case of perfectly free spins on the magnet surface ($b_\pm=0$), the dispersion relation (6) is completely identical with that found in Ref. 1. The characteristic equation for the elastoexchange boundary value problem discussed here was obtained and analyzed previously in Ref. 1, where, in particular, it was shown that in the case $\mathbf{u}\perp(\mathbf{l}\|\mathbf{n})$ ($(\mathbf{l}\perp(\mathbf{n})\perp\mathbf{u}$), depending on the frequency (ω) and projection of the wave vector k on the film plane ($k_\|$), there are four fundamentally different types of propagating bipartite elastoexchange spin waves, whose structure is independent of the magnitude of the surface magnetic anisotropy but is characterized by the parameter $q^2=-(kn)^2$.

As a result, for $(\mathbf{l}\|\mathbf{n})\perp\mathbf{u}$, the following are possible ($\omega_1^2\equiv\omega^2-\omega_0^2$):

A) volume elastoexchange spin waves I ($q_{1,2}^2<0$)

$$c^2k^2(1-\gamma)+2\omega_{me}ck_\|\sqrt{\gamma}<\omega_1^2<\omega_{me}^2+c^2k_\|^2\ (ck_\|\leqslant\omega_{me}), \quad (8)$$

B) surface elastoexchange spin waves ($q_{1,2}^2>0$)

$$c^2k^2(1-\gamma)+2\omega_{me}ck_\|\sqrt{\gamma}<\omega_1^2<\omega_{me}^2+e^2k_\|^2,\quad (ck_\|\geqslant\omega_{me}),$$

$$c^2k^2(1-\gamma)-2\omega_{me}ck_\|\sqrt{\gamma}>\omega_1^2>0,$$

C) quasisurface elastoexchange spin waves

$$c^2k^2(1-\gamma)-2\omega_{me}ck_\|\sqrt{\gamma}<\omega_1^2<c^2k^2(1-\gamma)+2\omega_{me}ck_\|\sqrt{\gamma},$$

D) volume elastoexchange spin waves II ($q_1^2q_2^2<0$)

$$\omega_1^2>\omega_{me}^2+c^2k_\|^2.$$

An analysis of Eq. (8) shows that one of the characteristic features of the elastoexchange spin wave spectrum in this geometry is the presence of ''high'' and ''low'' frequency bands in which surface elastoexchange spin waves exist, separated by a region of quasisurface (generalized) elastoexchange spin waves. In order for the high and low frequency bands of these surface spin disturbances to exist simultaneously for a given value of $k_\parallel$, it is necessary that the value of $k_{\parallel *}$, determined by the equation $c^2k_\parallel^2(1-\gamma) - 2\omega_{me}ck_\parallel\sqrt{\gamma}+\omega_0^2=0$, exceed $\omega_{me}/\sqrt{\gamma}$. Depending on the relationship of the elastic anisotropy parameters (γ), the structure of these bands on the $\omega,k_\parallel$ plane may be substantially distorted,[1] but remains topologically unchanged.

Using Eqs. (6)–(8), we can now proceed to a more detailed study of the influence of surface magnetic anisotropy on the elastoexchange dynamics of rhombic antiferromagnetic films. First of all, we consider the role of surface magnetic anisotropy effects in the elastoexchange spin dynamics of a semi-infinite antiferromagnetic crystal, the dispersion relation for which, when $\mathbf{n}\|0X$ for both $k\in XZ$ and $k\in XY$, is determined by Eqs. (6) and (7) in the limit $d\to\infty$ ($b=b_+$ for $\xi>0$ and $b=b_-$ for $\xi<0$):

$$q_1^2+q_2^2+q_1q_2-\gamma k_\parallel^2=b(q_1+q_2). \tag{9}$$

It can be shown that for completely unpinned spins on the surface of a semi-infinite magnet ($b=0$), the dispersion relation (9) describes a bipartite surface elastoexchange spin wave which has been found and studied elsewhere[6,1] and whose dispersion relation can be found from Eq. (9) in the explicit form

$$\omega^2=\omega_0^2+c^2k_\parallel^2+\frac{\gamma}{2}\left\{-c^2k_\parallel^2+ck\left(\frac{4\omega_{me}^2}{\gamma}+c^2k_\parallel^2\right)^{1/2}\right\}. \tag{10}$$

If the spins are partially immobilized, then it is not possible to find an exact explicit solution of the dispersion relation (9); however, for arbitrary $k_\parallel$ a dispersion relation for the spectrum of surface elastoexchange spin waves including partial surface pinning of the spins $b<0$ ($b>0$) can be found in the approximation $\omega^2\ll\omega_0^2+\omega_{me}^2+c^2k^2$:

$$\omega^2\cong\omega_0^2+2c^2k_\parallel^2(1+\gamma)+2ck_\parallel(\gamma(\omega_{me}^2+c^2k_\parallel^2))^{1/2} -\left\{\frac{|b|c}{2}\pm\left(\frac{c^2b^2}{4}+\gamma c^2k_\parallel^2+\sqrt{\frac{\omega_{me}^2+c^2k_\parallel^2}{c^2}}\,\gamma k_\parallel^2\right)^{1/2}\right\}^2. \tag{11}$$

A joint analysis of the dispersion curves (11) and the conditions for the existence of the different types of elastoexchange spin waves (8) shows that the main effects through which the surface magnetic anisotropy affects the localization character of a propagating elastoexchange surface spin wave (10) involve relatively small magnitudes of the wave vector $k_\parallel$. In this limit, an easy-plane surface magnetic anisotropy (11) at small $k_\parallel$ leads to strong [compared to Eq. (10)] localization of the traveling spin wave near the surface, so that it becomes a purely surface wave ($q_{1,2}^2>0$), which, according to Eq. (8), lies in the low-frequency band of the surface elastoexchange spin-wave disturbances. If, on the other hand, the surface magnetic anisotropy is easy-axis in character ($b>0$), then for small $k_\parallel$ there is no localization whatever of spin-wave disturbances of this type near the surface of the magnet, and Eq. (11) in this case corresponds to a quasiuniform volume mode over the thickness of the magnetic field which coincides with the region $A1$ of bipartial elastoexchange spin waves (8). Gradually, with increasing $k_\parallel$, the effect of immobilization on the character of the dispersion curves (11) is reduced and for $k_\parallel=k_{*1}(b)$ it shifts into the region of quasisurface (generalized) elastostatic spin waves in both cases. As $k_\parallel$ increases, this localized surface spin wave, defined by Eq. (11), smoothly crosses into the band of ''high'' frequency elastostatic spin waves (8) at $k=k_{*2}(b)$. The critical values $k_{*1,2}(b)$ are defined by Eqs. (11) and (8). The boundary curve is $\omega^2=\omega_0^2+c^2k^2(1-\gamma) - 2\omega_{me}ck_\parallel\sqrt{\gamma}$ for $b<0$ and $\omega^2=\omega_0^2+c^2k^2(1-\gamma) + 2\omega_{me}ck_\parallel\sqrt{\gamma}$ for $b>0$. Thus, the dispersion relation for traveling bipartite elastoexchange surface spin waves in the case of completely free spins (10) found in Refs. 6 and 1 determines the lower (upper) boundary in Eq. (9) for the existence of localized elastoexchange spin waves near the surface in the case of an easy-axis (easy-plane) surface magnetic anisotropy. An analysis of Eqs. (10) and (11) shows that the magnitude of the easy-plane anisotropy constant cannot exceed some critical value, since, otherwise, the direction of the equilibrium magnetization would be different in the volume and near the surface of the magnet.[7]

It follows from Eqs. (6)–(9) that the results of above analysis of Eq. (9) in the limit $q_{1,2}d\gg1$ also determine the spectrum of the elastoexchange hyperbolic modes of a thin antiferromagnetic film of thickness d (Eq. (6)). In the approximation of pure exchange, the spectrum of the hyperbolic modes of this thin antiferromagnetic film in this geometry is determined by the relation

$$(q^2-b_+b_-)\mathrm{th}qd=q(b_++b_-),$$
$$q^2=(\omega_0^2+c^2k_\parallel^2-\omega^2)/c^2. \tag{12}$$

An analogous expression has been studied previously.[7,8] The conditions for the existence of and the number of real roots (surface exchange spin waves) are conveniently presented in the form of a table[9] (Table I). In order to present the results of the analysis of the spectrum of the elastoexchange hyperbolic modes of an antiferromagnetic film given by Eqs. (6)–(9) for $q_{1,2}d\gg1$ in an analogous form, we introduce some notation. Let A and B denote the modes of the elastoexchange surface spin wave (11) which are realized, respectively, for easy-plane ($b<0$) and easy-axis ($b>0$) surface magnetic anisotropy, and A_0 denote the version (10) examined in Refs. 6 and 1 of the formation of a surface elastoexchange spin wave in the case of completely free

TABLE I.

	$(b_++b_-)>0$	$(b_++b_-)=0$	$(b_++b_-)<0$
$b_+b_->0$	B,B	0	A,B
$b_+b_-=0$	A_0,B	A_0,A_0	A_0,A
$b_+b_-<0$	A,B	A,B	A,A

TABLE II.

	$(b_+ + b_-) > 0$	$(b_+ + b_-) = 0$	$(b_+ + b_-) > 0$
$b_+ b_- > 0$	0	0	1
$b_+ b_- = 0$	0	1	1
$b_+ b_- < 0$	1	1	2

spins on the magnet surface. In this case, the conditions for the existence, number, and type of possible elastoexchange hyperbolic modes in a rhombic antiferromagnetic film for $q_{1,2}^2 \gg 1$ can be listed in the form shown in Table II.

A comparison of Tables I and II shows that indirect spin–spin exchange through an elastostatic phonon field can cause fundamental changes in the structure of the spectrum of exchange hyperbolic modes with $k_\parallel \neq 0$ in a thin uniformly magnetized antiferromagnetic film.

With the aid of Eqs. (6)–(8) we now examine how the presence of a surface magnetic anisotropy affects the spectrum of the trigonometric (volume) elastoexchange spin-wave disturbances of a thin magnetic film. An analysis of Eq. (6) shows that in the limit of a large surface magnetic anisotropy ($b_\pm \to \infty$) the dispersion relations for the spectrum of the volume elastoexchange spin modes of a rhombic antiferromagnetic film can be found explicitly:

$$\omega_n^2 = \omega_0^2 + \omega_{me}^2 \frac{k_\parallel^2}{k_\parallel^2 + \gamma(\pi n/d)^2} + c^2(k_\parallel^2 + (\pi n/d)^2),$$

$$n = 1, 2, \ldots . \tag{13}$$

Equation (14) implies that indirect spin–spin exchange through an elastostatic phonon field leads to possible degeneracy of the dispersion relations for the lowest modes of the elastoexchange volume spin waves (13) for $d > c\pi n/\omega_{me}$. This degeneracy is related to the fact that for given ω and $k_\parallel$, volume and elastostatic volume spin waves with the same spin polarization can propagate independently in this magnetic film. In order to confirm this, it is sufficient to examine Eq. (13) in two limiting cases: 1) $c^2 \to 0$ and 2) $p \to 0$. In the first case (the ''zero-exchange'' approximation), Eq. (13) describes the spectrum of propagating volume elastostatic spin modes induced in the magnetic film by indirect spin–spin exchange through an ''elastostatic'' phonon field. These modes form a countably infinite set, their frequencies ω lie within a certain interval, $\omega_0^2 \leqslant \omega^2 \leqslant \omega_0^2 + \omega_{me}^2$, and for $0 < k_\parallel < \infty$ these limiting frequencies are the points where the spectrum of the volume zero-exchange spin waves ''bunches up.'' In the second case ($p \to 0$), Eq. (13) describes the spectrum of volume exchange spin waves propagating along the magnetic film in neglect of the magnetoelastic interaction. These excitations also form a countably infinite set, but now for $0 < k_\parallel < \infty$ their spectrum is limited only from below, i.e., $\omega^2 \geqslant \omega_0^2 + c^2 k_\parallel^2$. In both cases, the magnetic wave is the straight line $k_\parallel \partial\omega/\partial k_\parallel > 0$. As a result, for a given $k_\parallel$, the existence regions of the two types of volume spin-wave excitations overlap, which also shows up in the existence of degeneracy points in the spectrum of the volume elastoexchange spin wave (13). If the spins on the magnet surface are not rigidly pinned (at least on one of the surfaces), then, as Eq. (6) implies, the degeneracy in the spectrum of the elastoexchange volume spin waves (13) is lifted owing to a resonant interaction of these propagating volume spin wave modes (a inhomogeneous spin–spin resonance). Branches of the volume elastoexchange spin waves develop in the neighborhood of the resonance.[1,10] Then a frequency interval appears in which volume spin-wave excitations of this type cannot develop. (In the dipole-exchange dynamics of thin magnetic films, these gaps are referred to as ''dipole'' gaps.[9]) The spectral structure of these excitations can easily be obtained using Eqs. (6) and (13), if the quantity $1/b$ is used as the small parameter (where b is the magnitude of the surface magnetic anisotropy on that surface of the magnet on which the Kittel condition for pinning of the moments is not satisfied).

Having studied the spectrum of the elastoexchange volume spin waves by the methods of spin-wave microwave electronics, we can now evaluate the character of the surface spin pinning. Equation (13) implies that complete pinning of the magnetic moments on both surfaces of a thin magnetic film prevents conversion of the mode with number n of these volume elastoexchange spin waves into a quasisurface wave. If for $k_\parallel \to 0$ the eigenfrequency of this mode lay in the range of the elastoexchange volume spin waves I (8), then, as Eqs. (8) and (13) imply, as $k_\parallel$ increases, its dispersion curve would cross over smoothly into the region of elastoexchange volume spin waves II, with a wave vector $k_{\parallel n}^2 = \omega_{me}^2/c^2 - (\pi n/d)^2 > 0$ at the transition point. If, on the other hand, $\omega_{me}^2/c^2 - (\pi/d)^2 < 0$, then all modes of the volume spin-wave excitation spectrum will be volume elastoexchange spin waves from region II (8) for arbitrary $k_\parallel$.

The elastoexchange dynamics of thin rhombic antiferromagnetic films is qualitatively different if $\mathbf{n} \perp (\mathbf{u} \parallel \mathbf{l} \perp \mathbf{k})$. It now follows from the boundary conditions (5) that an elastoexchange spin wave propagating in the YZ plane is now not a 4-part, as in Eq. (6), but a 6-part wave. However, in this case, as well, the boundary value problem for a volume elastoexchange spin wave propagating in the YZ ($\mathbf{u} \parallel 0X$) plane can be solved exactly with the following values of the surface magnetic anisotropy parameters:

$$\frac{\partial l_z}{\partial y} = 0; \quad l_y = 0; \quad \sigma_{yx} = 0;$$

$$\xi = 0, d; \quad k \in YZ, \quad \mathbf{n} \parallel 0Y,$$

$$\frac{\partial l_y}{\partial z} = 0; \quad l_z = 0; \quad \sigma_{zx} = 0;$$

$$\xi = 0, d; \quad k \in YZ, \quad \mathbf{n} \parallel 0Z. \tag{14}$$

The corresponding dispersion relation is a biquadratic equation in the frequency ω of the spin-wave excitations and, therefore the structure of the spectrum of this boundary value problem can also be found explicitly:

$$\omega_{\pm n}^2 = \frac{P_1}{2} \pm \left\{ \left(\frac{P_1}{2} \right)^2 - P_2 \right\}^{1/2},$$

$$P_1 = ((\tilde{\omega}_{0y}^2 + \tilde{\omega}_z^2)\beta + \tilde{\omega}_{0z}^2 + \tilde{\omega}_y^2)/(1+\beta),$$

$$P_2=(\tilde{\omega}_{0y}^2\tilde{\omega}_z^2\beta+\tilde{\omega}_{0z}^2\tilde{\omega}_y^2)/(1+\beta),$$
$$\tilde{\omega}_{0y}^2\equiv\omega_{0y}^2+c^2((\pi n)^2/d^2+k_\parallel^2),\quad \tilde{\omega}_y^2\equiv\tilde{\omega}_{0y}^2+\omega_{mey}^2,$$
$$\tilde{\omega}_{0z}^2\equiv\omega_{0z}^2+c^2((\pi n)^2/d^2+k_\parallel^2),\quad \tilde{\omega}_z^2\equiv\tilde{\omega}_{0z}^2+\omega_{mez}^2,$$
$$\omega_{mey}^2\equiv g^2c^2p_{66}^2M_0^2/(c_{66}\alpha),\quad \omega_{0y}^2=g^2c^2\tilde{\beta}_y/\alpha,$$
$$\omega_{mez}^2\equiv g^2c^2p_{55}^2M_0^2/(c_{55}\alpha),\quad \omega_{0z}^2=g^2c^2\tilde{\beta}_z/\alpha,$$
$$\beta\equiv\begin{cases}c_{55}(\pi n/d)^2/c_{66}k_\parallel^2, & \mathbf{n}\parallel 0Z,\\ c_{55}k_\parallel^2/c_{66}(\pi n/d)^2, & \mathbf{n}\parallel 0Y,\end{cases}\qquad(15)$$

An analysis of Eqs. (15) in the zero-exchange approximation ($c^2\to 0$) shows that in this case the spectrum of the volume elastostatic spin waves propagating in the plane of the film is, as in the case of Eq. (13), discrete, and these modes form a countably infinite set. Now, however, these modes have four points where the spectrum bunches: ω_{1-4}, and for $0<k_\parallel<\infty$ they are grouped into two bands which do not overlap in frequency and which we shall arbitrarily refer to as ''high'' (ω_+) and ''low'' (ω_-) frequency bands. If $\tilde{\omega}_z^2>\tilde{\omega}_y^2$, then for a given orientation of the vectors **n** and **l**, the type of wave (forward or backward) will be the same for the volume elastostatic spin wave modes lying in one frequency band and the opposite for the modes lying in the other frequency band. For the same values of the crystal parameters, a change in the type of wave to the opposite in each of the bands takes place on changing the orientation of the vector **n** from one Cartesian axis to another in a plane with its normal along the X axis. If, on the other hand, $\tilde{\omega}_z^2<\tilde{\omega}_y^2$, then all the volume modes for a given crystal, in both the high and low bands of the spectrum (15), are simultaneously forward ($k_\parallel\partial\omega/\partial k_\parallel>0$) or backward ($k_\parallel\partial\omega/\partial k_\parallel<0$) waves, depending on whether the vector **n** is directed along the hard ($0Z$) or medium ($0Y$) axis of the crystal.

