Model for the shock actuation of a wedge-shaped nozzle with allowance for detachment of the flow

V. G. Maslennikov and V. A. Sakharov

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

The motion of the starting gasdynamic discontinuities and of the detached flow arising upon the shock actuation of a wedge-shaped nozzle is examined in the framework of the problem of the decay of an arbitrary discontinuity at a jump in the cross-sectional area of the channel. The results of the calculation are compared with experimental measurements of the spatial and time dependence of the starting discontinuities upon the shock actuation of submerged jets over a wide range of variation of the parameters governing the flow: at pressure ratios $100 < \bar{p} < 1700$ and sound velocities $0.21 < \bar{a} < 2.4$ of the issuing and immersing gases, for issuing gas species with $1.16 < \gamma < 1.67$, and for various combinations of these parameters. The possibility of using this model to determine the relaxation time to an undetached supersonic flow in a wedge-shaped nozzle upon the shock actuation of slightly underexpanded and overexpanded jet flows is analyzed. © *1997 American Institute of Physics.* [S1063-7842(97)00106-2]

To describe the processes accompanying the shock actuation of hypersonic nozzles and highly underexpanded jets, a model of a suddenly actuated steady-state radial source has been developed.[1] There, as in numerical calculations[2,3] describing the shock actuation of submerged jets, the establishment of a supersonic flow in the nozzle is linked to the arrival, at a specified cross section, of a stagnation wave, which is a surface normal to the streamlines. In actuality the relaxation time of a flow in a nozzle upon actuation of slightly underexpanded and overexpanded jets is determined by the dynamics of the detached flow inside the nozzle, one of the elements of the shock-wave structure of which is an oblique shock separating the supersonic flow in the nozzle and the detached jet flow.

Studies of the shadow patterns and interferograms of the flow occurring upon the actuation of wedge-shaped nozzles and a comparison of these with the patterns of steady detached flow in such nozzles[4] has shown that the jet flow behind the detachment point is a planar flow moving nearly parallel to the plane of symmetry of the nozzle at a certain distance downstream from the detachment point. Based on this fact, let us consider a simplified gasdynamic structure of the flow arising upon the shock actuation of a wedge-shaped nozzle. Suppose that in a certain cross section of the nozzle there is detached flow, as a result of which the flux of the issuing gas upon going through the oblique shock is rotated through an angle equal to the half angle of the nozzle. Then a region of one-dimensional flow forms behind the detachment point. The reflection of the oblique wave from the plane of symmetry of the nozzle disrupts the one-dimensional character of the flow in the central region. We will neglect the influence of this part of the flow on the overall flow pattern. The supersonic flow ahead of the oblique shock will be assumed radial. Three-dimensional flow occurs in the region behind the starting shock wave and in the neighborhood of the front of the issuing gas.

Experimental studies of the actuation of submerged jets of various gases[5] have shown that the velocity of the detachment cross section in the nozzle is largely determined by the velocity of the gasdynamic discontinuities preceding it. This experimental fact, along with the results of Ref. 6, suggests that a solution might be found using a slightly different statement of the problem than in the conventional approach.

We will determine the velocities of the starting wave and contact surface from the solution of the problem of the decay of a discontinuity in a channel with an abrupt change in cross-sectional area. Here the Mach number M_1 of the shock wave behind the jump in cross-sectional area is given by the relation

$$\bar{p} = \frac{2\gamma_1 M_1^2 - (\gamma_1 - 1)}{(\gamma_1 + 1)g}$$
$$\times \left[1 - \frac{\gamma - 1}{\gamma_1 + 1} \frac{g}{\bar{a}}^{-\frac{\gamma-1}{2\gamma}} (M_1 - 1/M_1) \right]^{-\frac{2\gamma}{\gamma-1}}, \quad (1)$$

where $\bar{p} = P_0/P_1$ and $\bar{a} = a_0/a_1$ are the ratios of the pressures and sound velocities of the gases on different sides of the discontinuity, and γ and γ_1 are the adiabatic exponents of these gases.

In reference to the shock actuation of a supersonic nozzle these parameters are determined by the initial conditions of the issuing flow. Since in the regimes under study the critical values of the flow parameters are always reached at the nozzle throat, here the parameter g is a single-valued function of the relative change in cross-sectional area of the channel. For example, for the subsonic part of the nozzle g is determined by the ratio of the cross-sectional areas of the shock tube channel and the nozzle throat. The velocities of the shock wave and contact surface at the nozzle throat, which are at the same time the input parameters for the supersonic part of the nozzle, are, respectively

$$V_1 = M_1 a_1, \tag{2}$$

$$V_2 = \frac{2a_1}{\gamma_1 + 1}(M_1 - 1/M_1). \tag{3}$$

We will use solution (1)–(3) to calculate the velocities of the starting wave and contact surface in the supersonic part of the nozzle after determining the parameter g as a function of the ratio of the cross-sectional areas at the throat and at the cross section where the boundary of the steady flow region is located at that point in time. For regimes in which a gas is issuing into a rarefied space (vacuum) this boundary is a surface of weak discontinuity of the gasdynamic parameters, in the crossing of which the velocity vector of the flow does not change direction. In the actuation of submerged jets a system of intense secondary shock waves arises, causing detachment of the flow. The nozzle cross section that passes through the point of detachment of the flow, which we will call the detachment cross section, will be assumed coincident with the boundary of the steady flow region. The values calculated for the starting gasdynamic discontinuities in this case correspond to flow into a straight channel with a cross-sectional area equal to the area of the detachment cross section. The three-dimensional character of the flow in the region between the starting wave and the contact surface leads to a decrease in the velocities of these discontinuities. To take this into account the values obtained for V_1 and V_2 must be corrected in accordance with their position in the nozzle. The velocity of the starting wave at a given point in time is determined by the Chisnell relation from the known ratio of the areas of the detachment cross section to the cross-sectional area at the location of the starting wave at that time, and the velocity of the contact surface is determined from relation (3).

Analysis of the interferograms of the flow show that the maximum intensity of the oblique shock occurs near the detachment point. This suggests that the motion of the oblique shock is determined by the flow near that point and can therefore be treated as motion of a plane oblique wave in the uniform steady supersonic flow of an ideal gas. The change in the normal components of the velocities on passage through the oblique shock is given by the relation

$$\frac{U_1 \sin(\alpha) - W \sin(\alpha - \theta)}{(U_2 - W) \sin(\alpha - \theta)} = \bar{\rho}_{21}, \tag{4}$$

where α is the angle between the velocity vector U_1 of the oncoming flow and the plane of the oblique shock, θ is the angle of rotation of the flow, W is the velocity of the oblique shock wave in the direction of the velocity vector U_2 of the flow behind the discontinuity, and $\bar{\rho}_{21}$ is the ratio of the densities to the right and to the left of the discontinuity.