As to the spectrum of the volume exchange spin waves propagating in this plane for $\hat{p}\to 0$, Eq. (15) implies that it corresponds to volume spin oscillation modes which do not interact among themselves and are coupled to two mutually orthogonal polarizations of the amplitude of the spin-wave oscillations (two sublattice model for an antiferromagnetic crystal). They are all forward waves and for fixed $k_\parallel$ ($0<k_\parallel<\infty$) their existence regions are bounded only from below. Since the existence regions found for the volume elastostatic and volume spin waves can overlap in frequency for a given $k_\parallel$, given the above analysis of the dispersion relation (13), in the case of Eq. (15) we can expect degeneracy points to develop in the spectrum of the propagating elastoexchange spin waves. Now, however, there will be two types of degeneracy points in the spectrum. The first, as in the case of Eq. (13), will correspond to crossing of branches which describe the dispersion relation of modes with different numbers (n and ν), but with the same spin polarization: $\omega_{+n}=\omega_{+\nu}$ or $\omega_{-n}=\omega_{-\nu}$. The second type of degeneracy point corresponds to the crossing of dispersion curves corresponding to modes of the elastoexchange spin wave spectrum with different spin polarizations $\omega_{+n}=\omega_{-\nu}$. Naturally, a change in the surface magnetic anisotropy parameter leads, as in the case of Eq. (13), to a resonant interaction of the spin-wave modes at the degeneracy points and to the development, near them, of two types of nonuniform spin-spin resonance of the resonantly interacting modes which differ in their spin polarizations. Depending on the type of degeneracy point indicated above, the character of the dispersion curves of the elastoexchange spin waves formed as a result of the resonant interaction, will also differ. This is because, for degeneracy points of the first type, the resonantly interacting modes are of a single type, while in the case of degeneracy points of the second type, an interaction of the forward and backward spin waves is possible, which leads to significant anomalies, compared to Eqs. (6) and (13), in the behavior of the group velocity of these waves near the resonance region.

As to the question of experimentally observing these effects of surface magnetic anisotropy in the volume elastoexchange dynamics of thin magnetic films, it should be emphasized that in order to detect them, it is necessary that for a given value of $k_\parallel$, the width of a single chosen mode of the spectrum of the normal elastoexchange oscillations $\Delta\omega_n$ should satisfy the condition $\Delta\omega_n\ll|\omega_n-\omega_{n+1}|$. The spin-wave excitation modes examined in this paper are the result of hybridization, in a bounded magnetic material, of Heisenberg and ''phonon'' spin–spin exchange mechanisms. Thus, we might expect that they would be efficiently excited by the methods belonging to the physics of magnetostatic oscillations, as well as by methods traditionally used in acoustoelectronics for generating traveling volume and surface acoustic waves.

In conclusion, the author wishes to express his deep gratitude to E. P. Stefanovskiĭ, T. N. Tarasenko, and A. N. Bogdanov for their support of the idea of this paper and for fruitful discussions.

[1] A. L. Sukstanskiĭ and S. V. Tarasenko, Zh. Éksp. Teor. Fiz. **105**, 928 (1994) [JETP **78**, 498 (1994)].
[2] V. Novatskiĭ, *Theory of Elasticity* [in Russian], Mir, Moscow (1975), 872 pp.
[3] B. N. Filippov, ''Surface spin and magnetoelastic waves in ferromagnetic materials'' [in Russian], Preprint No. 80/1, IFM UnTs AN SSSR, Sverdlovsk (1980), 62 pp.
[4] M. K. Balakirev and I. A. Gilinskiĭ, *Waves in Piezoelectric Crystals* [in Russian], Nauka, Novosibirsk (1982), 239 pp.
[5] A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* [in Russian], Nauka, Moscow (1994), 462 pp.
[6] S. V. Tarasenko, Fiz. Tverd. Tela (Leningrad) **33**, 3021 (1991) [Sov. Phys. Solid State **33**, 1705 (1991)]
[7] N. M. Salanskiĭ and M. Sh. Erukhimov, *Physical Properties and Applications of Thin Magnetic Films* [in Russian], Nauka, Moscow (1975), 220 pp.
[8] R. F. Soohoo, *Magnetic Thin Films* [Harper and Row, New York (1965); Mir, Moscow (1967), 423 pp.].
[9] B. A. Kalinikos and A. N. Slavin, J. Phys. C **19**, 7013 (1986).
[10] A. L. Sukstanskiĭ and S. V. Tarasenko, Pis'ma Zh. Tekh. Fiz. **15**(4), 28 (1989) [Sov. Tech. Phys. Lett. **15**(2), 132 (1989)].

Translated by D. H. McNeill

Sulfide passivation of III–V semiconductor surfaces: role of the sulfur ionic charge and of the reaction potential of the solution

V. N. Bessolov, Yu. V. Zhilyaev, E. V. Konenkova, and M. V. Lebedev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted March 18, 1998)
Zh. Tekh. Fiz. **68**, 116–119 (August 1998)

A model is proposed for describing the effect of a solution on the electronic properties of sulfided surfaces of III–V semiconductors which treats the adsorption of sulfur in terms of a Lewis oxide–base interaction. According to this model, the density of states on a sulfided surface, which pin the Fermi level, decreases as the global hardness of the electron shell of the adsorbed sulfide ions is increased. The Thomas–Fermi–Dirac method is used to calculate the global hardness of sulfide ions with different charges as a function of the dielectric constant of the medium. It is shown that the hardness of a sulfur ion is greater when its charge is lower and the dielectric constant of the solvent is lower. © *1998 American Institute of Physics.* [S1063-7842(98)02008-X]

1. Because of the presence of an oxide layer, the properties of real GaAs surfaces are characterized by a high density of states which cause rigid pinning of the Fermi level near the middle of the band gap and, as a consequence, lead to a high rate of surface radiationless recombination. Processing the semiconductor in solutions of inorganic sulfides removes the oxide layer and causes formation of a thin sulfide coating, which substantially reduces the density of surface states, and leads, in particular, to an increase in the luminescence quantum yield and an improvement in the characteristics of many semiconductor devices.[1]

The reduction in the density of surface states by sulfiding is attributed to the formation of an ordered layer of sulfur adatoms on the surface. Theoretical predictions of the electronic structure of a sulfided surface are based on the choice of a suitable reconstruction, optimizing the position of the atoms on the surface by minimizing the total energy, and calculating the distribution of the density of electronic states for the given atomic structure. In this way, the electronic structure of a sulfided GaAs(100) surface with a (1×1) reconstruction has been calculated,[2] as well as the distribution of electronic states on a sulfided GaAs(100) surface with the (2×1) reconstruction, more often observed experimentally.[3,4]

On the other hand, the electronic properties of the surface are modified during the chemical interaction of the ions in the solution with the surface atoms of the semiconductor. Thus, those factors which affect the chemical processes should have an influence on the electronic properties of the sulfided surface. In particular, the electronic properties of a sulfided GaAs surface can depend strongly on the composition of the passivating solution. For example, using spirits with a low dielectric constants as solvents has made it possible to increase the efficiency of passivation, which showed up as an increase in the photoluminescence intensity[5,6] and a decrease in the bending of the GaAs bands near the surface,[7] as well as an improvement in the characteristics of high-power semiconductor lasers.[8] These data suggest that the properties of a sulfur ion in a liquid dielectric medium will differ from those of a sulfur atom, whose adsorption is usually considered in calculations of electronic structure[2–4] and this difference should affect the distribution of electronic states on a sulfided surface.

In this paper we propose a model which explains the effect of a dielectric medium on the electronic properties of sulfided III–V semiconductor surfaces.

2. The chemical reaction leading to adsorption of an atom on a semiconductor surface is accompanied by charge transfer and, therefore, a change in the electronic structure of the surface. From a chemical point of view, the properties of the electronic system can be characterized by the electronegativity[9] χ (the electrochemical potential μ) and chemical hardness η,[10] which characterizes the change in the electrochemical potential of a system as the number of electrons in it changes. In terms of the theory of the density functional,[11] the electronegativity and hardness are derivatives of the total energy E of the electrons in the system with respect to the number N of electrons in the system for a fixed external potential ν, i.e.,

$$\chi=-\mu=-\left(\frac{\partial E}{\partial N}\right)_{\nu}, \tag{1}$$

and

$$\eta=\left(\frac{\partial^2 E}{\partial N^2}\right)_{\nu}=\left(\frac{\partial \mu}{\partial N}\right)_{\nu}. \tag{2}$$

The electronegativity of a semiconductor surface can be identified with the work function of the semiconductor (i.e., with the Fermi level relative to the vacuum level).[12,13]

The hardness of an ideal semiconductor surface (i.e., a surface in which the highest filled and lowest vacant electronic states are separated by a band gap) equals the width of its band gap.[14] The actual surfaces of many semiconductors, especially GaAs, is metallic, since there are a large number

of states in the band gap. The hardness of such a surface can be expressed[15] in terms of the density ρ_{ss} of surface states pinning the Fermi level,

$$\eta_{\text{surf}} \approx 1/2\rho_{ss}. \tag{3}$$

In accordance with the principle of hard and soft acids and bases,[16] only those atoms whose stiffness is roughly equal to the hardness of the semiconductor surface can react with the semiconductor surface.

The hardness of an atom or ion can be defined[17] in terms of the local hardness $\eta(\mathbf{r})$ and the Fukui function $f(\mathbf{r})$, which characterizes the reactivity of the region of an atom at position $\mathbf{r}$,

$$\eta_{\text{at}} = \int \eta(\mathbf{r}) f(\mathbf{r}) dr, \tag{4}$$

where the integral is taken over all space.

According to the principle of maximum hardness,[18] when an atom or ion is adsorbed on a semiconductor surface, the hardness of the semiconductor–adatom system should be higher than the initial hardness of either the semiconductor surface or the atom. This means that adsorption of an atom on the surface of a semiconductor with a metallic electronic structure should reduce the density of surface states that pin the Fermi level, while adsorption on a surface with a semiconductor electronic structure should lead to an increase in the surface band gap.

If the hardness of an atom or ion exceeds the hardness of the semiconductor surface, then the electronic structure of a semiconductor surface coated with adatoms will be determined by the hardness of the adatoms. Since the hardness of an atom or ion is determined exclusively by the distribution of the electron density and the electronic potential,[17] it is possible, thereby, to change the electronic structure of a semiconductor surface by acting on the electron cloud of an atom or ion prior to adsorption.

One of the methods for changing the distribution of the electron density in an ion is to place it in a dielectric medium (e.g., a liquid solvent). The ionic charge polarizes the solvent and this gives rise to an additional potential which acts on the electron cloud of the ion (the so-called reaction potential[19]). The existence of a reaction potential leads to a change in the hardness of the ion.[20]

3. In order to clarify the degree of influence of the charge of the ion and solvent, some numerical estimates of the global hardness of sulfur atoms and negative ions have been made. According to Eq. (4), for calculating the global hardness of an atom or ion, one must calculate the local hardness $\eta(\mathbf{r})$ and the Fukui functional $f(\mathbf{r})$, which are functionals of the electron density $\rho(\mathbf{r})$. The electron density $\rho(\mathbf{r})$ was calculated using the Thomas–Fermi–Dirac method.[21] The Fukui function was determined with the aid of a gradient expansion[22]

$$f(\mathbf{r}) = \frac{\rho(\mathbf{r})}{N} + \frac{\alpha}{N}\rho_0^{-2/3} \times \left\{ \left[\left(\frac{\rho_0}{\rho} \right)^{2/3} - 1 \right] \nabla^2 \rho - \frac{2}{3} \left(\frac{\rho_0}{\rho} \right)^{2/3} \frac{\nabla\rho\nabla\rho}{\rho} \right\}, \tag{5}$$

TABLE I. Global hardness of neutral atoms in the second period of the periodic table.

Atom	Hardness, calculated in Ref. 22, eV	Hardness, calculated using Eq. (4), eV
Na	8.25	7.99
Al	8.45	9.42
Si	9.42	10.38
P	10.41	10.78
S	11.38	11.40
Cl	12.35	11.62

where ρ_0 is the electron density in the core and α is a parameter which depends on ρ_0 and the atomic number.

The local hardness of an atom and ion in the absence of a solvent were calculated by analogy with an earlier paper.[22] Here the contribution of the derivatives of the kinetic energy of the electrons, of the classical Coulomb repulsion of the electrons, and of the exchange and correlation energies, which are functionals of the energy, were taken into account. The accuracy of the method was estimated by calculating the hardness of neutral atoms in the second period of the periodic system.

A comparison of the calculated global hardnesse of various neutral atoms in the absence of a solvent, η_{at}^0, with Hartree–Fock calculations[22] (see Table I) shows that the Thomas–Fermi–Dirac method is suitable for calculating the hardness of atoms in the second period. A calculation of the hardness S^{q-} of the sulfur ion as a function of charge q showed that the hardness of an ion is greater if its charge is lower (Fig. 1).

The effect of the solvent was taken into account using an induced charge model.[19] According to this model, an atom or ion is assumed to be in a dielectric continuum inside a spherical cavity of radius slightly greater than the radius of the atom or ion. Outside this cavity the dielectric constant ε equals that of the solvent, and inside it, $\varepsilon = 1$. When a charged particle is placed inside the cavity, a charge develops at its boundary with a surface density σ that depends on ε. If we assume that the ion is spherically symmetric (which is entirely allowable for monatomic ions) and the ion itself is at the center of the cavity, then we can assume that σ is

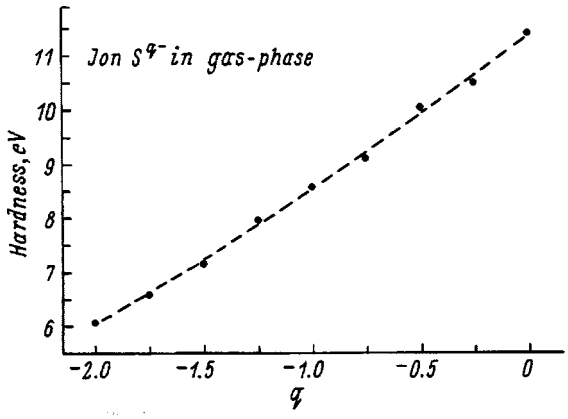

FIG. 1. Global hardness of sulfur ions in a medium with dielectric constant $\varepsilon = 1$ as a function of ionic charge q.

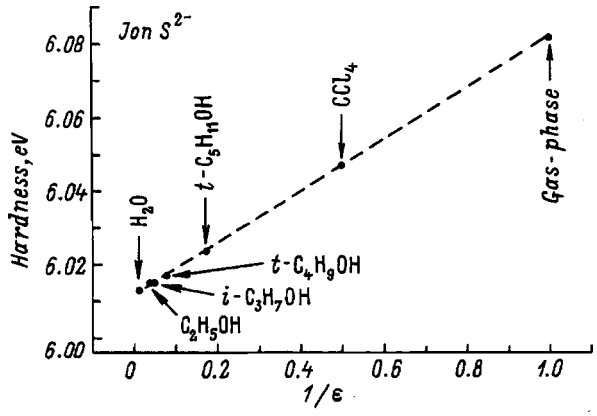

FIG. 2. Global hardness of the S^{2-} sulfide ion as a function of the reciprocal of the solvent dielectric constant.

constant on the entire cavity surface. The presence of a surface charge on the cavity surface causes an additional reaction potential Φ_σ, which also depends on ε, to develop inside the cavity,

$$\Phi_\sigma = \int_\Sigma \frac{\sigma(s)}{|\mathbf{r}-\mathbf{s}|}\, d\mathbf{s}, \tag{6}$$

where Σ is the surface area of the cavity and $\mathbf{s}$ is a vector defining a point on the cavity surface.

When a spherical atom lies at the center of the spherical cavity and, accordingly, σ is constant, the reaction potential Φ_σ will be independent of $\mathbf{r}$ inside the cavity,[23] but will depend only on ε and the ionic charge. Here, this potential will not affect the density distribution $\rho(\mathbf{r})$ in the atom or ion.

One of the components of the local hardness is the potential Φ_e of the electrons of the atom or ion.[17] In order to account for the effect of the solvent one must use $\Phi_e + \Phi_\sigma$ in place of Φ_e. The global hardness of an ion in solution will then be

$$\eta^*_{\text{at}} = \eta^0_{\text{at}} + \frac{\Phi_\sigma}{2N}. \tag{7}$$

It is clear that the hardness of the S^{2-} ion decreases as the dielectric permittivity of the solvent is raised (Fig. 2).

These results indicate that as the charge of the sulfur ion is lowered, its hardness increases, i.e., the efficiency of electronic passivation of the semiconductor should become greater. In fact, when solvents with a lower dielectric constant (such as alcohols instead of water), the effective charge of the sulfur ion in the solution is reduced and at the same time the photoluminescence intensity of sulfided GaAs increases.[24]

It should be noted that the chemical bonds which develop in a solution (e.g., hydrogen bonds) will cause a change in the spatial distribution of the electron density of an ion, which, in turn, should cause a still larger change in the hardness of the sulfide ion than that estimated using the induced charge model. In particular, a proton attached to an S^{2-} ion (HS^- ion formation) causes a drop in the hardness of sulfur from 6.08 (Fig. 1) to 4.98 eV,[20] even without a solvent.

4. It is clear, therefore, from the proposed model that a reduction in the ionic charge and dielectric constant of the solvent cause a rise in the hardness of a sulfide ion in the solvent, increased hardness of the semiconductor–atom system, and, therefore, a reduction in the density of surface states which pin the Fermi level. This conclusion is in good agreement with experimental data on the sulfide passivation of GaAs surfaces.[5–8,24]

This work was supported by the State Program on ''Current Problems in Solid State Physics: Atomic Structure of Surface'' (Grant No. 97-3.22) and by the Volkswagen Stiftung (Germany).

[1] C. J. Sandroff, R. N. Nottenburg, J.-C. Bischoff, and R. Bhat, Appl. Phys. Lett. **51**, 33 (1987).
[2] T. Ohno and K. Shiraishi, Phys. Rev. B **42**, 1194 (1990).
[3] S.-F. Ren and Y.-C. Chang, Phys. Rev. B **41**, 7705 (1990).
[4] K. N. Ow and X. W. Wang, Phys. Rev. B **54**, 17661 (1996).
[5] V. N. Bessolov, A. F. Ivankov, E. V. Konenkova, and M. V. Lebedev, Pis'ma Zh. Tekh. Fiz. **21**(1), 46 (1995) [Tech. Phys. Lett. **21**(1), 20 (1995)].
[6] V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, J. Vac. Sci. Technol. B **14**, 2761 (1996).
[7] V. N. Bessolov, M. V. Lebedev, and D. R. T. Zahn, J. Appl. Phys. **82**, 2640 (1997).
[8] V. N. Bessolov, M. V. Lebedev, Yu. M. Shernyakov, and B. V. Tsarenkov, Pis'ma Zh. Tekh. Fiz. **21**(14), 53 (1995) [Tech. Phys. Lett. **21**(7), 561 (1995)].
[9] R. G. Parr, R. A. Donnelly, M. Levy, and W. E. Palke, J. Chem. Phys. **68**, 3801 (1978).
[10] R. G. Parr and R. G. Pearson, J. Am. Chem. Soc. **105**, 7512 (1983).
[11] R. G. Parr and W. Yang, Annual Rev. Phys. Chem. **46**, 701 (1995).
[12] P. Piquini, A. Fazzio, and A. Dal Pino Jr., Surf. Sci. **313**, 41 (1994).
[13] V. N. Bessolov, A. F. Ivankov, and M. V. Lebedev, J. Vac. Sci. Technol. B **13**, 1018 (1995).
[14] R. G. Pearson, Proc. Natl. Acad. Sci. USA **83**, 8440 (1986).
[15] W. Yang and R. G. Parr, Proc. Natl. Acad. Sci. USA **82**, 6723 (1985).
[16] R. G. Pearson, Uspekhi Khimii **40**, 1259 (1971).
[17] M. Berkowitz, S. K. Ghosh, and R. G. Parr, J. Am. Chem. Soc. **107**, 6811 (1985).
[18] R. G. Pearson, Acc. Chem. Res. **26**, 250 (1993).
[19] J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).
[20] J. Lipoinski and L. Komerovski, Chem. Phys. Lett. **262**, 449 (1996).
[21] P. Gombás, *Die Statistische Theorie des Atoms und Ihre Anwendungen* [Springer-Verlag, Vienna (1949); IL, Moscow (1951), 399 pp.].
[22] P. K. Chattaraj, A. Cedillo, and R. G. Parr, J. Chem. Phys. **103**, 10621 (1995).
[23] H. Jeffreys and B. Swirles, *Methods of Mathematical Physics* (Cambridge Univ. Press, 1966), 424 pp.
[24] V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, Fiz. Tverd. Tela (St. Petersburg) **39**, 60 (1997) [Phys. Solid State **39**, 54 (1997)].

Translated by D. H. McNeill

Time-of-flight measurements in a molecular beam generated from a jet of condensing carbon dioxide

S. Ya. Khmel' and R. G. Sharafutdinov

Institute of Thermal Physics, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia
(Submitted November 4, 1996; resubmitted May 5, 1997)
Zh. Tekh. Fiz. **68**, 120–124 (August 1998)

The technique is described and results are presented for time-of-flight (TOF) measurements of the number density of molecules using electron-beam fluorescence diagnostics of a cluster-containing molecular beam extracted from a jet of condensing CO_2. The possibility of using these methods to record the velocity distribution function of the molecules (monomers) in a cluster beam is substantiated. Methods for measuring the velocity of the clusters in a CO_2 cluster beam based on their fragmentation upon impact on a solid surface are proposed and implemented. The domains of applicability of these methods and their shortcomings and merits in comparison with the conventional methods are discussed. The proposed methods are used in measurements of the velocity and temperature of the gas and the velocity of the clusters in a jet of condensing CO_2. © *1998 American Institute of Physics.* [S1063-7842(98)02108-4]

INTRODUCTION

Time-of-flight (TOF) measurements of the number density of molecules are widely used to obtain data on the velocity distribution function of molecules in molecular beams, including gas-dynamic beams extracted from supersonic free jet expansions.[1] Under certain conditions[1–3] the formation of a molecular beam from a jet takes place without disturbing the jet. In this case the molecular beam carries information about the jet which is not distorted by the sampling process. Thus, by performing a TOF analysis in a molecular beam, we can reconstruct the velocity distribution function of the molecules in the jet and, in particular, find the hydrodynamic velocity and temperature of the gas.

Such measurements are performed quite widely in molecular beams extracted from single-phase jets;[1] however, several methodological difficulties arise in the case of a condensing gas. They are associated with the appearance of clusters in the molecular beam. The first and foremost problem is the detection of molecules in the presence of clusters. The conventional detection methods employed in molecular beams (ion detectors, mass spectrometers) lead to the strong fragmentation of clusters upon electron-impact ionization. As a result, it is difficult to separate the signals from monomers and clusters and, thus, to correctly perform TOF measurements. For this reason, it is difficult to perform TOF measurements for clusters of a definite size. The second problem is associated with the formation of a molecular beam from a gas jet containing clusters. While a model for the formation of a molecular beam has been devised for single-phase jets,[2,3] only a few papers devoted to this subject are known for jets of a condensing gas.[4,5]

A mass spectrometer is usually employed as the detector for TOF analyses in cluster beams, and the measurements are performed at the monomer mass.[4,6] However, due to the fragmentation of clusters, especially van der Waals clusters, the results of such measurements are distorted appreciably, as was noted by Bailey *et al*.[4] Therefore, the phenomenon of the velocity slip of clusters relative to monomers[7,8] is utilized to correctly measure the velocity distribution function of the molecules in a cluster beam extracted from a gas expansion. In this case the mass spectrometer records a bimodal signal at the monomer mass. The peak corresponding to shorter times of flight is formed by true monomers, and the second peak is formed by monomers released from clusters by fragmentation. The velocity distribution function of the monomer component and a mean cluster velocity can thus be measured. However, such measurements can be performed only on TOF systems with good resolution. The velocities of the monomers and clusters in a molecular beam extracted from a jet of condensing Ar were measured in a similar manner in Ref. 8.

In the present work an electron-beam fluorescence detector[9] was used to record the TOF signals. In practice this calls for the use of the electron-beam fluorescence technique[10] for measurements in a molecular beam. The advantage of this method over the others stems from the possibility of using optical spectrum analysis, which permits, for example, measurement of the velocity distribution functions of molecules in individual quantum states[11] or the velocity distribution functions of different molecules of identical mass (N_2 and CO) in a gas mixture.

It was shown in Ref. 12 that the fairly large CO_2 clusters in a molecular beam make a far smaller contribution to the fluorescence of the molecular beam excited by electron impact than do the monomers. This circumstance is the basis for using electron-beam fluorescence diagnostics for a CO_2 cluster beam. The purpose of the present work was to obtain data on the distribution functions of molecules and clusters in a jet of condensing carbon dioxide by performing TOF measurements in the molecular beam using the electron-beam fluorescence technique.

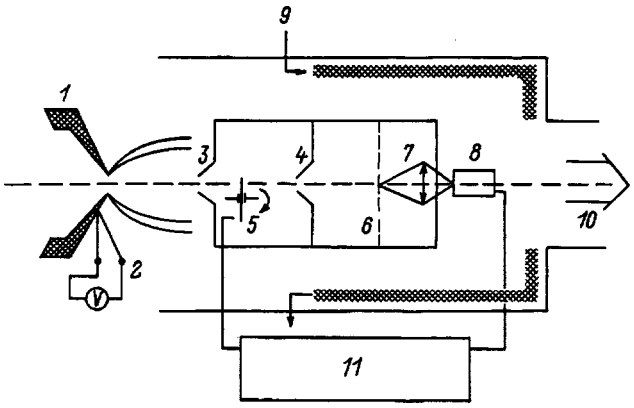

FIG. 1. Schematic representation of the experimental setup.

EXPERIMENT

The experiments were performed on the VS-4 low-density gas-dynamic facility of the Institute of Thermal Physics of the Siberian Branch of the Russian Academy of Sciences.[13] The equipment used in the present work is schematically represented in Fig. 1. Axisymmetric sonic nozzles of diameters $d_*=0.95$ and 2.11 mm served as the gas source *1*. The pressure in the forechamber of the nozzle (the stagnation pressure) was varied in the range 8–609 kPa, the stagnation temperature being held at the level of room temperature and monitored by thermocouple *2*. The gas from the source expanded into a vacuum chamber evacuated by booster pumps with a pumping speed of 35 000 liter/s and cryogenic pump *10* in liquid nitrogen *9* with a pumping speed for CO_2 up to 20 g/s. This permitted maintenance of the pressure in the vacuum chamber at the $0.1-1$ Pa level in working regimes with gas injection. The working gas, technical-grade carbon dioxide, was used without additional purification. A molecular-beam system was installed within the working chamber of the VS-4 facility for the TOF measurements. The formation of a molecular beam from the jet was effected by a skimmer *3* and a collimator *4*. The diameter of the skimmer (a conical diaphragm) was equal to 3.23 mm, and the collimator had a rectangular shape and measured 2.5×9.6 mm.

The TOF analysis was performed according to the usual scheme. A chopper *5* — a disk with two slits rotating at a frequency of 180 Hz — ''cut off'' packets of molecules, which were recorded by the detector after traversing a definite distance, which is called the time-of-flight base.

In the present work we used the electron-beam fluorescence detector that we developed[9] instead of a conventional ionization detector. It consists of electron beam *6* (diameter, $\sim 1-2$ mm; electron energy, 5.5 keV; current, ~ 20 mA; distance from the site of emergence of the beam from the electron gun to the collector, >190 mm), an optical system *7* for collecting radiation, and an FÉU-39A photomultiplier *8*. The distance between the chopper and the electron beam (the time-of-flight base line) was equal to 220 mm, and the distance between the skimmer and the chopper was 245 mm. The photomultiplier was mounted on a two-coordinate positioner, which enabled us to adjust its position and obtain transverse profiles of the molecular beam. The optical system and the photomultiplier selected ensure the detection of radiation in the spectral range $200-600$ nm. Electron-beam-activated carbon dioxide emits in this region at the wavelengths of the $\tilde{B}^2\Sigma_u^+ \Rightarrow \tilde{X}^2\Pi_g$ CO_2^+ and $\tilde{A}^2\Pi_u \Rightarrow \tilde{X}^2\Pi_g$ CO_2^+ band systems.[14]

The main difficulty arising when using electron-beam fluorescence diagnostics of a molecular beam is the low level of the net signal; the signal-to-noise ratio is usually <1. This difficulty was overcome with the aid of a signal accumulation system. An accumulation system based on a multichannel analyzer in a standard CAMAC unit controlled by an Élektronika-60 minicomputer *11* permits the accumulation of 1000 signals during 90 sec, the TOF signal being divided into 1000 time channels.

RESULTS

It was shown in Ref. 12 that when a molecular beam with CO_2 clusters of fairly large size is excited by a high-energy electron beam, the clusters make a smaller contribution to the fluorescence detected than do the monomers. Hence it follows that the TOF signal obtained using an electron-beam fluorescence detector in a CO_2 cluster beam is formed predominantly by emission from the monomers. It can be used to obtain the velocity distribution function of the molecules (monomers) in a cluster beam. Clusters do make a certain contribution to the emission of a molecular beam, but, as was shown in Ref. 12, it is small, at least for large clusters. For this reason it can be hoped that the error in the results obtained is small.

The distribution function was reconstructed from the TOF signal by solving the ill-posed inverse problem using statistical regularization.[15] This procedure was carried out on a computer using software specially written for these purposes.[16,17] The instrumental function was also obtained from the TOF signal for a model system with an *a priori* known velocity distribution function for the molecules (a gas jet of argon, carbon dioxide, etc.) by solving an ill-posed inverse problem. The reconstructed distribution functions were fairly close to the Maxwellian velocity distribution functions of the molecules. This enabled us to use them to obtain the rate of directed motion or the hydrodynamic velocity and the translational temperature, i.e., the so-called parallel temperature, of the gas in a CO_2 jet. The error in the determination of the velocity was 2%, and the error in the determination of the temperature was 40%.