Taking the velocity of the flow behind it as the scale for the velocity of the detachment cross section, we divide the numerator and denominator of the left-hand side of (4) by $U_2 \sin(\alpha - \theta)$ and, with allowance for the equality of the tangential components of the velocities, $U_1 \cos(\alpha) = U_2 \cos(\alpha - \theta)$, we obtain

$$\frac{\tau - w}{1 - w} = \bar{\rho}_{21}. \tag{5}$$

Here we have introduced the notation $\tau = \tan(\alpha)/\tan(\alpha - \theta)$ and $w = W/U_2$. Dividing the right- and left-hand sides of Eq. (5) by w, we require an extremum of the resulting ratio

$$\frac{\tau/w - 1}{1 - w} = \frac{\bar{\rho}_{21}}{w}. \tag{6}$$

Since the flow is assumed steady on both sides of the discontinuity, the differentiation of the left-hand side is done only with respect to the variable w. Then an extremum of (6) will be reached when the following condition is met:

$$w = \tau - (\tau^2 - \tau)^{1/2}. \tag{7}$$

With the given choice of sign in front of the term in parentheses, condition (7) gives a maximum of the ratio $\bar{\rho}_{21}/w$.

The physical meaning of this last condition, which we take as a postulate, becomes clear when this ratio is expressed in dimensional quantities. The quantities ρ_2 and U_2 appearing in the numerator of this ratio monotonically increase and decrease, respectively, as the wave velocity W decreases. Their product $\rho_2 U_2$ has an extremum corresponding to the flow with the maximum specific flow rate of the gas behind the detachment cross section.

By substituting (7) into (5) we can express the velocity w of the detachment cross section in terms of the density ratio $\bar{\rho}_{21}$:

$$w = \frac{1}{2 - 1/\bar{\rho}_{21}}. \tag{8}$$

Here w takes on the value 1 when the shock wave degenerates into an acoustic disturbance and asymptotically approaches the value 0.5 with increasing $\bar{\rho}_{21}$. The strongest dependence of w on $\bar{\rho}_{21}$ occurs for values of $\bar{\rho}_{21}$ close to 1. We note that relation (8) is independent of the character of the flow and can be obtained by considering the motion of a non-oblique shock.

It is clear that the motion of the detachment cross section will be determined by the choice of the function $\bar{\rho}_{21}(t,x)$. We will assume that the intensity of the oblique shock is a function of its position in the nozzle only. Then rotation of the flow through a specified angle θ is possible at values of the Mach number of the oncoming flow greater than a certain minimum value M_m. The density ratio $\bar{\rho}_m$ corresponding to this case in expression (5) occurs at the minimum value of τ, which is attained at the maximum value of the angle

$$\alpha_m = \tan^{-1}(\tan(\theta) + 1/\cos(\theta)). \tag{9}$$

The maximum value of the density ratio depends on the species of gas and is equal to $\bar{\rho}_{max} = (\gamma + 1)/(\gamma - 1)$. Since for an angle of rotation of the flow $\theta = 15°$ and a value $\gamma = 1.16$ the range of variation of the density ratio is $2.79 < \bar{\rho}_{21} < 13.5$, which corresponds to a change in velocity of the detachment cross section over the range $0.609 < w < 0.519$. We note that the given velocity estimate agrees satisfactorily with the measurements of the initial velocity of the detachment point.[5] The ratio of the densities on the two sides of the shock is

$$\bar{\rho}_{21} = \frac{(\gamma+1)[M\sin(\alpha)]^2}{(\gamma-1)[M\sin(\alpha)]^2+2}, \qquad (10)$$

where M is the Mach number of the flow ahead of the detachment cross section.

In the general case the Mach number depends on the difference between the velocities of the supersonic flow and detachment cross section. However, the weak dependence of w on $\bar{\rho}_{21}$ in relation (8) gives some freedom in the choice of the latter. For simplicity in finding the solution we set the number M in expression (10) equal to the Mach number of the steady flow in the nozzle.

Thus the solution of the problem of actuation of a nozzle is constructed as follows. The initial conditions of the issuing flow are used to calculate the velocities of the starting wave and contact surface at the nozzle throat, and these velocities determine the position of these corresponding discontinuities in the first time step. At later times Eqs. (5), (7), (9), and (10) are used to give the values of the angles α and θ and the velocity w of the oblique shock for a known position of the latter in the nozzle prior to the time when the angle of rotation θ of the flow equals a specified value, i.e., the half angle of the nozzle. Here the density $\bar{\rho}_{21}$ varies from 1 to $\bar{\rho}_m$ with increasing M. With allowance for the character of the function (8) at small values of $\bar{\rho}_{21}$ and the assumptions made concerning this quantity in relation (10), one can expect that w will be determined only to within a large error in the initial time period. However, the the size of the corresponding part of the nozzle is of the order of the height of the throat, which amounts to only an insignificant fraction of its total length. A calculation shows that this does not have an appreciable influence on the determination of the actuation time of the nozzle. The subsequent motion of the detachment cross section is governed by the solution of equations (5), (7), and (10) at a constant angle of rotation of the flow. The velocities of the starting shock wave and contact surface are calculated from the known position of the detachment cross section in the nozzle. The trajectories of the starting discontinuities in the nozzle are constructed by numerical integration over time, and the time of its actuation is determined.

Figure 1 shows the calculated trajectories of the starting wave (*1*) and detachment point (*2*) together with the experimental data points corresponding to the actuation of two nozzles with different degrees of expansion, with an angle $\theta = 15°$. The horizontal lines on the graphs indicate the exit section of the nozzle. Figure 1a shows a comparison of the calculated and experimental data for the flow of a jet of argon (filled points, $\bar{p} = 114$, $\bar{a} = 2.38$) and a jet of carbon dioxide (open points, $\bar{p} = 100$, $\bar{a} = 2.39$) from a nozzle with a relative expansion of 10 into a space filled with a gas of the same species. Figure 1b shows a comparison of the results of the calculation and experiment for the formation of a jet of tetrafluoromethane flowing out of a nozzle with a relative expansion of 31 into a space filled with the same gas (filled points, $\bar{p} = 157$, $\bar{a} = 1.49$) and with a gas of a different species, viz., hydrogen (open points, $\bar{p} = 184$, $\bar{a} = 0.21$). The calculated data points pertain to a value $\theta = 15°$ and are in good agreement with the experimental data. The difference
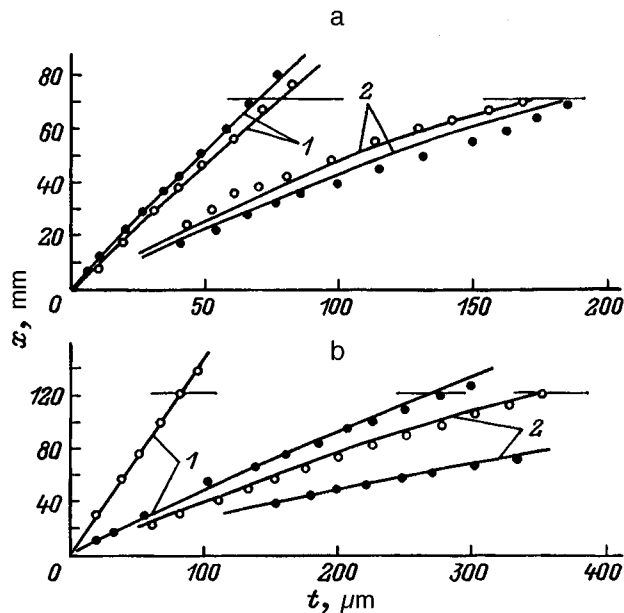

FIG. 1.

between the calculated actuation time of the nozzle and the time measured in the experiments for these regimes does not exceed 5%. The calculated values of the angle of inclination α of the shock to the geometric generator of the nozzle also agree with the experimental measurements.