As was pointed out in the Introduction, the extraction of molecular beams from jets of condensing gas remains little studied. It is clear, however, that the TOF measurements should be carried out under conditions where there are no influences from distorting factors: skimmer interference and scattering by the background gas. Therefore, we performed an investigation to determine these conditions,[5] and all the measurements performed within the scope of the present work were performed under these conditions. Because of some specific features of our system (the insufficiently high vacuum in the working chamber and the relatively large skimmer diameter), it turned out that in obtaining the depen-

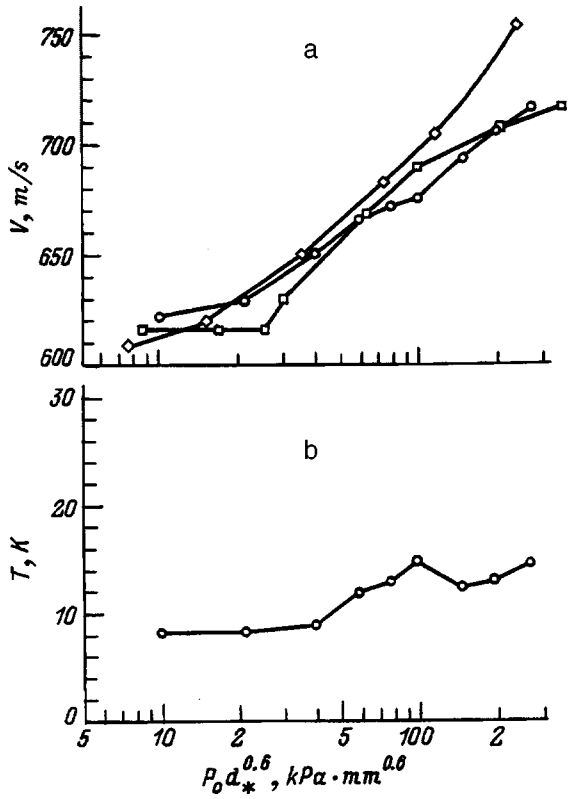

FIG. 2. Dependence of the velocity (a) and temperature (b) of the monomers in a jet of condensing carbon dioxide on $P_0 d_*^{0.6}$: ○ — present work for $d_* = 2.11$, $\Delta T_0 = 292.5 - 285.5$ K, and $x = 125 - 490$ mm; □ — Ref. 6 for $d_* = 0.147$ mm, $T_0 = 285$ K, and $x \leqslant 230$; ◇ — Ref. 4 for $d_* = 0.386$ mm, $T_0 \approx 295$ K, and $x = 386$ mm.

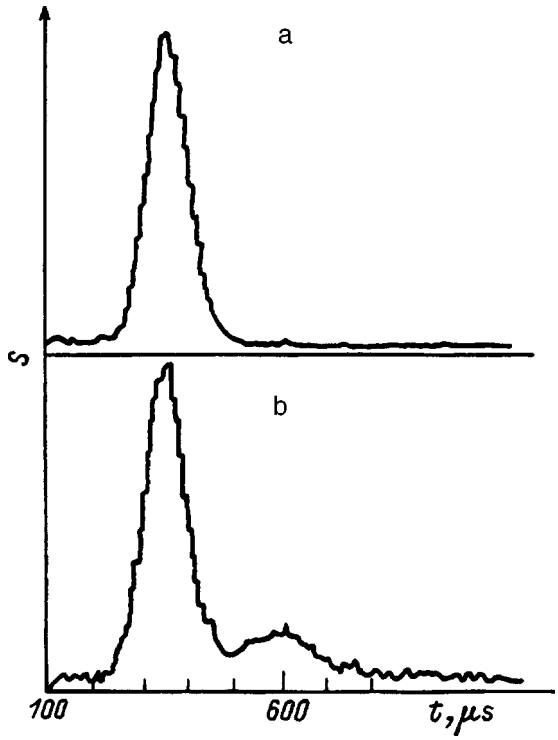

FIG. 3. Evolution of the form of a TOF signal as a function of the detector geometry: a — electron-beam–lens distance $z = 158$ mm; b — $z = 158$ mm, but with a quartz glass plate between the electron beam and the lens at a distance of 37 mm; $d_* = 0.95$ mm, $T_0 = 285$ K, $P_0 = 608$ kPa, $x = 360$ mm.

dence of the velocity distribution function of the molecules on P_0 each pressure value must correspond to a certain nozzle–skimmer distance and that the higher this pressure, the greater this distance must be.

Figure 2a shows a plot of the velocity of the monomers in a jet of condensing carbon dioxide as a function of $P_0 d_*^{0.6}$. As we know, this quantity is a similarity parameter for the condensation process,[6] and such a choice for the horizontal axis facilitates the comparison with the data of other authors below. The figure reveals that as the stagnation pressure is increased, the velocity of the monomers begins to increase after the onset of condensation and that in the regime with developed condensation it exceeds the velocity of the molecules in a jet without condensation by 15%. Time-of-flight measurements of the velocity of molecules in jets of condensing carbon dioxide have also been performed in Refs. 4 and 6; the results obtained in those studies are shown in Fig. 2a, and they are in qualitative agreement with the data from the present work. Figure 2b presents the dependence of the temperature of the monomers on the stagnation pressure. The temperature also rises with increasing stagnation pressure after the onset of condensation. However, the rise is fairly small, apparently because of the features of the measurement procedure. The fact is that the nozzle–skimmer distance must be increased as the stagnation pressure is increased in order to avoid the disturbing influence of the skimmer. However, as the distance is increased, the temperature decreases appreciably as long as the flow is not ''frozen,'' while the velocity varies only slightly.

It was shown in Refs. 12 and 18 that in the TOF system described above a barrier (a quartz glass plate) placed at a small distance after the electron beam can serve as a detector of the cluster component of the molecular beam. When the barrier is present, the TOF signal exhibits a second peak. An example is shown in Fig. 3. The mechanism for the appearance of the second peak is as follows: clusters colliding with the surface of the barrier break up into fragments, the monomers and smaller clusters formed recoil from the surface, enter the molecular beam, undergo excitation, and fluoresce, and the TOF signal becomes bimodal as a result. The second peak is clearly also present in an ordinary molecular beam, but it is very weak (at the same parameters of the recording system), and the presence of clusters sharply enhances it.

The velocity of the clusters in a molecular beam is measured using the following procedure: the bimodal TOF signal is first recorded with the barrier (Fig. 3b), then the ordinary TOF signal is recorded without the barrier (Fig. 3a), and finally the reflected TOF signal formed by the clusters is obtained by subtracting the second signal from the first. The velocity of the clusters can be estimated from the position of the maximum of this signal.[19] The following expression can be written with consideration of the data from Ref. 19 for the time T corresponding to the maximum of the TOF signal:

$$T - t/2 = (L + l)/V + l/v + \tau, \tag{1}$$

where V is the velocity of the clusters, v is the velocity of

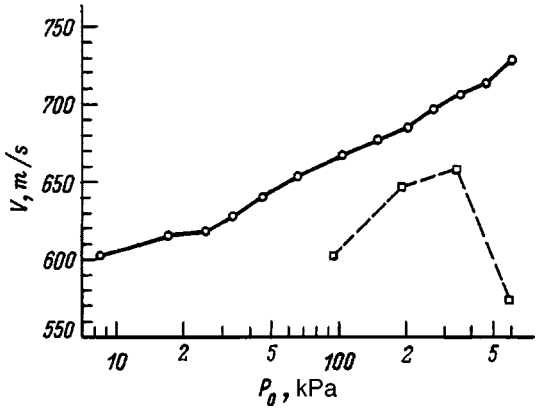

FIG. 4. Velocities of the monomers and clusters in a jet of condensing carbon dioxide versus the stagnation pressure: ○ — monomers; □ — clusters; $d_* = 0.95$ mm, $\Delta T_0 = 285-282$ K, $\Delta x = 65-360$ mm.

the cluster fragments reflected from the barrier, L is the time-of-flight base, l is the electron-beam–barrier distance, τ is the lifetime of the cluster fragments on the barrier surface, and t is the width of the instrumental function at the base.

It is known from Ref. 20, in which the scattering of CO_2 clusters in a gas-dynamic molecular beam on a solid surface was investigated, that clusters with a size smaller than a certain value fragment predominantly into monomers, the scattering has a diffuse character, and the lifetime of the fragments on the surface $\tau \approx 1\ \mu$s. The barrier is at room temperature; therefore, $v = 380$ m/s, and the remaining quantities (T, t, L, and l) are measured experimentally. According to estimates, the error in a velocity measurement is 5%.

The results of the measurement of the velocities of the clusters and monomers as a function of P_0 are shown in Fig. 4. It is seen that the velocity of the clusters is $\approx 10\%$ less than the velocity of the monomers, and velocity slip accordingly occurs. The velocities of both components increase with the pressure. These results are qualitatively consistent with the data for a jet of condensing argon.[8] For clusters only the point at the maximum stagnation pressure deviates from the general picture. This is apparently a result of a change in the character of the interaction of the clusters with the solid surface.[20,21] According to the data in Ref. 20, as the size of CO_2 clusters (and, accordingly, P_0) increases, the scattering becomes lobed instead of diffuse, and the velocity of the scattered components decreases significantly (by a factor of about 2). The pressure and thus the cluster size at which the velocity anomaly is observed in Fig. 4 are consistent with the data in Ref. 20 on lobed scattering. Accordingly, the use of Eq. (1) for determining the velocity under such conditions leads to a large error. We note that the results presented on the velocity of the cluster component have only an illustrative character primarily because of the inadequate extent to which the interaction of clusters with a surface has been studied. As for the measuring system under consideration, it can be optimized. In particular, the electron-beam–barrier distance can be reduced by about an order of magnitude with resultant expansion of the range of pressures (cluster sizes) for measuring the velocity and reduction of the error.

DISCUSSION

The method described above for measuring the velocity distribution function of monomers in a CO_2 cluster beam has one ambiguity. We have assumed on the basis of Ref. 12 that monomers make a contribution to the fluorescence of a molecular beam, while clusters of large size scarcely emit. In addition, it is known that the radiated intensity per cluster molecule decreases with increasing cluster size.[12] This means that clusters of small size (dimers, trimers, etc.) can, in principle, emit efficiently and accordingly make a contribution to the TOF signal just as monomers.

However, clusters of small size, unlike large clusters, have only an insignificant velocity slip relative to the monomers, and, in addition, under real conditions there are generally few of them relative to the total amount of the substance. In fact, a molecular beam formed from a jet with a developed condensation process (a condensate fraction $\approx 20-30\%$ and a large mean cluster size) consists practically only of clusters due to the enrichment of the beam as a consequence of the difference in the thermal expansion of monomers and clusters. If the condensation process has just begun in the jet (a condensate fraction $\approx 1\%$ and a small mean cluster size), there will be few clusters in the molecular beam, since they are considerably smaller and enrichment is relatively weak. For these reasons, the clusters of small size should not noticeably distort the results of measurements of the velocity distribution function of the molecules in a cluster beam.

This method for measuring the velocity distribution function can be employed not only for CO_2, but also for any substance whose clusters do not emit or emit inefficiently at the wavelengths of the monomers.

In the proposed method for measuring the velocity of clusters in a beam, the measurement result is the hydrodynamic velocity averaged over the size distribution function of the clusters. This method can be employed for van der Waals and any other weakly bound clusters. The alternative method, which employs a mass spectrometer, does not have such a restriction, but the measurement result is similar to the former,[8] since the same phenomenon, viz., cluster fragmentation, is utilized in both cases, although fragmentation occurs upon collisions with a solid surface in the former case and upon ionization by electron impact in the latter case.

We note that there is another possibility for measuring the velocity of clusters using an electron-beam fluorescence detector in a molecular beam, viz., direct recording of the cluster TOF signal. However, it can be realized only for clusters which emit efficiently upon excitation by electron impact and whose emission spectrum differs from the emission spectrum of the monomers.

In the present work we tested new methods for measuring the velocity distribution function of monomers and the velocity of clusters in a cluster beam. Their potential and fundamental advantage lie in the possibility of using optical spectrum analysis even for such fine measurements as obtaining the velocity distribution function of molecules in individual quantum states.

CONCLUSIONS

A method for measuring the velocity distribution function of the monomers (molecules) in CO_2 cluster beams with the aid of a TOF analysis employing an electron-beam fluorescence detector has been substantiated and implemented in the present work. A method for measuring the velocity of the clusters in a CO_2 cluster beam has been proposed and implemented. These measurement methods can significantly supplement the conventional methods employing a mass spectrometer, since the former, unlike the latter, permit the use of optical spectrum analysis. The velocity and temperature of the molecules and the velocity of the clusters in a jet of condensing CO_2 have been measured using the proposed methods. The results obtained are consistent with the published data.

We thank P. A. Skovorodko for taking an interest in this work and for some fruitful discussions, as well as S. V. Poroseva for writing the software for processing the TOF signals and assisting in processing the experimental results.

[1] *Atomic and Molecular Beam Methods*, G. Scoles (Ed.) (Oxford University Press, New York–Oxford, 1988), Vol. 1, p. 721.
[2] J. B. Anderson, R. P. Andres, and J. B. Fen, in *Molecular Beams (Advances in Chemical Physics, Vol. 10)*, J. Ross (Ed.) [Interscience, New York (1966), pp. 275–317; Mir, Moscow (1969), pp. 299–345].
[3] A. E. Zarvin and R. G. Sharafutdinov, Zh. Prikl. Mekh. Tekh. Fiz. (6), 107 (1979).
[4] A. B. Bailey, R. Dawbarn, and M. R. Busby, AIAA J. **14**, 91 (1976).
[5] A. E. Belikov, S. Ya. Khmel', and R. G. Sharafutdinov, in *Rarefied Gas Dynamics, Proceedings of the 17th International Symposium* (Aachen, Germany, 1990), pp. 1220–1226.
[6] D. Golomb, R. E. Good, A. B. Bailey *et al.*, J. Chem. Phys. **57**, 3844 (1972).
[7] E. W. Becker, K. Bier, and W. Henkes, Z. Phys. **146**, 333 (1956).
[8] J. Covellier and A. Binet, Rev. Phys. Appl. **23**, 91 (1988).
[9] A. E. Belikov, E. S. Voronel', Ya. Ya. Tomsons *et al.*, Zh. Prikl. Mekh. Tekh. Fiz. (2), 18 (1986).
[10] L. A. Gochberg, in *Proceedings of the 18th Aerospace Ground Testing Conference* (Colorado, 1994), AIAA-94-2635, pp. 1–43.
[11] A. E. Belikov, S. Ya. Khmel', and R. G. Sharafutdinov, in *Flow of Rarefied Gases with Nonequilibrium Physicochemical Processes. Proceedings of the 8th All-Union Conference on the Dynamics of Rarefied Gases* [in Russian] Moscow (1985), pp. 116–120.
[12] S. Ya. Khmel' and R. G. Sharafutdinov, Zh. Tekh. Fiz. **67** (7), 63 (1997) [Tech. Phys. **42**, 775 (1997)].
[13] A. A. Bochkarev, E. G. Velikanov, A. K. Rebrov *et al.*, in *Experimental Methods in the Dynamics of Rarefied Gases* [in Russian], Novosibirsk (1974), pp. 6–29.
[14] A. A. Bochkarev, V. A. Kosinov, A. K. Rebrov, and R. G. Sharafutdinov, *ibid*., pp. 98–137.
[15] Yu. E. Voskoboĭnikov, N. G. Preobrazhenskiĭ, and A. I. Sedel'nikov, *Mathematical Treatment of Experiments in Molecular Gas Dynamics* [in Russian], Nauka, Novosibirsk (1984), 239 pp.
[16] S. V. Poroseva, in *Abstracts of the 4th All-Union Conference of Junior Scientists on Current Topics in Thermal Physics and Physical Gas Dynamics* [in Russian], Novosibirsk (1991), pp. 59–60.
[17] S. V. Poroseva, in *Abstracts of the All-Union Conference on Ill-Posed Problems in Mathematical Physics and Analysis* [in Russian], Novosibirsk (1992), pp. 114–116.
[18] A. E. Belidov, S. Ya. Khmel', and R. G. Sharafutdinov, in *Thirteenth International Symposium on Molecular Beams, Book of Abstracts* (El Escorial, Madrid, 1991), p. A.10.
[19] W. S. Young, Rev. Sci. Instrum. **44**, 715 (1973).
[20] A. A. Vostrikov, S. G. Mironov, and B. E. Semyachkin, Zh. Tekh. Fiz. **52**, 1164 (1982) [Sov. Phys. Tech. Phys. **27**, 705 (1982)].
[21] R. J. Holland, G. O. Xu, A. Robertson *et al.*, J. Chem. Phys. **88**, 7952 (1988).

Translated by P. Shelnitz

Novel compact 60-kV Mott detector for spin-polarization electron spectroscopy

V. N. Petrov, M. S. Galaktionov, B. V. Yushenkov, and Yu. A. Mamaev

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

M. Landolt

Eidgenössische Technische Hochschule, CH-8093, Zürich, Switzerland
(Submitted June 5, 1997)
Zh. Tekh. Fiz. **68**, 125–130 (August 1998)

A compact 60-kV Mott polarimeter designed specially for the local analysis of surface and two-dimensional magnetism by spin-resolved electron spectroscopy is developed and tested. The use of a design which combines a spherical accelerating field and the absence of a retarding potential after scattering of the electron beam ensures high stability of the measured polarization even when the potential and diameter of the beam being investigated vary. As a result of optimization of the scattering angle (118°) and the use of surface-barrier detectors with a large collection angle (~48°), the efficiency or figure of merit of the polarimeter, which determines the signal-to-noise ratio $\varepsilon=(I/I_0)\cdot(S_{\text{eff}})^2$, equals 2.5×10^{-4}. Specially developed electronic circuits and optimum positioning of the detectors provide a maximum electron counting rate as high as 5×10^6 counts/s. Consequently, it is possible to calibrate the polarimeter (to find the effective Sherman function S_{eff}) by extrapolating the measured asymmetry to a high discrimination level. This instrument can also be used in other areas of solid-state physics, atomic physics, and high-energy physics.
[S1063-7842(98)02208-9]

INTRODUCTION

Measuring the spin polarization of an electron beam is still a difficult task. For the most part, the scattering of electrons on gold films is used for this purpose.[1,2] In high-voltage Mott detectors the electrons in the beam being analyzed are accelerated to an energy of $\sim100-120$ keV, and the scattered electrons are detected by two (or four) detectors arranged symmetrically relative to the beam at $\pm120°$ angles. As a result of the interaction of the spin of each electron with its orbital angular momentum, the effective scattering cross sections for electrons with opposite spins are different. In other words, spin-orbit coupling gives rise to left-right scattering asymmetry A_{LR}, which can be defined as the normalized difference between the signals of the left-hand (N_L) and right-hand (N_R) detectors:

$$A_{LR}=(N_L-N_R)/(N_L+N_R). \tag{1}$$

In this case the polarization of the beam is defined as

$$P_0=A_{LR}/S_{\text{eff}}, \tag{2}$$

where S_{eff} is the effective Sherman function, i.e., the asymmetry which should be observed for 100% electron polarization.

To obtain the absolute value of the polarization, we must determine S_{eff} by an independent measurement or utilize the results of a reliable theory[3] and extrapolate the values of the measured asymmetry in some way to the case of single elastic scattering.

Let us examine some parameters of different polarization analyzers presently in use. Only data on so-called high-voltage Mott detectors are presented here, since the other types of analyzers, such as diffraction polarimeters[4] or analyzers based on diffuse electron scattering,[5] still cannot compete with Mott analyzers. This is due to the insufficient temporal stability of these instruments. The need to periodically clean the tungsten surface by heating to high temperatures in the case of diffraction detectors or by depositing fresh layers of gold in the case of a diffuse-scattering analyzer renders the other superior parameters of these instruments worthless. It is assumed that an electron polarization analyzer, which is a unique and fairly expensive instrument, should operate for several years without adjusting its principal parameters or special servicing. These requirements are known to be satisfied only by high-voltage Mott analyzers.[6,7]

A classical high-voltage Mott polarimeter (''conventional'' polarimeter) consists of a linear accelerating column and a scattering chamber, which is under a high potential of ~100 kV and contains a target and two (or four) surface-barrier detectors located at $\pm120°$ angles relative to the axis of the instrument. As the scattered electrons move in the drift space, they are picked up by these detectors, which also perform energy selection of the electrons. The instrument has a high efficiency. The main deficiencies are: its large size, high cost, and possible instability of the measured polarization as the diameter and position of the electron beam at the entrance to the analyzer vary.[8]

The next group of instruments consists of retarding-potential polarimeters. The name is associated with the method for the energy selection of electrons after they are scattered on a target. The main advantages are the small dimensions and simplicity of design.[9,10]

Another group of instruments can be called spherical field-free polarimeters. In this case we have a variant of a conventional analyzer, in which the beam is accelerated in the space between two hemispheres and the scattered electrons also move in a drift space and undergo energy selection in a surface-barrier detector. These instruments exhibit high stability of the measured polarization as the diameter and position of the electron beam at the entrance to the analyzer are varied.[11,12]

The main parameters of various types of Mott polarimeters are presented below in order of their decreasing (in our opinion) significance. We shall comment briefly on several points.

1) Efficiency or figure of merit of the polarimeters. This parameter, which determines the signal-to-noise ratio, is calculated from the formula

$$\varepsilon = (I/I_0) \cdot (S_{\text{eff}})^2, \tag{3}$$

where I/I_0 is the fraction of the intensity of the incoming beam reaching the detectors after the electrons are scattered on the target.

The maximum figure of merit is generally exhibited by $100-120$-kV Mott detectors, but the efficiency of several models of retarding-potential polarimeters is also fairly high.

2) Sensitivity toward displacement of the electron beam being analyzed and to variation of its diameter. This parameter is extremely important, since in some experiments the characteristics of the beam (its diameter and spatial position) can undergo slight variations, which can, in turn, cause uncontrollable variations in the measured polarization. The spherical field-free polarimeters are the undisputed leaders here.

3) Dimensions of the instrument. The retarding-potential analyzers have no competitors with regard to size. Some designs permit employment of the polarimeters within a standard vacuum chamber, while conventional $100-120$-kV Mott detectors have a volume of several cubic meters.

4) Maximum counting rate. This parameter is generally not mentioned when the designs of different instruments are described. Nevertheless, it is one of the most important characteristics for solving a number of experimental problems, in which a set of a sufficient number of pulses during a short time is required to achieve a certain statistical error. This is important, in particular, in the investigation of different types of magnetic relaxation. In retarding potential analyzers the maximum counting rate is determined by the properties of the channel electron multipliers or microchannel plate detectors used. It reaches a maximum value of $\sim 10^6$ counts/s. The maximum counting rate of conventional polarimeters is 10^4-10^6 counts/s. The bound on this parameter is due the large capacitance of surface-barrier detectors. In addition, the distance between the detectors and preamplifiers is large in most designs of conventional polarimeters. The considerable capacitance of the connecting cable also leads to a decrease in the maximum electron counting rate.

5) Energy resolution of the electron detectors. In this case we are referring to the width of the energy ''window'' used to detect scattered electrons. In conventional polarimeters it is determined by the parameters of the surface-barrier-detector/preamplifier unit and amounts to ≈ 10 keV. In retarding-potential analyzers this ''window'' is assigned by the magnitude of the applied retarding potential and can reach a value of several electron-volts. As a consequence, it is possible to calibrate retarding-potential analyzers by extrapolating the measured asymmetry to zero energy losses.

6) Potential of the electron detectors and preamplifiers. In this case the retarding-potential polarimeters has definite advantages, since problems associated with the high voltage must be solved for conventional $100-120$-kV Mott detectors.

As a result of a similar analysis performed several years ago, we settled on a spherical field-free polarimeter. We then endeavored to take into account and rectify the deficiencies inherent in the instruments described. The employment of a 60-kV potential enabled us to create a fairly compact analyzer, but, at the same time, required definite efforts to optimize the dimensions of the instrument, the parameters of the detectors, and the configuration of the electronic circuits. The novel Mott polarimeter described in this paper is considerably smaller than the former instruments and has several characteristics which surpass those of the previous analyzers. The creation of special electronic circuits and determination of the optimum arrangement of the surface-barrier detectors enabled us to increase the maximum electron counting rate to 5×10^6 counts/s, as a result of which it has become possible to calibrate the polarimeter (to find the effective Sherman function S_{eff}) by extrapolating the measured asymmetry to a high discrimination level.

DESIGN OF THE INSTRUMENT

A schematic representation of the polarimeter is shown in Fig. 1. The main components of the analyzer are two polished metallic hemispheres with a common center. The outer hemisphere operates at or near ground potential. A potential of about 60 kV is supplied to the inner hemisphere. For this purpose it is secured by means of two cylindrical ceramic insulators connected in series, which were designed for a total voltage of ≈ 80 kV. The electron beam whose polarization is to be determined is directed into the region between the hemispheres through a diaphragm with a circular opening, and then, after being accelerated by the spherical field, it impinges on the inner hemisphere. Calculations of the electron trajectories show that such a field focuses the incoming beam well. Four surface-barrier detectors with a large working surface of $\sim 1.5\,\text{cm}^2$ each, a gold foil, and guiding diaphragms are installed within the small hemisphere. The optimum scattering angle chosen according to calculations is 118°, since the figure of merit ε of the polarimeter is larger in this case than for scattering at a 120° angle. The angle for the collection of electrons by each detector is 48°. The scattering film is an 800 Å thick gold layer deposited on a thin free tungsten film. After scattering, the electrons pass through holes in diaphragms and are picked up by the detectors.

In an analysis of the scattering asymmetry is analyzed, the reflected electrons must be energy-selected, since, as we know, maximum asymmetry arises in the case of elastic re-

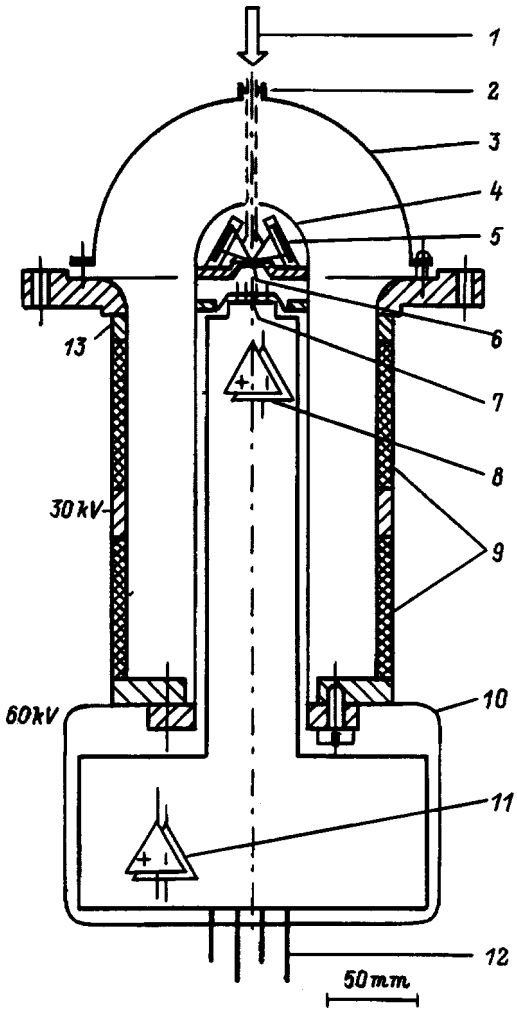

FIG. 1. Schematic representation of the instrument: *1* — incoming beam, *2* — entrance diaphragm, *3* — outer hemisphere, *4* — inner hemisphere, *5* — four surface-barrier detectors, *6* — gold foil, *7* — vacuum electric input, *8* — four preamplifiers, *9* — ceramic insulators, *10* — Duralumin housing, *11* — four discriminators and optical signal formers, *12* — fiber-optic system, *13* — ground potential.

flection. This is because information on the spin is lost as a result of inelastic scattering, which can occur before the elastic interaction event. In addition, the beam of inelastic electrons emerging at a scattering angle equal to 118° can contain electrons which preliminarily underwent elastic scattering at other angles (at high energies the elastic and inelastic scattering events can be considered independently of one another). In our polarimeter such selection is effected by the surface-barrier detectors, whose output pulses have an amplitude proportional to the energy of the incident electrons. After the amplification and discrimination of these pulses at a definite level, their repetition frequency can be assumed to be proportional to the number of elastically reflected electrons. Unfortunately, such detectors (in pairs with special charge-sensitive amplifiers) have poor energy resolution. In our case it amounts to ≈ 18 keV. This resolution is not equal to the best achievable value due to the use of surface-barrier detectors with a large working surface and, accordingly, a large capacitance. As we have already stated above, we developed and fabricated special preamplifiers, which permit operation of the detectors in a high-counting-rate regime. For the same reason the polarimeter design was developed so that the distance between the surface-barrier detectors and the preamplifiers would be minimal. The preamplifiers are mounted in a special cylindrical housing and placed within the polarimeter. The vacuum joint is located in direct proximity to each surface-barrier detector. The length of each of the connecting cables is no greater than 50 mm. The output signals from the charge-sensitive preamplifiers are fed into discriminators and shapers. All the amplifiers are under a high working potential of ≈ 60 kV. For further processing the signals are transmitted by a fiber-optic system to the input of amplifying shapers at ground potential. A special compact stabilized 70-kV power source, which measured $340 \times 280 \times 160$, was fabricated to power the analyzer and all the amplifiers.

TESTING AND CALIBRATING THE POLARIMETER

The testing of the polarimeter was carried out in several stages. Its high-voltage properties were investigated first. The criterion for the maximum potential which can be supplied to the instrument was the appearance in the counting channels of pulses associated with the detection of electrons arising from discharge as a result of high-voltage breakdown in vacuum or slight short-term degradation of the vacuum in the polarimeter chamber. A maximum value of ≈ 68 kV for the high voltage was recorded. All the tests were subsequently carried out at 60 kV.

In the next stage the surface-barrier detectors and the amplification circuit were tested. A source of unpolarized electrons in the form of a tungsten coil was used. As we have already stated above, the energy resolution of the electron detectors is determined by the surface-barrier-detector/preamplifier unit and is related mainly to the presence of internal noise in the preamplifier. Its value is fairly large for amplifiers operating with high counting rates or, more precisely, with very short pulses. In our case the sum of the pulse duration and the dead time of the preamplifier is of the order of 65 ns (the capacitance of a surface-barrier detector of large area is ≈ 60 pF, and the creation of amplifiers with such parameters is a formidable task). To analyze the energy resolution of the surface-barrier-detector/preamplifier unit, we recorded the dependence of the electron counting rate in the channels on the potential on the inner hemisphere (on the energy of the electrons). Then the curves obtained were differentiated and approximated by a Gaussian function. The full width at half-maximum of this function was taken as the energy resolution of the system. Although this method is not absolutely rigorous, we were still able to estimate the energy resolution. Its mean value was equal to ≈ 18 keV for all the detectors. This value coincides with the data obtained when the system was tested using a standard signal generator.

Then we investigated the dependences of the noise pulse counting rate without an electron beam $N_n(A)$ and the electron counting rate $N_{e+n}(A)$ on discrimination level, where A is the pulse amplitude or the discrimination level in arbitrary units (a.u.). The signal (without noise pulses) electron counting rate was determined by simple subtraction: $N_e(A)$

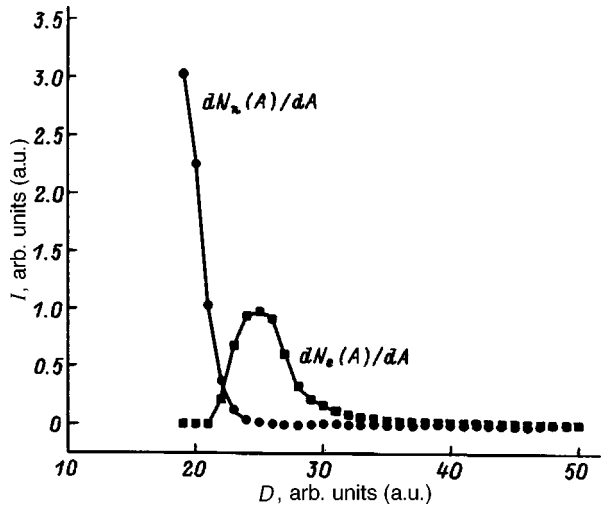

FIG. 2. Dependence of the derivatives of the signal $[dN_e(A)/dA]$ and noise $[dN_n(A)/dA]$ counting rates on discrimination level for one of the channels.