Increasing the initial pressure differential $\bar{p}$ under otherwise equal conditions leads to a decrease in the intensity of the secondary shock wave and to its degeneration into an acoustic disturbance in the case of flow into vacuum. In the latter case an undetached flow is realized in the nozzle. It is clear that for outflow regimes with a sufficiently large ratio of the initial pressures the angle of rotation of the flow will be less than the half angle of the nozzle. One of the main parameters of the problem which determine the dynamics of the starting process is the number M_1 in relation (1). It would seem advisable to relate the value of the flow rotation angle θ with the value of this parameter at the nozzle throat. For example, for flow into a rarefied space, as $\bar{p}$ tends toward infinity while the Mach number tends toward its limiting value M_m, the flow rotation angle θ tends toward zero. As $\bar{p}$ and, accordingly, M_1 decrease, the angle θ takes on values close to the nozzle half angle θ_0. One can approximate the dependence in the following way, for example:

$$\theta/\theta_0 = 1 - M_1/M_m. \qquad (11)$$

The points in Fig. 2 show the experimentally obtained trajectories of the starting discontinuities corresponding to the flow of tetrafluoromethane into the same gas at $\bar{p} = 1717$, $\bar{a} = 1.53$, and also the results of a calculation (curves *1* and *2*) carried out for different angles of rotation of the flow. It is seen that the trajectory of the detachment cross section calculated for the value $\theta = 11°$ obtained from (11) is in better agreement with experiment than the trajectory constructed for $\theta = 15°$.

Using relation (11), we calculated the trajectories of the starting gasdynamic discontinuities in the shock actuation of
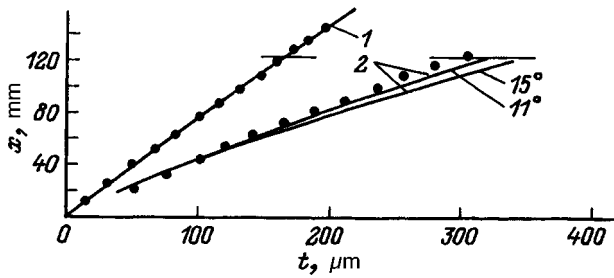
FIG. 2.

FIG. 3.

a wedge-shaped nozzle with a half angle $\theta = 15°$ for flow regimes with various gases, with $1.16 < \gamma < 1.67$, issuing into a space immersed either in the same gas or in hydrogen. The ratio of the initial pressures and sound velocities at the nozzle entrance were varied in the ranges $100 < \bar{p} < 1700$ and $0.21 < \bar{a} < 2.4$, respectively. In the entire range of variation of the governing parameters of the flow the disagreement between the calculated actuation time of the nozzle and the experimental value was never more than 10%.

In the proposed approach to the problem of the actuation of a nozzle it is assumed that the three-dimensional flow in the region between the starting wave and the contact surface does not affect the parameters of the flow behind the detachment point. This condition will be violated at subsonic velocities of the issuing gas behind the detachment cross section or of the entrained flow of the immersing gas behind the shock wave. The presence of subsonic flow velocities behind the detachment point indicates that the boundary of the domain of applicability of the given model is being reached in the region of small ratios of initial pressures. Experiments have shown that the greatest influence on the actuation of the region of three-dimensional flow is observed upon an abrupt change in the shape of the channel, e.g., at the emergence of the starting discontinuities from the nozzle. The points in Fig. 3 show the experimental[7] positions of the points of intersection of the oblique shock waves along the plane of symmetry of a wedge-shaped nozzle transitioning into a channel of constant cross section, at successive points in time (the numbers on the graph give the time in μs). Marked off along the geometric generator of the nozzle (half angle $\theta = 30°$) are the coordinates of the detachment point of the flow, as obtained from a calculation according to relations (5), (7), (9), and (10). The dashed lines join points on the axis and on the generator of the nozzle which correspond to the same points in time from the start of the outflow. For the outflow of nitrogen at values $\bar{p} = 170$, $\bar{a} = 3.12$ (unfilled points) the flow behind the detachment cross section is supersonic. In that case one observes equality of the velocities of the detachment point and the crossing point of the oblique
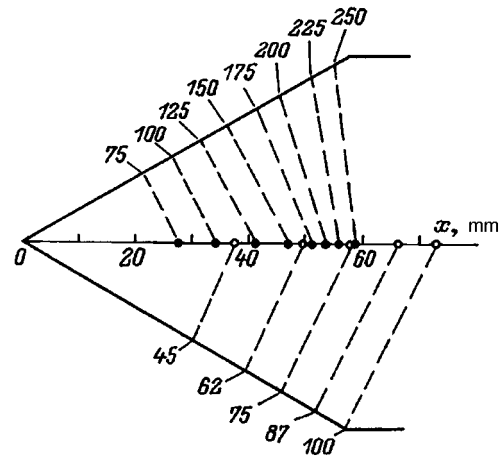
shocks along the entire length of the nozzle, as is reflected in the fact that the lines remain parallel. For a regime with subsonic detached flow, $\bar{p} = 20$, $\bar{a} = 1.62$ (filled points in the upper part of Fig. 3) the parallelism of the lines breaks down at approximately 175 μs. A comparison with experiment shows that for such regimes the calculated velocity of the detachment point is greater than the value realized in experiment, starting from the time when the contact surface emerges from the nozzle. This tendency is preserved both for a nozzle transitioning to a channel of constant cross section and for a nozzle opening into an unbounded space. Results of a similar nature were obtained for a nozzle with a half angle of 5°.