$=N_{e+n}(A)-N_n(A)$. Figure 2 presents plots of the derivatives of the signal and noise counting rates as functions of discrimination level. It is seen that the noise can be clearly separated from the real signal by properly choosing the discrimination level despite the low resolution of the detectors used.

In the next stage we tested the operation of the polarimeter and calibrated it using a beam of spin-polarized electrons with a known value of P_0. The work was carried out in the department of experimental physics of St. Petersburg State Technical University. The polarimeter was mounted in an experimental setup equipped with a source of polarized electrons based on photoelectron emission from semiconductor samples under the action of circularly polarized light. The polarization of the beam was $P_0=35\pm2\%$. To calibrate the detector, the polarization direction of the incident beam was modulated so that each channel could be investigated individually. The asymmetry was determined from formula (1), where N_L and N_R are the counting rates in the channel as the polarization direction of the incident beam is varied. The values of asymmetry obtained coincided to high accuracy in each of the channels. Figure 3 presents the measured spectral dependence of the photoemitter used in the calibration in one of the channels at the 25 a.u. discrimination level (Fig. 2). The polarization $P_0=35\pm2\%$ corresponded to a quantum energy ≈1.6 eV. The effective Sherman function amounted to ≈0.16. The fraction of the intensity of the incident beam I/I_0 reaching the detectors after scattering on a target was measured at the same discrimination level. As expected, because of the large detector working surface and the small distance between the detectors and the gold foil, it was fairly high and equal to $\sim10^{-2}$. The efficiency of the polarimeter calculated from formula (3) was $\approx2.5\times10^{-4}$. Although the improvement in the figure of merit was achieved entirely by increasing I/I_0, rather than the Sherman function, we still believe that the use of a spherical field-free polarimeter design with its low sensitivity to displacements of the electron beam being analyzed permits minimization of various instrumental asymmetries. We shall show below that the effective Sherman function of an analyzer can be varied by varying the discrimination level until the maximum possible value for the instrument geometry and electron energy under consideration, $S_{\text{eff theor}}=0.33\%$, is achieved. The counting rate in the measurements of the dependence shown in Fig. 3 was about 2×10^5 counts/s due to design features of the source of polarized electrons used in the calibration. We note that in work with a counting rate $\sim10^6$ counts/s (which is allowed by the polarimeter under consideration) the time needed to record the entire dependence would be equal to 10 s with a statistical error smaller than the diameter of a point on the graph.

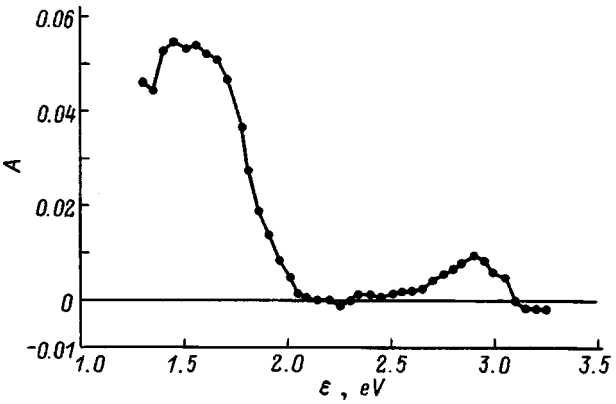

FIG. 3. Spectral dependence of the photoemitter of the polarized-electron source measured in one of the polarimeter channels at the 25 a.u. discrimination level (Fig. 2). The counting rate $\approx2\times10^5$ counts/s.

In addition, we propose a new method for calibrating a polarimeter based on extrapolation of the measured scattering asymmetry to a high discrimination level. The ability of the instrument to operate with high counting rates, up to 5×10^6 counts/s, is utilized in the method. To explain the proposed method, we briefly recall the main principles that have been employed to calibrate Mott detectors in the case where the polarization of the electron beam is unknown.[3]

In high-voltage Mott detectors the function S_{eff} depends strongly on the thickness of the foil. This is mainly due to the low energy resolution of surface-barrier detectors, since inelastically and multiply scattered electrons are detected along with elastically scattered electrons in fairly thick foils, and the information on the spin is consequently lost. Therefore, to calibrate a polarimeter, measurements are performed using foils of different thicknesses, and the results are extrapolated to zero thickness. Then S_{eff} is determined by comparing the extrapolated results with calculations for elastic scattering. In retarding-potential analyzers the dependence of S_{eff} on foil thickness is determined by the size of the energy ''window'' used to detect the scattered electrons. When the energy ''window'' is smaller than 75 eV, the dependence of S_{eff} on foil thickness becomes so weak that its decrease as a consequence of multiple scattering is attributed almost completely to inelastic effects. In this case extrapolation to zero foil thickness is often replaced by extrapolation to zero energy losses.

The idea of our calibration method is based on the fact that in a system with a low energy resolution mostly elec-

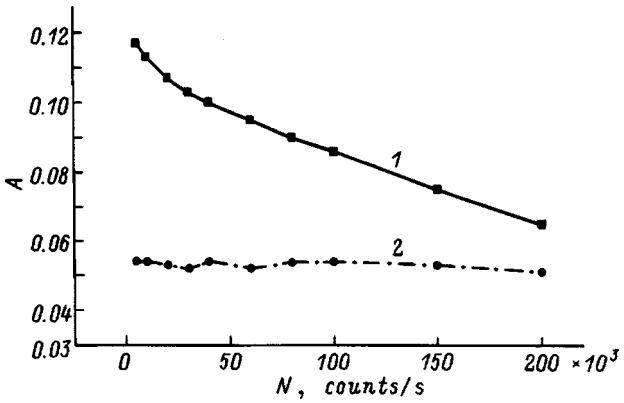

FIG. 4. Dependence of the scattering asymmetry for one of the channels as a function of measured counting rate for two discrimination levels, in arbitrary units (a.u.): *1* — 50, *2* — 25 (Fig. 2).

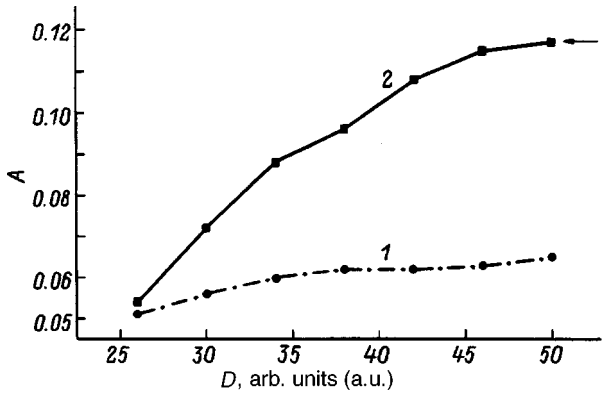

FIG. 5. Dependence of the scattering asymmetry on discrimination level for different counting rates, counts/s: *1* — 2×10^5, *2* — 5×10^3. The arrow on the graph marks the asymmetry value equal to the product $P_0\cdot S_{\text{eff theor}}=0.35\cdot0.33=0.116$.

trons which have undergone elastic scattering will be detected when the discrimination level is increased significantly, i.e., the appearance of signals associated mainly with elastic electron scattering should be expected on the ''tail'' of the signal ''Gaussianlike'' curve at a point at a distance of several multiples of the standard deviation σ (for example, at a discrimination level $\approx$50 a.u.; Fig. 2). The following is important here: the real number of electrons arriving must not exceed the maximum counting rate, which is a characteristic of a given detector. In order to conserve the former measured counting rate as the discrimination level is increased, the electron current at the entrance to the polarimeter must be increased. Clearly, when a certain current, which is determined by the frequency parameters of the detector–preamplifier unit, is achieved, the system will commit counting errors in the statistically distributed pulses. Consequently, the asymmetry should depend strongly on the measured counting rate at high discrimination levels.

Figure 4 presents plots of the dependence of the scattering asymmetry on counting rate for two discrimination levels, viz., the working level and a high level. We note that a significant dependence of the asymmetry on counting rate is, in fact, observed at the high discrimination level. According to our data, the measured counting rate at the 50 a.u. discrimination level is roughly 500 times smaller than the value at the 25 a.u. working level with the same current at the entrance to the polarimeter; therefore, the maximum possible counting rate for the 50 a.u. discrimination level equals $5\times10^6/500=10^4$ counts/s.

Figure 5 shows plots of the dependence of the scattering asymmetry as a function of discrimination level for various counting rates. It is seen that the plot for 5×10^3 counts/s achieves saturation at a discrimination level $\leqslant$45 a.u. The arrow shows the point corresponding to the product of the polarization $P_0=35\pm2\%$ and the theoretical value of the Sherman function calculated for the instrument geometry under consideration and 60 kV, $S_{\text{eff theor}}=0.33$:

$$A_{LR}=P_0\cdot S_{\text{eff theor}}=0.35\cdot0.33=0.116. \quad (4)$$

The value obtained agrees well with the experimentally measured asymmetry at the high discrimination level.

Hence it should be concluded that the present method for extrapolating the measured scattering asymmetry to a high discrimination level can be used to calibrate classical Mott polarimeters equipped with surface-barrier detectors having a low energy resolution. For this purpose, the value of the asymmetry measured at the high discrimination level (when saturation of the asymmetry has been achieved) must be divided by the theoretical value of the Sherman function calculated for the case of monatomic scattering at the respective energy. The value obtained will be equal to the polarization of the electron beam under investigation [formula (2)]. The values of the Sherman function for various energies can be taken, for example, from the tables in Ref. 13.

CONCLUSION

1. A compact 60-kV Mott polarimeter has been developed, tested, and calibrated. The maximum overall dimensions of the instrument are 250×480 mm.

2. The use of a design which combines spherical field acceleration and the absence of a retarding potential after acceleration of the electron beam has permitted the creation of an instrument having high stability of the measured polarization even when the position and diameter of the beam being investigated vary.

3. As a result of optimization of the scattering angle (118°) and the use of surface-barrier detectors with a large collection angle ($\approx$48°), the figure of merit of the polarimeter $\varepsilon=(I/I_0)\cdot(S_{\text{eff}})^2$ equals 2.5×10^{-4}.

4. Specially developed electronic circuits and optimum positioning of the preamplifiers (in direct proximity to the surface-barrier detectors) have made it possible to obtain a maximum electron counting rate as high as 5×10^6 counts/s.

5. A new method for calibrating conventional Mott polarimeters equipped with surface-barrier detectors having a low energy resolution has been proposed. The method is based on extrapolation of the measured asymmetry to a high discrimination level.

We thank Yu. P. Yashin for supplying the semiconductor crystals for the polarized-electron source.

This work was performed as part of the Russian Government Program ''Surface Atomic Structures,'' Project No. 95-1.23.

[1] N. F. Mott, Proc. R. Soc. London **124**, 425 (1929).
[2] J. Kessler, *Polarized Electrons* (Springer, New York, 1985).
[3] T. J. Gay, M. A. Khakoo, F. B. Dunning *et al.*, Rev. Sci. Instrum. **63**, 114 (1992).
[4] J. Kirschner and R. Feder, Phys. Rev. Lett. **42**, 1008 (1979).
[5] M. R. Scheinfein, J. Unguris, D. T. Pierce *et al.*, Rev. Sci. Instrum. **61**, 2501 (1990).
[6] T. J. Gay and F. B. Dunning, Rev. Sci. Instrum. **63**, 1635 (1992).
[7] F. B. Dunning, Nucl. Instrum. Methods Phys. Res. A **347**, 152 (1994).
[8] M. Kalisvaart, M. R. O'Neill, T. W. Riddle *et al.*, Phys. Rev. B **17**, 1570 (1978).
[9] L. G. Gray, M. W. Hart, F. B. Dunning *et al.*, Rev. Sci. Instrum. **55**, 88 (1984).
[10] X. Zhang, H. Hsu, F. B. Dunning *et al.*, Phys. Rev. B **44**, 9133 (1991).
[11] M. Landolt, R. Allenspach, and D. Mauri, J. Appl. Phys. **57**, 3626 (1985).
[12] M. Getzlaff, J. Bansmann, and G. Schönhense, J. Magn. Magn. Mater. **131**, 304 (1994).
[13] G. Holzwarth and H. J. Meister, Nucl. Phys. **59**, 56 (1964).

Translated by P. Shelnitz

Effectiveness of the use of oxygen-containing niobium in thermionic energy converters

V. P. Kobyakov

Institute of Structural Macrokinetics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow District, Russia

A. G. Kalandarishvili

Kurchatov Institute Russian Science Center, 123182 Moscow, Russia

(Submitted June 30, 1997)
Zh. Tekh. Fiz. **68**, 131–137 (August 1998)

Features of the structure and properties of oxygen-containing niobium obtained by vapor-phase crystallization in the presence of an oxygen addition are studied. The effectiveness of the use of this material in fabricating collectors for thermionic converters is evaluated, and the properties and efficiency of oxygen-containing niobium obtained by various methods are compared. © *1998 American Institute of Physics.* [S1063-7842(98)02308-3]

INTRODUCTION

Thermionic converters with an oxygen addition have been the subject of a considerable number of studies, including investigations into the use of various oxygen-containing layers in the collectors in thermionic converters (for example, Refs. 1–3). The most interesting results in this area were obtained in Refs. 4–6, where experimental thermionic converters with an oxygen-containing niobium collector were investigated. In the present work features of the structure and properties of oxygen-containing niobium are studied.

EXPERIMENTAL METHOD

Samples of ''vapor-phase'' oxygen-containing (VPOC) niobium were obtained by reducing niobium pentachloride in hydrogen on a molybdenum substrate at temperatures in the range 1350–1450 K.[7] Oxygen was introduced into the chloride–hydrogen mixture at various fixed flow rates. Niobium interacted with oxygen from the vapor phase during its crystallization. The total oxygen content in the samples was determined by neutron-activation analysis. X-ray structural investigations of the samples were performed on a DRON-2.0 diffractometer at room temperature and with the aid of a URVT-2200VTs camera attachment. The high-temperature camera was outfitted with a built-in quadrupole mass spectrometer. A POLYUS-3 ion microprobe was used to investigate the composition of the subsurface layer of the samples by secondary-ion mass spectrometry (SIMS). The latter two devices were developed and fabricated in the Institute of the Physics of Metals of the Ukrainian Academy of Sciences.

The work function was measured on cylindrical samples by the total-current method[8] in an all-metal apparatus with oil-free evacuation to a residual pressure no higher than 10^{-5} Pa. The measurement error was 0.05 eV. VPOC niobium was investigated directly as a collector material within an experimental thermionic converter with cylindrical electrodes. The forming process, particularly the forming of the cylindrical collector unit, was nonstandard mainly on account of the low plasticity of VPOC niobium. Nevertheless, by combining the special techniques of the vapor-phase deposition process with gas-static conversion, we were able to obtain reliable ceramic-metal collector units with an emission layer inside that can be used to fabricate an experimental thermionic converter. A similar device was made using the SB-1 niobium alloy (1 wt % Zr) for comparative investigations. The cylindrical emitter units of these devices were formed on the basis of single-crystal tungsten tubes with a $\langle 111 \rangle$ axial orientation, which were obtained by a chloride vapor-phase technology.[9,10] The testing of the devices was performed on a high-temperature bench.[11] After the device was outgassed and the interelectrode space was filled with cesium vapor, the current–voltage characteristics were recorded at various emitter and collector temperatures. Optimization with respect to the cesium vapor pressure was performed each time.

RESULTS AND DISCUSSION

1. The investigations previously performed on VPOC niobium[7,12] revealed significant differences between it and other types of oxygen-containing niobium. In particular, it was pointed out that it contains a new metastable oxide phase with an anomalously high hardness, which is capable of releasing oxygen upon heating. Figure 1 presents the x-ray diffractogram ($\mathrm{Cu}\,K_\alpha$ radiation) of one of the samples of VPOC niobium with a total oxygen content equal to 1.9 wt %. An analysis of the phase composition (the JCPDS database) showed that niobium monoxide is present in the sample in an appreciable amount along with the solid solution of oxygen in niobium. It was also found that the lattice constant of niobium is increased by more than 1%, while that of the monoxide is increased by just under 0.5%. The latter may attest to the dissolution of oxygen in niobium monoxide as well. An hypothesis regarding the existence of such a metastable phase, i.e., a solid solution of oxygen in niobium monoxide, was previously advanced in Ref. 12.

These results are consistent with the data from the mi-

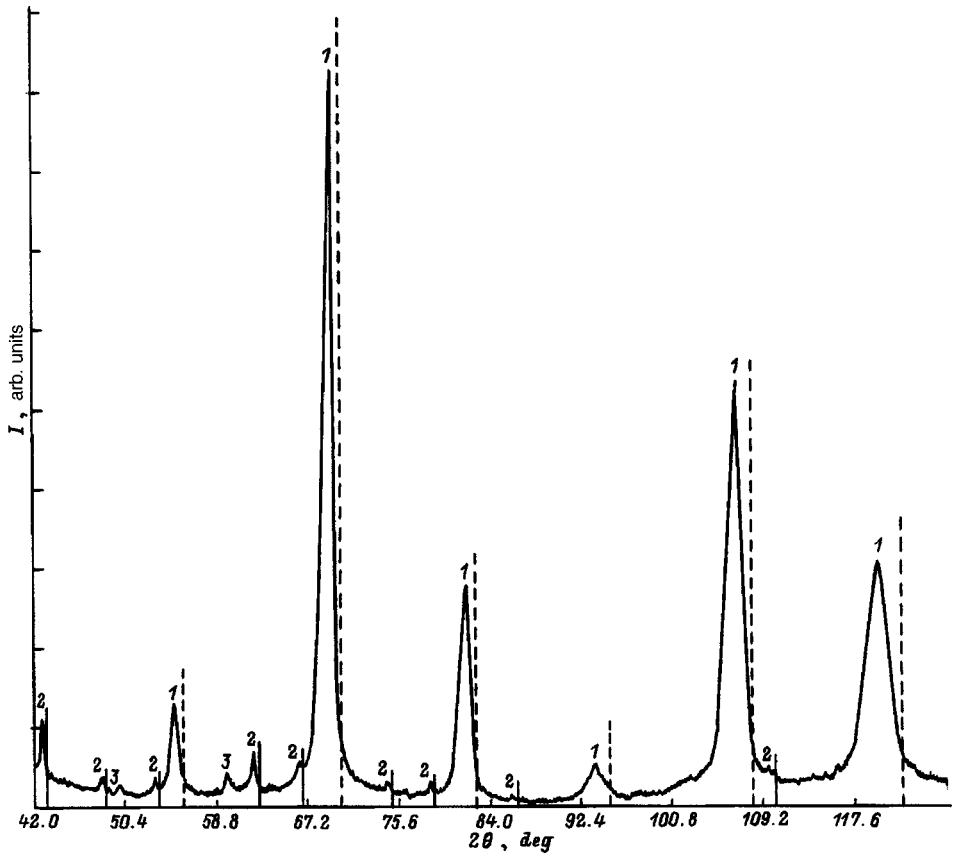

FIG. 1. Diffractogram of vapor-phase oxygen-containing (VPOC) niobium containing 1.9 wt % oxygen. *1* — Nb, *2* — NbO, *3* — presumably NbO_2 (according to card 34-898 in the JCPDS). Vertical dashed lines — position of the reflections of high-purity Nb (card 34-370), vertical solid lines — same for the reflections of NbO (card 15-535).

croprobe investigations (by SIMS) presented in Fig. 2, which show that the subsurface layer of VPOC niobium contains NbO, whose content correlates with the increase in the total amount of oxygen in the samples. As was noted in Ref. 12, the presence of O^+ ions in the mass spectrum is unexpected here, since it is known that it is impossible to knock out oxygen from niobium with dissolved oxygen, even when oxide precipitates are present on the surface, by either heat treatment[13–16] or bombardment with argon ions,[17] and niobium monoxide vaporizes congruently.[14] This phenomenon was attributed in Ref. 12 to the presence of a nonequilibrium oxide phase in the oxygen-containing niobium investigated. The data presented below confirm the hypotheses previously advanced.

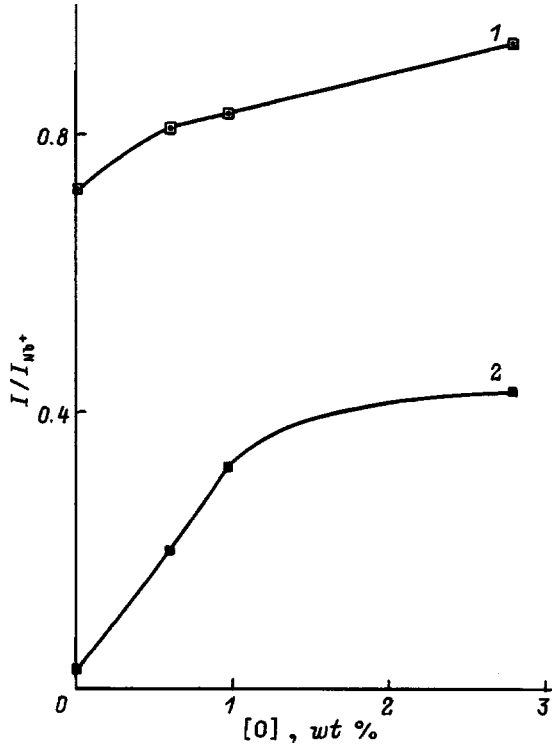

FIG. 2. Intensity of the secondary-ion peaks as a function of the oxygen concentration in VPOC niobium (normalized to the ion current of Nb^+): *1* — NbO^+, *2* — O^+.

The results obtained in the present work can be compared with the data in Ref. 18 on niobium diffusively saturated with oxygen, sintered niobium obtained from a preliminarily compacted mixture of niobium and niobium pentoxide powders, and single-crystal oxygen-implanted niobium. The investigation in Ref. 18, which was performed using x-ray photoelectron spectroscopy (XPS), demonstrated the presence of niobium and oxygen in both the free and bound states in the subsurface layer of all the oxygen-containing materials just cited. The oxides NbO and Nb_2O_5 were identified in the samples in various proportions. It can be asserted that these data do not contradict the results presented above. In addition, it must be taken into account that while a 2–3 nm thick subsurface layer participates in the formation of the electron spectrum in XPS, a layer which is approximately three orders of magnitude thicker produces the x-ray diffrac

tion pattern. The higher niobium oxide discovered by XPS is confined to a very thin surface layer, whose contribution is beyond the sensitivity range of x-ray structural analysis. As for highly sensitive SIMS, the ion microprobe used in the present work did not permit the detection of masses above 220 amu. At the same time, XPS, not being a structural method, cannot tell us anything about the interrelationships between the niobium and oxygen atoms which are detected as bound atoms by this method. The x-ray structural data presented in Fig. 1 are direct evidence of the existence of a new nonequilibrium phase, viz., a solid solution of oxygen in niobium monoxide. As we have already noted here and before in Refs. 7 and 12, the presence of this phase is characteristic only of VPOC niobium, and its formation is apparently due to the specific conditions of the ''atom-by-atom'' formation of a niobium layer from the vapor phase.

It is noteworthy that the presence of oxide phases in VPOC niobium does not have any effect on its electric resistance at either room or high temperature. This is in good agreement with the data in Ref. 19, according to which niobium monoxide is a metallic conductor. This finding provides additional evidence that VPOC niobium does not contain niobium pentoxide, which is an electrical insulator.

To complete this section, we should present the results previously published in a difficultly accessible publication (Ref. 12) on the thermal stability of the new nonequilibrium oxide phase, which will be needed for the ensuing comparisons. These results are presented here in a more complete form than in Ref. 12. We are referring to an investigation of samples of VPOC niobium using a high-temperature x-ray camera outfitted with a mass spectrometer. In particular, a sample containing 0.95 wt % oxygen, which had a thin oxygen-free niobium layer on its surface after synthesis, was investigated. In the first stage of the experiment, in which the sample was heated to 2000 K, the x-ray diffraction pattern contained only niobium reflections, and the composition of the residual gases corresponded to the gas background in the camera without a sample (the hydrogen peak, which increased appreciably when the sample was heated, was an exception). During a high-temperature anneal with continuous evacuation, the partial pressures of all the components of the gas mass spectrum decreased proportionally as a result of the outgassing of the camera and the sample. This attested to the high thermomechnical properties and airtightness of the oxygen-free niobium layer. The sample was subsequently withdrawn from the camera and exposed to air, and after the thin oxygen-free layer was removed from its surface, it was again mounted in the camera. Repeated photographing showed that the x-ray diffraction patterns contain not only niobium reflections, but also reflections of niobium monoxide, which are characteristic of VPOC niobium. Significant changes were also observed in the residual medium over the sample. In particular, an intense oxygen peak appeared (Fig. 3). As the temperature was raised, the two-phase composition of the sample was conserved over the entire temperature range investigated. The release of oxygen increased slightly as the temperature was raised to about 850 K, but it decreased somewhat during further heating to 1400–1450 K. Then a smooth increase in the intensity of the oxygen peak

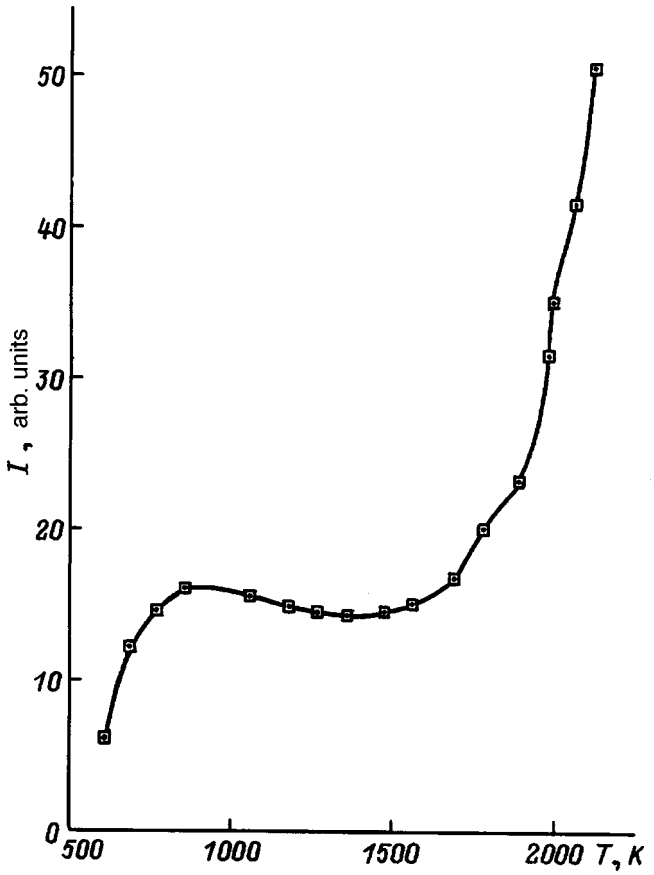

FIG. 3. Temperature dependence of the amplitude of the oxygen peak in the mass spectrum of the vapor phase over VPOC niobium containing 0.95 wt % oxygen.

was observed at first, which became considerably stronger above 1600 K and then sharply intensified at temperatures above 1900 K. It can be assumed that processes involving conversion of the solid solutions of oxygen in niobium and niobium monoxide into oxides with a higher oxygen state predominate on the sample surface at 850–1450 K. However, x-ray diffraction monitoring did not confirm this. We can probably say that destruction of the metastable oxide state begins at the end of this temperature range.

2. The temperature dependence of the effective work function of the electron for samples of VPOC niobium with different total oxygen concentrations was previously investigated in Refs. 12 and 20. Here these data have been supplemented and compared with the results discussed in Subsec. 1. It is seen in Fig. 4 that VPOC niobium exhibits a sharp temperature dependence of the effective work function (curve *1*), as was previously shown for oxygen-containing tungsten in Refs. 21 and 22. However, in the case of tungsten the anneal time did not have any significant influence on the reproducibility of the course of the curve in repeated measurements. In the case of niobium, curve *1* successively transformed during the annealing of the sample into less steep curve *2* and then into nearly linear curves *3–5*, which were faithfully reproduced in repeated measurements. The slope of ''steady-state'' curves *3–8* and their level along the vertical axis correlated with the oxygen content in the samples, as can be seen in Fig. 4.

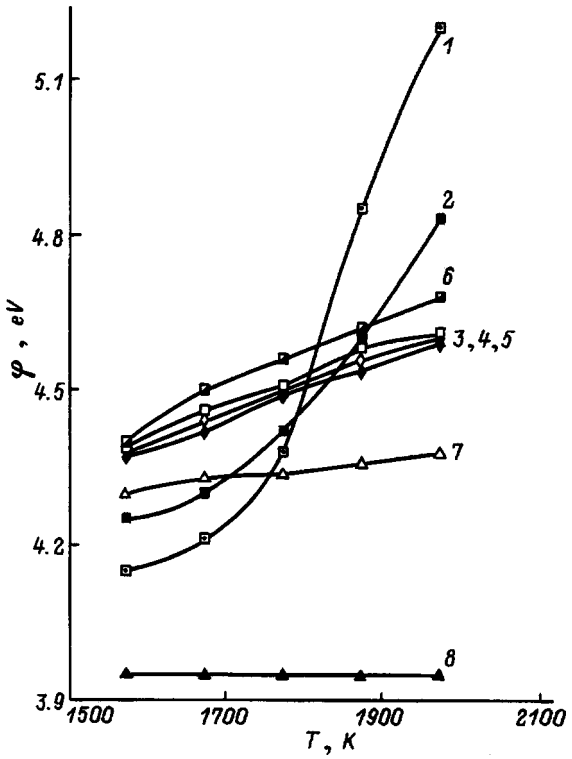

FIG. 4. Temperature dependence of the work function of VPOC niobium with various oxygen concentrations: *1–5* — 3.82, *6* — 4.2, *7* — 0.36, *8* — 1.2×10^{-3} wt %; preanneal in vacuum at 2023 K for 1 (*1*), 2 (*2*), 5 (*3,6–8*), 17 (*4*), and 21 h (*5*).

Figure 5 shows plots of the concentration dependence of the work function for VPOC niobium at various measurement temperatures. It can be seen that when the total oxygen content in the samples reaches 0.4–0.5 wt %, the steeply ascending curves reach a slightly inclined plateau with a positive temperature coefficient $d\varphi/dT=5\times10^{-4}$ eV/K [this value differs significantly from the mean tabulated value for single-crystal oxygen-free niobium: -3.8×10^{-5} eV/K (Ref. 23)]. It should be assumed that the critical oxygen concentration indicated above corresponds to its limiting solubility under the nonequilibrium conditions considered. For comparison, Fig. 5 presents the analogous concentration dependences for oxygen-containing niobium obtained by other methods.[24,25,6] The main differences between the curves characterizing these samples is that they do not reach a plateau and have a less steep course. Curves *6* and *8* correspond to oxygen-containing niobium obtained by the arc remelting of niobium with an oxide dopant, and curve *7* corresponds to oxygen-containing niobium sintered from a mixture of Nb and Nb_2O_5 powders. Curves *6* and *7*, unlike curves *1–5* and *8* were plotted using values of the work function determined from measurements of the contact potential difference.

If we compare the results of the thermionic emission measurements (curves *1–5* and *8*), we should conclude that VPOC niobium has a work function at least 0.2 eV higher than that of arc-remelted oxygen-containing niobium. If we compare the results obtained by measuring the contact potential difference with one another, we might conclude that arc-remelted oxygen-containing niobium is significantly more efficient than VPOC niobium, but there is some basis to assume that curve *6* was obtained without preliminary annealing of the samples to bring their surface into the ''steady'' state.