The influence of the entrained subsonic flow on the motion of the detachment cross section is also manifested most strongly at a change in shape of the channel. Experiments show that when the starting discontinuities emerge from the nozzle into an unbounded space, the velocity of the detachment cross section can increase or decrease, depending on the parameters of the flow behind the starting shock wave. Similar effects have been noted previously.[5]

[1] S. F. Chekmarev and N. V. Stankus, Zh. Tekh. Fiz. **54**, 1576 (1984) [Sov. Phys. Tech. Phys. **29**, 920 (1984)].
[2] A. B. Britan, Tr. Inst. Mekh. MGU, No. 43, pp. 48–57 (1976).
[3] V. L. Grigorenko, Izv. Akad. Nauk SSSR Mekh. Zhidk. Gaz., No. 1, pp. 120–127 (1980).
[4] B. M. Dobrynin, V. G. Maslennikov, and V. A. Sakharov, Zh. Tekh. Fiz. **58**, 2390 (1988) [Sov. Phys. Tech. Phys. **33**, 1459 (1988)].
[5] V. G. Maslennikov and V. A. Sakharov, Zh. Tekh. Fiz. **65**(8), 190 (1995) [Tech. Phys. **40**, 854 (1995)].
[6] B. M. Dobrynin, V. G. Maslennikov, and V. A. Sakharov, Zh. Tekh. Fiz. **57**, 118 (1987) [Sov. Phys. Tech. Phys. **32**, 69 (1987)].
[7] T. V. Bazhenova and L. G. Gvozdeva, *Nonstationary Interactions of Shock Waves* [in Russian], Nauka, Moscow (1977), p. 274.

Translated by Steve Torstveit

Cylindrical MHD induction device in the "ideal pressure source" regime

Yu. A. Polovko, E. P. Romanova, and É. A. Tropp

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg

The causes for the onset of a specific operating regime of MHD induction devices which arises after a loss of stability of the uniform flow are investigated. A modification of the original method of studying singularly perturbed dynamical systems in the neighborhood of a bifurcation point is used to construct the asymptotic behavior of the characteristic pressure for various limiting cases. © *1997 American Institute of Physics.* [S1063-7842(97)00206-7]

INTRODUCTION

Even in the first experiments[1] confirming the predicted[2] loss of stability of the flow in MHD induction devices a characteristic property of the secondary flows that arise was noted: after the loss of stability an MHD induction device will operate as an ideal pressure source over a wide range of flow rates. The external characteristics of the device (the curve of the head developed by the device versus the flow rate of the electrically conducting fluid flowing through it) has an extended "plateau" $p=p_c=$ const, which corresponds to a stepped profile of the velocity of the secondary flows. This same phenomenon was later detected in a number of other experiments on cylindrical and planar MHD devices in the constant supply current and constant grid voltage modes. This effect has also been noted in calculations using one-dimensional jet models both for pumps and for MHD generators and chokes.[3–5]

At least three different ways of constructing the horizontal part of the branched external characteristic have been proposed (Fig. 1): through the minimum of the characteristic of the electrodynamic approximation[6] (segment *1*), through the point of zero flow rate[7] (segment *2*), and in accordance with the three-velocity flow model, when $AB=BC$, as is shown in Fig. 1.[8] However, none of these ways has been rigorously justified, and it remains a relevant problem to explain the causes of this phenomenon and to develop a method of estimating *a priori* the value of the characteristic pressure for an MHD device with specified parameters.

The solution of this problem has both theoretical and applied significance, since, on the one hand, the phenomenon in question can be observed in a wide class of MHD flows with a definite geometry of the channel, and, on the other hand, MHD devices of this type are used in the cooling loops of fast-neutron reactors, and it is important to know what kinds of flow regimes can arise in these devices, which must be highly reliable.

In Ref. 9 an attempt was first made to obtain an asymptotic formula for estimating the characteristic pressure at high supply currents. The result was based on the solution of a nonlinear eigenvalue problem arising in the construction of the internal boundary layer of the secondary flows. The use of additional, insufficiently well justified hypotheses as to the structure of the boundary layer has led to an expression which, although strictly speaking is not an asymptotic formula, is entirely suitable for practical estimates.

In this paper we examine a mathematical model which takes into account the friction between the fluid jets and which was used in Ref. 5 for studying the secondary flows in MHD devices. Here the derivation is of a sufficiently rigorous character, and it results in two different asymptotic formulas.

MATHEMATICAL MODEL

To describe the flows in MHD induction devices, the following system of nonlinear differential equations

$$\varkappa d^2\mathbf{b}(\varphi)/d\varphi^2=[1-\mathbf{i}\varepsilon(1-v(\varphi))]\mathbf{b}(\varphi)+\mathbf{i},$$

$$\nu d^2 v(\varphi)/d\varphi^2=-\varepsilon^{-1}j^2\mathrm{Re}\,\mathbf{b}(\varphi)$$
$$+v^2(\varphi)\mathrm{sign}(v(\varphi))+p \tag{1}$$

was used in Ref. 5, with the periodic boundary conditions $\mathbf{b}(0)=\mathbf{b}(2\pi)$, $\mathbf{b}'(0)=\mathbf{b}'(2\pi)$, $v(0)=v(2\pi)$, $v'(0)=v'(2\pi)$, where $\mathbf{b}$ is the relative value of the complex amplitude of the magnetic induction $\mathbf{B}(\varphi,t)=\mathbf{b}(\varphi)\times\exp(\mathbf{i}(\alpha x-\omega t))$, v is the relative value of the flow velocity, p is the relative value of the pressure developed by the device, $\varkappa=1/\alpha^2 R^2$, $\varepsilon=\mu_0\sigma\omega\delta_h/\alpha^2\delta_m$ is the magnetic Reynolds number, j is the dimensionless supply current, ν is the reduced hydrodynamic viscosity, τ is the
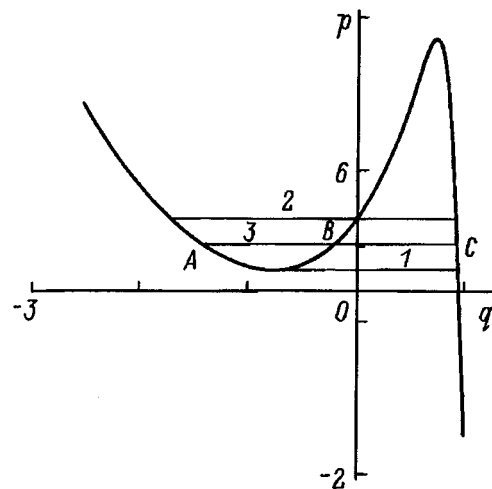

FIG. 1. Previously proposed methods of constructing the segment of the external characteristic $p=p_c$ of the secondary flows in an MHD device, corresponding to its operation as an ideal pressure source (pump mode).

1063-7842/97/060591-04\$10.00

pole pitch, L is the length of the device, δ_h is the width of the hydraulic gap, δ_m is the width of the magnetic gap between the central and outer magnetic circuits, σ is the conductivity of the working medium, R is the radius of the device, and $\mathbf{i}$ is the imaginary unit.

In contrast to earlier systems,[2-4,9] which used the so-called one-dimensional jet model of the flow, system (1) takes into account the friction between the individual fluid jets, and the balance equation for the pressures goes over from a finite-difference to a differential equation, and the order of the system increases from fourth to sixth.