Thus, the observed differences can attest to the higher thermionic emission efficiency of VPOC niobium in comparison to the other types of oxygen-containing niobium. The comparison with oxygen-containing tungsten mentioned above allows us to conclude that the original chemical state of the VPOC niobium surface is not reproduced after the anneal during the first measurement cycle. As was indicated in Refs. 21 and 22, in the case of tungsten the factor controlling the emission properties is the diffusion of oxygen from the bulk. In the case of VPOC niobium the emission properties are apparently controlled by the presence of the nonequilibrium oxide phase, whose ''steady'' state, as follows from

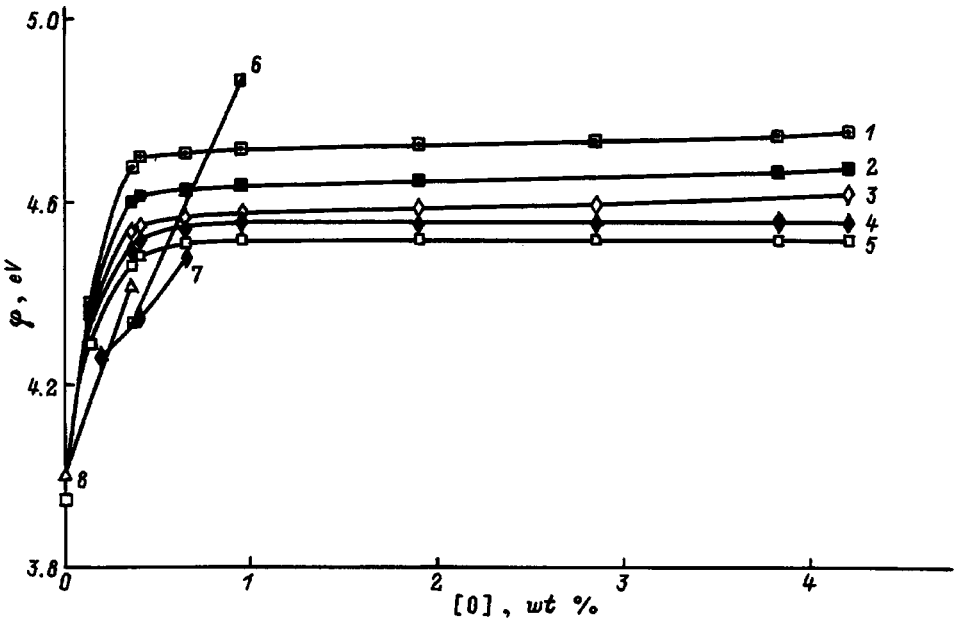

FIG. 5. Concentration dependence of the work function of oxygen-containing niobium at various measurement temperatures: *1* — 1970, *2* — 1870, *3* — 1770, *4* — 1670, *5* — 1570 K (VPOC niobium, measurement of the total current); *6* — data from Ref. 23 (arc-remelted oxygen-containing niobium, measurement of the contact potential difference); *7* — data from Ref. 6 (sintered oxygen-containing niobium, measurement of the contact potential difference); *8* — data from Ref. 22 (arc-remelted oxygen-containing niobium, measurement of the total current).

the data in Fig. 3, is not established at once. During the establishment of thermal equilibrium, the excess oxygen is apparently released into the gas phase and evacuated. After several hours of annealing, a steady state is achieved, which is determined not only by the oxygen concentration in the sample and its temperature, but also by the partial pressure of oxygen in the surrounding residual medium. The compromise between these three factors specifies the steady-state coverage of the sample surface with chemisorbed oxygen, which, in turn, determines the effective work function of the surface.

3. Results of the testing of experimental thermionic converters with an oxygen-containing niobium collector were published in Refs. 4–6 and 26. In most cases (Refs. 4–6) the release of oxygen from the collector into the interelectrode space of the thermionic converter was mentioned during the discussion of the results obtained. The authors based this only on the increase in the steepness of the envelopes of the current–voltage characteristics without presenting any direct experimental evidence of the release specifically of oxygen. At the same time, it is fairly well known (see above) that the outgassing of niobium is accompanied by the removal of oxygen from it only in the form of the monoxide NbO and that appreciable vaporization of the latter takes place without dissociation at temperatures significantly exceeding the collector temperatures.[16] Other hypotheses were advanced only in Ref. 26, where the information in a considerable body of experimental material from both Russian and non-Russian sources was considered. In particular, it was postulated in Ref. 26 as a result of an analysis of the influence of additions of oxygen on the increase in the electron temperature and on the voltage drop in the interelectrode space that these effects are associated with the presence of some volatile oxides of refractory metals in the interelectrode space, which form negative ions in the discharge plasma and thereby degrade the space-charge neutralization conditions. When any oxygen-containing material is used in the collector, it follows from the results presented above, as well as from the data in Ref. 16, that niobium monoxide molecules can appear in the interelectrode space at collector temperatures in the range 1000–1100 K only as result of bombardment of the collector surface and only in the form of positively charge ions. At the same time, when VPOC niobium is used in the collector, the temperature level just indicated is perfectly adequate for the appearance of oxygen in the interelectrode space. Both the ionization of oxygen atoms with the formation of O^+ ions and charge exchange on ions of the monoxide with the formation of negative NbO ions are possible to some extent in the interelectrode plasma. While in the former case the space-charge neutralization conditions are improved, in the latter case they are degraded, as was suggested in Ref. 26. In addition, under these conditions both oxygen neutrals which have escaped being ionized and the monoxide ions that have undergone charge exchange impinge on the emitter, promoting emission, as has been reported.[1–6,22] It can be asserted that among the entire variety of processes in the plasma and on the electrodes of a working thermionic converter, the processes associated with the release of oxygen from the collector material play the main role. However, it follows from the

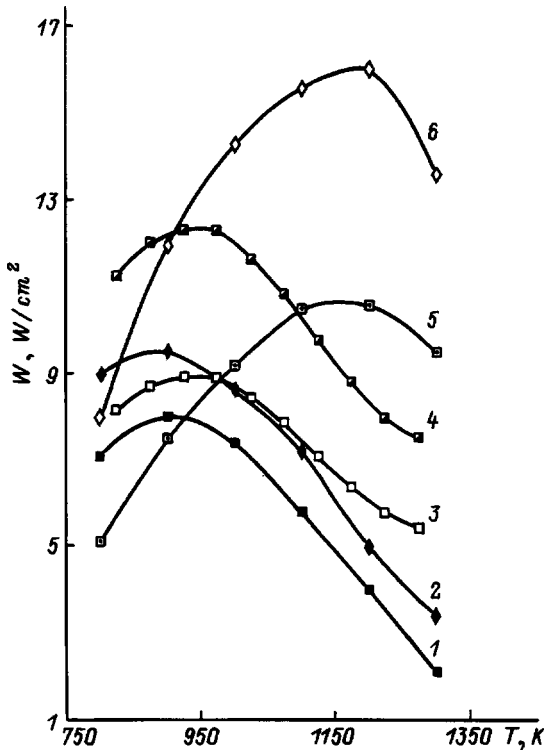

FIG. 6. Dependence of the maximum specific output power of experimental thermionic converters on collector temperature. Collector material: SB-1 niobium alloy (*1*,*2*); sintered oxygen-containing niobium (*3*,*4*); VPOC niobium (*5*,*6*). Emitter temperature, K: 2000 (*1*,*5*); 2100 (*2*,*6*); 1873 (*3*); 2073 (*4*).

data presented in this paper that only VPOC niobium is capable of releasing a considerable amount of oxygen at the collector temperatures. Hence, it would be expected that a thermionic converter with a collector made from such a material will have improved output characteristics.

Figure 6 shows the output characteristics of an experimental thermionic converter with a collector made from VPOC niobium containing 0.65 wt % oxygen. For comparison, the figure also shows the analogous data for a device with a collector made from the SB-1 niobium alloy with an oxygen concentration in the niobium matrix at the 10^{-3} wt % level (the remaining oxygen is bound in ZrO_2 precipitates) and for a device with a collector made from sintered oxygen-containing niobium containing 0.37 wt % oxygen. In the last case the data were taken from Ref. 6. Single-crystal ''vapor-phase'' tungsten chloride served as a the emitter material. The electrode gap measured 0.25 mm.

It should, first of all, be noted that there is a general tendency for a significant increase in the level of the specific output power of thermionic converters when oxygen-containing niobium is used in the collector. This is in good agreement with the facts which have been demonstrated repeatedly before. Another significant point is the shift of the peak power of the thermionic converters toward higher collector temperatures. It is noteworthy that the temperature coefficient of the output power increases somewhat upon passage from the device with the ''oxygen-free'' collector to the device with the collector made from sintered oxygen-containing niobium and rises sharply upon passage to the

device with the collector made from VPOC niobium. When devices with different oxygen-containing niobium materials are compared with one another, two circumstances must be taken into account. On the one hand, the device with VPOC niobium was tested at somewhat higher emitter temperatures (by 27 K for the maximum emitter temperature). On the other hand, the device with VPOC niobium (like the device with the SB-1 niobium alloy) had a cylindrical electrode geometry, unlike the device with sintered oxygen-containing niobium. The latter had flat electrodes. Apart from the known advantages of ''flat'' experimental devices, in the present case the fact that the flat single-crystal tungsten emitter had the form of a single (110) face with maximum emission-adsorption efficiency was significant. As we know, an emitter made from the same material, but with a cylindrical emission surface, has six sites corresponding to $\langle 110\rangle$ directions on the lateral surface in the best case with the $\langle 111\rangle$ axial orientation. A comparison of these surfaces at the maximum possible work function[23] reveals that $\varphi = 5.3$ eV for a flat W(110) face and that $\varphi = 5.0$ eV for a cylindrical W $\langle 111\rangle$ surface. This circumstance is certainly more significant than the small difference in temperature. Nevertheless, the device with VPOC niobium was far more efficient, according to all the papers. The peak specific output power of this device is 16 W/cm^3, 30% higher than in the device with sintered oxygen-containing niobium. Also, the optimum collector temperature is about 1170 K, as opposed to 950 K for the device with sintered oxygen-containing niobium. A 100 K difference in emitter temperatures corresponds to an almost 50% increase in power at the maximum of the curves for the device with VPOC niobium, while it corresponds to only an 18% increase for the device with sintered oxygen-containing niobium.

4. It would certainly be difficult to find an unequivocal interpretation for the results obtained. We should, first of all, note the difference between the collector materials used with respect to total oxygen concentration: 0.65 wt % in VPOC niobium and 0.37 wt % in sintered oxygen-containing niobium. However, according to the data in Fig. 6 for VPOC niobium, both of these compositions are in the region of the plateau on the concentration dependence of the work function, and the curve for sintered oxygen-containing niobium is far lower at any oxygen concentration. Thus, it can be assumed that the significant differences in the thermionic efficiency of the experimental thermionic converters are attributable to the presence in VPOC niobium of an oxide phase, i.e., a solid solution of oxygen in niobium monoxide, which is capable of stably generating oxygen at the collector temperatures.

It follows from the data considered that the total gain in the output power of the devices with VPOC niobium in comparison to the ''oxygen-free'' device of the niobium series amounts to almost 70%. If, following the perfectly obvious arguments advanced in Ref. 6, we attribute this entire gain to improvement of the situation on both the collector and the emitter, we can estimate each contribution separately on the basis of the data presented above. According to our ideas, the effect of the use of sintered oxygen-containing niobium in the collector should be assigned almost entirely to the ''collector'' component of the increase in the power of the device, since the release of the oxygen-containing components in thermionic converters can be neglected in this case. According to the data in Fig. 6, this effect amounts to less than 30%. Thus in the case of the device with VPOC niobium, more than 40% of the power increase corresponds to the increase in emission due to the participation of the oxygen released from the collector in the adsorption dynamics on the emitter surface. However, it should be noted that the contribution of the ''collector'' component can, in fact, be even larger. The fact is that the presence of oxygen and oxygen-containing compounds near and on the collector surface can appreciably diminish the ''oxygen'' effect due to the formation of a thin semiconductor film consisting of cesium oxides or of cesium/oxygen/electrode-material oxide composites on the collector surface and to the increase in electron reflectivity.

In conclusion, it can be presumed that the characteristics of a highly efficient thermionic converter with an oxygen-containing tungsten emitter[22] can be improved considerably by using VPOC niobium with the optimum oxygen content as the collector material in such a device.

We thank B. Ya. Dynkina, V. N. Taranovskaya, I. V. Basilaya, Yu. L. Éloshvili, B. I. Ermilov, and V. G. Kashiya for many years of fruitful cooperation, without which this work could not have been carried out.

[1] D. M. Ernst, in *IEEE Conference Record of 1970 Thermionic Conversion Specialist Conference: Papers Presented at the 9th Annual Conference, Miami Beach, Florida, October 26–29, 1970*, IEEE, New York (1970), pp. 492–497.
[2] J. Dunlay, S. Matsuda, and V. Poirier, in *Third International Conference on Thermionic Electric Power Generation*, West Germany (1972), Vol. 3, pp. 1085–1090.
[3] D. Lieb and F. Rufeh, *ibid.*, pp. 1091–1102.
[4] I. G. Gverdtsiteli, N. E. Menabde, V. K. Tskhakaya *et al.*, in *Thermionic Conversion of Thermal Energy into Electric Power* [in Russian], Izd. FÉI, Obninsk (1980), pp. 160–170.
[5] T. S. Kokosadze, L. M. Tsakadze, and V. K. Tskhakaya, Zh. Tekh. Fiz. **56**, 929 (1986) [Sov. Phys. Tech. Phys. **31**, 586 (1986)].
[6] N. E. Menabde and L. M. Tsakadze, in *Proceedings of the 2nd Specialist Conference ''Nuclear Power Engineering in Space''* [in Russian], Izd. SFTI, Sukhumi (1992), pp. 81–90.
[7] B. Ya. Dynkina, V. P. Kobyakov, S. P. Mitrofanova *et al.*, Izv. Akad. Nauk SSSR, Met. (4), 114 (1986).
[8] L. N. Dobretsov and M. V. Gomoyunova, *Emission Electronics*, Israel Program for Scientific Translations, Jerusalem, 1971.
[9] V. S. Emel'yanov, A. I. Evstyukhin, V. I. Shulepov *et al.*, *Molybdenum in Nuclear Power Engineering* [in Russian], Atomizdat, Moscow (1977), 160 pp.
[10] V. P. Kobyakov, Kristallografiya **41**, 552 (1996) [Crystallogr. Rep. **41**, 523 (1996)].
[11] A. G. Kalandarishvili, *Sources of Working Media for Thermionic Energy Converters* [in Russian], Énergoatomizdat, Moscow (1993), 304 pp.
[12] V. P. Kobyakov and V. N. Taranovskaya, in *Proceedings of the 2nd Specialist Conference ''Nuclear Power Engineering in Space''* [in Russian], Izd. SFTI, Sukhumi (1992), pp. 369–385.
[13] E. Gebhardt, E. Fromm, and D. Jakob, Z. Metallkd. **55**, 432 (1964).
[14] J. Niebuhr, J. Less-Common Met. **11**(3), 191 (1966).
[15] V. N. Ageev, Fiz. Tverd. Tela (Leningrad) **13**, 3353 (1971) [Sov. Phys. Solid State **13**, 2816 (1972)].
[16] E. K. Kazenas and D. M. Chizhikov, *Pressure and Composition of the Vapor over Oxides of Chemical Elements* [in Russian], Nauka, Moscow (1976), 342 pp.
[17] V. I. Shvachko, B. T. Nadykto, Ya. M. Fogel' *et al.*, Fiz. Tverd. Tela. (Leningrad) **7**, 1944 (1965) [Sov. Phys. Solid State **7**, 1572 (1966)].

[18] B. B. Shishkin, I. S. Omel'chuk, S. N. Voronov *et al.*, Fiz. Khim. Obrab. Mater., No. 5, 76 (1987).
[19] V. B. Lazarev, V. V. Sobolev, and I. S. Shaplygin, *Chemical and Physical Properties of Simple Metal Oxides* [in Russian], Nauka, Moscow (1983), 240 pp.
[20] V. P. Kobyakov, N. E. Menabde, A. M. Sirenko *et al.*, Poverkhnost', No. 2, 145 (1987).
[21] V. P. Kobyakov and S. A. Zaslavskiĭ, Vysokochist. Veshchestva, No. 5, 75 (1989).
[22] V. P. Kobyakov, Zh. Tekh. Fiz. **66**(7), 169 (1996) [Tech. Phys. **41**, 726 (1996)].
[23] V. S. Fomenko, *Emission Properties of Materials. Handbook* [in Russian], Naukova Dumka, Kiev (1981), 340 pp.
[24] I. G. Gverdtsiteli, N. E. Menabde, and L. M. Tsakadze, in *Abstracts of the 17th All-Union Conference on Emission Electronics* [in Russian], Leningrad (1978), p. 196.
[25] G. M. Drok, A. I. Kontorez, R. Ya. Kucherov *et al.*, Élektron. Tekh. Mater., No. 6(155), 12 (1981).
[26] V. Z. Kaĭbyshev, Yu. N. Dzhashiashvili, V. A. Koryukin *et al.*, in *Proceedings of the 2nd Specialist Conference ''Nuclear Power Engineering in Space''* [in Russian], Izd. SFTI, Sukhumi (1992), pp. 16–31.

Translated by P. Shelnitz