We will investigate the behavior of system (1) for wide channels, i.e., for small values of ν and $\varkappa$. Here we consider two different paths to the limit at the point $(\nu,\varkappa)=(0,0)$. In the first case the value of $\varkappa$ is assumed negligible, and the terms of system (1) containing this quantity as a factor are dropped, and then the asymptotic behavior of the pressure p_c as $\nu \to 0$ is found. In the second case, one first neglects the friction between jets ($\nu=0$) and then investigates the properties of the system at small $\varkappa$.

We use the notation $\mathrm{Re}\,b(\varphi)=x_1$, $\mathrm{Im}\,b(\varphi)=x_2$ and transform the initial system of equations and boundary conditions to the form

$$\varkappa x_1'' = x_1 + \varepsilon x_2 - \varepsilon x_2 \mathrm{v} = f_1(x_1,x_2,\mathrm{v},p,\varepsilon,j),$$

$$\varkappa x_2'' = x_2 - \varepsilon x_1 + \varepsilon x_1 \mathrm{v} + 1 = f_2(x_1,x_2,\mathrm{v},p,\varepsilon,j),$$

$$\nu \mathrm{v}'' = -\varepsilon^{-1} j^2 x_1 + \mathrm{v}|\mathrm{v}| + p = f_3(x_1,x_2,\mathrm{v},p,\varepsilon,j),$$

$$x_i(0) = x_i(2\pi), \quad x_i'(0) = x_i'(2\pi), \quad i=1...2,$$

$$\mathrm{v}(0) = \mathrm{v}(2\pi), \quad \mathrm{v}'(0) = \mathrm{v}'(2\pi). \tag{2}$$

CASE OF INFINITELY WIDE CHANNELS

To a first approximation as $\varkappa \to 0$ after the terms containing $\varkappa$ are neglected and x_1 and x_2 are expressed in terms of v, problem (2) is rewritten as

$$\nu \mathrm{v}'' = p - \frac{j^2(1-\mathrm{v})}{\varepsilon^2(1-\mathrm{v})^2+1} + \mathrm{v}|\mathrm{v}| = p - P(\mathrm{v}) = F(\mathrm{v},p,\varepsilon,j),$$

$$\mathrm{v}(0) = \mathrm{v}(2\pi), \quad \mathrm{v}'(0) = \mathrm{v}'(2\pi). \tag{3}$$

Here $P(\mathrm{v})$ denotes the external characteristic of the MHD device, which corresponds to uniform flow. System (3) belongs to the class of second-order autonomous dynamical systems, which were analyzed in detail in Ref. 10. For $p_1^i < p^i < p_2^i$, $i=1,2$ (Fig. 2), it has three equilibrium positions: v_1^i, v_0^i, and v_2^i, one of which (v_0^i) is of the ''center'' type, while the other two are saddle points. At a certain value $p_c^i \in (p_1^i,p_2^i)$ a bifurcation arises in system (3) — a separatrix appears, joining the two saddle points of rest (Fig. 3). As was shown in Ref. 10, in the neighborhood of the points p_c^i for small ν the solution of equation (3) (the velocity profile) will have a form which is close to stepped, specifically, there will be a sharp jump between the values v_1^i and v_2^i. The position of this jump, i.e., the extent of each step, is very sensitive to the deviation of p from p_c. For $|p^i - p_c^i| \sim \exp(-1/\nu)$ the ratio of the lengths of the steps and, hence, the value of the flow rate, change by an amount of the order of unity.[10] It is this effect that leads to a hori-
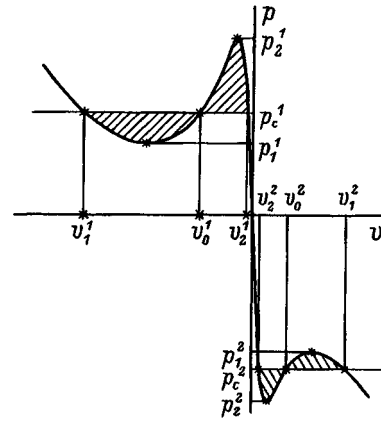
FIG. 2. Diagram for determining the characteristic values of the pressure p_c^i in the pump ($i=1$) and generator ($i=2$) modes. The shaded areas are equal.

zontal segment on the corresponding external characteristics. The bifurcation value p_c^i itself is determined in the pump ($i=1$) and generator ($i=2$) modes from a system of three nonlinear equations for the variables p_c^i, v_1^i, and v_2^i:

$$\int_{\mathrm{v}_1^i}^{\mathrm{v}_2^i} F(\mathrm{v},p_c^i,\varepsilon,j)d\mathrm{v}$$

$$= \frac{|\mathrm{v}|\mathrm{v}^2}{3} + \frac{j^2\ln(\sqrt{\varepsilon^2(1-\mathrm{v})^2+1})}{\varepsilon^2} + p_c^i \mathrm{v}\Big|_{\mathrm{v}=\mathrm{v}_1^i}^{\mathrm{v}=\mathrm{v}_2^i} = 0,$$

$$F(\mathrm{v}_k^i,p_c^i,\varepsilon,j) = p_c^i - \frac{j^2(1-\mathrm{v}_k^i)}{\varepsilon^2(1-\mathrm{v}_k^i)^2+1} + \mathrm{v}_k^i|\mathrm{v}_k^i| = p_c^i$$

$$- P(\mathrm{v}_k^i) = 0; \quad k=1,2. \tag{4}$$

System (4) is an analytical criterion for the existence of a cell (Fig. 3b) on the phase plane (v,v') of system (3), bounded by the separatrices joining the saddle points. Conditions (4) admit a simple geometric interpretation. They reflect the equality of the areas of the shaded figures (Fig. 2) bounded by the segment $p=p_c$ and by the external characteristic of the uniform flow.

At large values of j one can obtain the asymptotic behavior $\hat{p}_c^i$ of the solution of equation (4). We shall seek $\hat{p}_c^i$ in the form of a series:

$$\hat{p}_c^i = c_1^i j^m \ln^n j + c_2^i j^r \ln^l j \ln^s(\ln j) + \dots, \quad r \leqslant m.$$

Here $4/3 \leqslant m \leqslant 2$, and $n \leqslant 0$ for $m=2$ and $n \geqslant 0$ for $m=4/3$, since it is only in this case that the equation $F(\mathrm{v},p,\varepsilon, j) = 0$ for $j \to \infty$ can have three real roots v_k^i; $i=1,2$; $k=0,1,2$. Then the last two equations of system (4) give in the leading approximation

$$\mathrm{v}_1^i = (-1)^i \sqrt{(|p_c^i|)} + O(j^2/p_c^i),$$

$$\mathrm{v}_2^i = 1 - p_c^i/j^2 + O((p_c^i)^3/j^6). \tag{5}$$

We note that the cases $m=2$, $n=0$ and $m=4/3$, $n=0$ must in general be considered separately. Such an analysis shows that these cases are not realized. When these limits of integration are substituted into the first equation of system
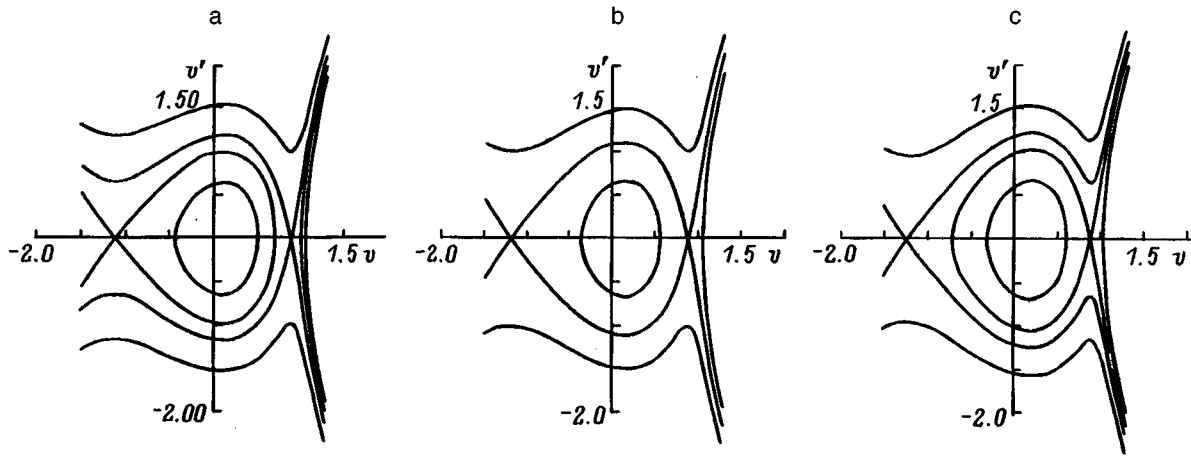
FIG. 3. The formation of a cell on the phase plane of system (3) for $p=p_c$ (pump mode): a: $p<p_c$, b: $p=p_c$, c: $p>p_c$; $j^2=40$, $\varepsilon=4$.

(4) and only the terms of the highest order in j are kept, with allowance for the *a priori* estimates for m in (5), one obtains the following equation for determining the leading terms in the asymptotic expression for $\hat{p}_c^i$:

$$\frac{j^2\ln|\hat{p}_c^i|}{2\varepsilon^2} - \frac{2}{3}|\hat{p}_c^i|^{\frac{3}{2}} + O(j^2) = 0,$$

which gives

$$\hat{p}_c^i = (-1)^{i+1}\left(\frac{j}{\varepsilon}\right)^{4/3}\ln^{2/3}j\left(1 + \frac{\ln(\ln j)}{3\ln j} + O\left(\frac{1}{\ln j}\right)\right);$$

$$i=1, 2. \tag{6}$$

Figure 4 shows the dependence of $p_c^1/\hat{p}_c^1$ on j for the case under consideration, illustrating the approach of the numerical solution of system (4) to the asymptotic behavior (5).

The asymptotic solutions of problems of the type (3) at small values of the singular perturbation parameter ν in the neighborhood of the bifurcation point p_c are constructed in Ref. 10. For system (3) they have the form

$$v(\varphi) = \begin{cases} v_j^i, & 0 \leqslant \varphi \leqslant \dfrac{2\pi\theta_j}{1+\theta_j}, \\ v_k^i, & \dfrac{2\pi\theta_j}{1+\theta_j} < \varphi < 2\pi, \end{cases}$$
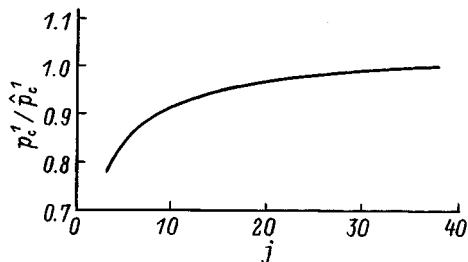
where

FIG. 4. Approach of the numerical solution of system (4) to the asymptotic expression (5); $\varepsilon=4$.

$$i=1, 2; (j,k) = \begin{cases} (1, 2) & p \leqslant p_c, \\ (2, 1) & p < p_c. \end{cases}$$

The specific value of the constant θ_j $(0<\theta_j\leqslant\lambda_j/\lambda_k)$ is determined by the relation between the parameters ν and $|p-p_c|$ as they go to zero.[10] In particular, for $\nu\to0$ with $p\equiv p_c$ one has $\theta_j=\lambda_j/\lambda_k$, where λ_j and λ_k are the absolute values of the eigenvalues of equation (3) linearized about the points v_j and v_k, i.e.,

$$\lambda_m = \frac{1}{\sqrt{\nu}}\sqrt{2|v_m| + \frac{j^2(1-\varepsilon^2(1-v_m)^2)}{(\varepsilon^2(1-v_m)^2+1)^2}}; \quad m=j,k.$$

Then the expression for the dimensionless flow rate at the points p_c^i to a first approximation for $\nu\to0$ have the form

$$q^i(p_c^i) = \frac{1}{2\pi}\int_0^{2\pi}v(\varphi)d\varphi = \frac{v_1^i\lambda_2^i + v_2^i\lambda_1^i}{\lambda_1^i+\lambda_2^i}.$$

Using the technique of Ref. 10, one can also estimate the value of the derivative dp^i/dq on the "plateau" of the external characteristic of the device:

$$\frac{dp^i}{dq} = \frac{R^i}{\nu}\exp\left(-\frac{\Lambda^i\delta}{\nu}\right) + O\left(\frac{\exp\left(-\dfrac{\Lambda^i\delta}{\nu}\right)}{\nu}\right), \tag{7}$$

where

$$i=1, 2; \quad 0<\delta\leqslant1, \quad \Lambda^i = \frac{\lambda_1^i\lambda_2^i}{\lambda_1^i+\lambda_2^i},$$

and R^i is some constant.

The above estimate explains completely how MHD induction devices can come to operate as ideal pressure sources. Indeed, according to Eq. (7), in the neighborhood of the point $p=p_c$ the slope of the external characteristic is an exponentially small quantity over a wide range of flow rates q.

ABSENCE OF INTERACTION BETWEEN JETS

Let us now take the limit $\nu \to 0$ first. Then in the leading approximation we have the following problem:

$$\varkappa x_1'' = x_1 + \varepsilon x_2 - \varepsilon x_2 |j^2 x_1/\varepsilon - p|^{0.5} \mathrm{sign}(j^2 x_1/\varepsilon - p)$$

$$= f_1(x_1, x_2, \mathrm{v}, p, \varepsilon, j),$$

$$\varkappa x_2'' = x_2 - \varepsilon x_1 + \varepsilon x_1 |j^2 x_1/\varepsilon - p|^{0.5} \mathrm{sign}(j^2 x_1/\varepsilon - p) + 1$$

$$= f_2(x_1, x_2, \mathrm{v}, p, \varepsilon, j),$$

$$x_i(0) = x_i(2\pi), \quad x_i'(0) = x_i'(2\pi), \quad i = 1, 2. \qquad (8)$$

Here the ''bifurcation scenario'' unfolds not on a plane but in a four-dimensional phase space. Nevertheless, the system of equations (8) in the appropriate intervals of p also has three points of rest (x_{1k}^i, x_{2k}^i), $k = 0, 1, 2$; $i = 1, 2$, two of which are hyperbolic ($k = 1, 2$), and for $p = p_c^i$ are joined by a heteroclinic trajectory. As before, we will seek an asymptotic expression for p_c in the form (5) with the same restrictions on the values of m and n. Although in the general case we can no longer write a nonlinear system analogous to (4), we can nevertheless obtain the asymptotic behavior of p_c^i for large j. The structure of the system is such that for $j \to \infty$ the heteroclinic trajectory which we seek lies, to leading order, on the surface $f_1(x_1, x_2, p, \varepsilon, j) = 0$. One can show this by making a nonlinear change of variables and transforming from the variable x_1 to $\Phi = f_1(x_1, x_2, p, \varepsilon, j)$. After such a substitution the first equation of system (8) in the leading approximation as $j \to \infty$ takes the form $\Phi = 0$ everywhere except in a small neighborhood (with radius of the order of $1/j$) of the straight line $x_2 = 0$. Then the heteroclinic trajectory for $j \to \infty$ is described approximately by the second-order system

$$\varkappa x_2'' = x_2 - \varepsilon x_1(x_2) + \varepsilon x_1(x_2)|j^2 x_1(x_2)/\varepsilon$$

$$- p|^{0.5}\mathrm{sign}(j^2 x_1/\varepsilon - p) + 1 = F(x_2, p, \varepsilon, j), \qquad (9)$$

where

$$x_1 = \varepsilon x_2$$

$$\times \begin{cases} j^2 x_2/2 + \sqrt{j^4 x_2^2/4 - p - j^2 x_2} - 1, & x_2 \leqslant -p/j^2, \\ -[j^2 x_2/2 + \sqrt{j^4 x_2^2/4 + p + j^2 x_2} + 1], & x_2 > -p/j^2 \end{cases}$$

for the pump mode and

$$x_1 = -\varepsilon x_2 [j^2 x_2/2 + \sqrt{j^4 x_2^2/4 + p + j^2 x_2} + 1]$$

for a generator. This system turns out to belong to the same class as system (3), according to the classification scheme investigated in Ref. 10. For this system the existence condition of a cell on the phase plane — the system of equations, analogous to (4), for determining p_c — has the form

$$\int_{x_{21}^i}^{x_{22}^i} F(x_2, p_c^i, \varepsilon, j) dx_2 = 0; \quad F(x_{2k}^i) = 0,$$

$$k = 1, 2; \quad i = 1, 2. \qquad (10)$$

Strictly speaking, since one of the roots in the leading approximation falls in a small neighborhood of zero ($x_{22}^i = -1/\varepsilon^2 p$), one must put c/j, where c is some constant,

as the upper limit of integration in (10) instead of x_{22}^i. This is because after it has fallen into the neighborhood of $x_2 = 0$ (within a radius of the order of $1/j$) the motion along the heteroclinic trajectory will no longer be described by system (9). However, this refinement does not affect the final result, viz., the value of p_c^i in the leading approximation.

For the analysis it is convenient to rewrite system (10) in terms of the velocity, using the following relations which are valid in this case:

$$x_1 = \frac{(p + \mathrm{v}|\mathrm{v}|)\varepsilon}{j^2}; \quad x_2 = \frac{p + \mathrm{v}|\mathrm{v}|}{(\mathrm{v} - 1)j^2}.$$

After a change of variables we will have

$$\int_{\mathrm{v}_1^i}^{\mathrm{v}_2^i} (P(\mathrm{v}) - p) \frac{\varepsilon^2 (\mathrm{v} - 1)^2 + 1}{j^4 (1 - \mathrm{v})^3} [(\mathrm{v} - 2)|\mathrm{v}| - p] d\mathrm{v} = 0. \qquad (11)$$

Then, as in the foregoing case of infinitely wide channels, we substitute (5) into (11) and equate to zero the terms of the highest order in j. For $j \to \infty$ we obtain for the pump and generator modes, after some manipulations,

$$p_c^i = (-1)^{i+1} \frac{j^2}{\sqrt{2}\varepsilon \sqrt{\ln j}} + O\left(\frac{j^2}{\sqrt{\ln j}}\right). \qquad (12)$$

Thus we have obtained two asymptotic formulas suitable for estimating the characteristic value of the pressure produced by an MHD device after the loss of stability of the uniform flow at high supply currents: formula (6), which is valid for wide channels, and formula (12), which is suitable for the case of low reduced viscosity. It should be emphasized that the foregoing analysis has revealed the main cause of the horizontal ''plateau'' on the branched external characteristic of MHD devices: the smallness of $\varkappa$ and ν. The structure of the initial system (2) is such that the only additional requirement for the possible onset of the ''ideal pressure source'' regime is that the parameter ε must be greater than 1. If this condition is met, then for sufficiently small $\varkappa$ and ν one will observe all the above effects related to the formation of secondary flows.

[1] A. P. Klavinya, O. A. Lielausis, and V. V. Riekstin'sh, *Abstracts of the Eighth Riga Conference on Magnetohydrodynamics* [in Russian], Riga (1975), Vol. 2, pp. 79–81.

[2] A. K. Gailitis and O. A. Lielausis, Magnitnaya Gidrodin., No. 1, pp. 106–111 (1975).

[3] Yu. A. Polovko and É. A. Tropp, Magnitnaya Gidrodin., No. 4, pp. 106–113 (1986).

[4] Yu. A. Polovko, Magnitnaya Gidrodin., No. 3, pp. 81–88 (1989).

[5] Yu. A. Polovko, E. P. Romanova, and É. A. Tropp, Zh. Tekh. Fiz. **66**, 36 (1996) [Tech. Phys. **41**, 315 (1996)].

[6] R. A. Valdmane, I. R. Kirillov, A. P. Ogorodnikov *et al.*, Magnitnaya Gidrodin., No. 3, pp. 98–104 (1982).

[7] R. R. Krishberg, Magnitnaya Gidrodin., No. 2, pp. 140–141 (1981).

[8] O. A. Lielausis, Ya. Ya. Lielpeter, and M. V. Pukis, *Abstracts of the Tenth Riga Conference on Magnetohydrodynamics* [in Russian], Riga (1981), Vol. 2, pp. 16–17.

[9] Yu. A. Polovko and E. A. Tropp, *Proceedings of the International Conference on Energy Transfer in Magneto-Hydrodynamic Flows*, Pamir. Cadarache, France, 1991, pp. 73–78.

[10] L. A. Bakaleĭnikov and Yu. A. Polovko, Diff. Urav. **25**, 903 (1989).

Translated by Steve Torstveit

Electron kinetics in a discharge plasma produced by a focused microwave beam in free space

A. A. Kuzovnikov, V. M. Shibkov, and L. V. Shibkova

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

A study is made of the electron kinetics in a discharge plasma produced by a high-power beam of electromagnetic radiation in the centimeter-wave region under conditions approaching free space, when the dimensions of the chamber are much greater than the wavelength of the microwave radiation. Two regimes of discharge production are investigated: the regime of short microsecond pulses at a repetition rate of 200 Hz, and a single millisecond pulse regime. It is shown that at threshold values of the microwave energy flux density the electron density in the initial stages of discharge formation reaches the critical value, and that the average energy of the electrons is of the order of 1.5–3 eV. © *1997 American Institute of Physics.*
[S1063-7842(97)00306-1]

The end of the 1940s saw the advent of cw microwave generators in the kilowatt range, which made it possible to produce long-lasting microwave discharges at atmospheric pressure and ushered in intensive research in a new area of plasma physics—the physics of microwave discharges. In Ref. 1 the formation and propagation of a discharge in a waveguide in air were investigated in the cw mode of the microwave generator. If by some means or other a discharge was excited at some place in the waveguide, the plasmoid would move toward the generator, counter to the microwave radiation. Approximately 75% of the power of the incident wave was absorbed by the plasma. Such a discharge can be sustained at a microwave power several orders of magnitude lower than is required for breakdown of atmospheric air, and the velocity of the discharge increases monotonically with increasing power. In Ref. 2 a physical interpretation of this phenomenon was given, based on the concepts of discharge propagation in the slow-burning regime. Microwave energy could be supplied to a plasma in another way as well.[3] A dielectric tube of a material transparent in the microwave range, e.g., quartz, was inserted in a rectangular waveguide (TE_{01} mode) perpendicular to its wide wall. A discharge would be ignited in the tube in the region where it crossed the waveguide. The plasma was sustained through the dissipation of energy from the traveling electromagnetic wave in it. The heat released was carried away by a gas blown through the tube or through thermal conduction to the cooled walls of the tube.

However, a specific feature of the microwave region makes it possible to do experiments in free space, with no contact of the discharge with the walls of the chamber and electrodes. In Ref. 4 a discharge was produced in a cavity in which a standing electromagnetic wave of the TH_{01} mode was excited. The structure of the field here is such that the electric field on the axis is directed along the axis and varies sinusoidally, while along the radius the field falls off with distance from the axis. The discharge was ignited on the axis in the region of maximum field. The plasma was drawn out along the electric field vector, and at high powers had the form of a pinch. In order to prevent the discharge from spreading under the influence of the buoyancy force and to lend stability to the plasma pinch, the gas was swirled in the cavity.

Progress in microwave electronics has made it possible to create a new form of discharge — an electrodeless microwave discharge in a focused beam of electromagnetic energy in free space.[5–21] The strong spatial nonuniformity of the energy release near electrodes is one of the main causes of discharge instability.[22] Going to electrodeless systems can improve the stability and increase the energy deposition in a gas discharge. A non-self-sustained microwave discharge in free space has a relatively high spatial uniformity of the energy deposition and affords the possibility of independently regulating the plasma parameters over relatively wide limits. The absence of contact with solids also makes the use of non-self-maintained microwave discharges extremely promising for obtaining highly pure substances in plasma chemistry.

In a gas acted on by a high-power focused beam of electromagnetic radiation in the microwave range, a breakdown of the gas occurs in the focal region of the discharge chamber when the electric field exceeds a threshold value. After the breakdown, the plasma formed at the focus of the beam begins to absorb intensely the energy delivered to the discharge in such a way that the electrons acquiring energy from the microwave field give it up in collisions with neutral gas molecules. The energy released in the gas is distributed over all the degrees of freedom (vibrational, rotational, translational, electronic excitation of molecules, dissociation, etc.), and, in turn, depending on the conditions (which can be controlled), gives rise to disequilibrium in the plasma and to various gasdynamic perturbations. Refinement of the theoretical models used to describe the spatial and temporal evolution of a discharge in wave fields requires reliable information about the parameters of the charged and neutral components of unsteady, nonequilibrium plasmas over a wide range of external conditions. To obtain such information, it is necessary to carry out comprehensive measurements of the main parameters of an electrodeless microwave discharge by contactless methods, if possible. The use of

optical diagnostics can yeild data on the parameters of a freely localized microwave discharge in air with high spatial and time resolution and best meets the above requirements.

For measurements of the electron density and temperature in this study we have used a modification[19] of the method proposed in Ref. 23, wherein n_e and T_e in a nonisothermal nitrogen plasma are determined simultaneously from a measurement of the absolute intensities of the emission in the bands of the second positive and first negative systems of nitrogen under conditions such that the excitation of these bands occurs by direct electron impact and the states are populated through radiative transitions. To find the absolute values of the populations of the investigated levels a photoelectric method is used to determine the radiance of the emission in comparison with that of a standard continuum source.

The object of study, a microwave discharge in a focused beam of electromagnetic waves, was produced on two test benches under conditions approaching free space, the dimension R of the chamber being much larger than the wavelength λ of the microwave radiation. The sources of microwave energy used were a pulsed magnetron on the first bench and centimeter-wave klystrons on the second bench. The magnetron oscillator of the first bench had the following characteristics: wavelength $\lambda_2 = 2.4$ cm, power delivered in the microwave tract $W \leqslant 600$ kW, and duration of the microwave pulse at half amplitude $\tau_u = 5-200$ μs. The oscillator could operate either in the single-pulse mode or in a frequent-pulse mode with a reciprocal duty factor of 1000.

On the first bench the microwave energy entered the discharge chamber along a waveguide tract. To avoid electrical breakdown of the waveguide the latter was filled with sulfur hexafluoride to pressures of up to 6 atm. The entire tract was sealed. The microwave energy was coupled into the chamber through a cone junction. The rectangular waveguide tract was terminated by a transition to a circular cross section. From this element of the waveguide tract a wave of type TE_{11} was directed into the cone transition. The funnel of the cone was terminated by a polystyrene lens 0.6 m in diameter and having a focal length of 0.8 m. The discharge was produced in a cylindrical barochamber, the vacuum system providing a wide range of pressures, from 10^{-5} to 760 torr, for doing the experiments of interest. The focusing lens doubled as a vacuum seal between the chamber and waveguide tract. The inner diameter of the vacuum chamber was 1 m, and its length was 3 m. The microwave discharge was obtained in a traveling wave mode. Matching was achieved by mounting a wall of microwave absorbers at the end of the discharge chamber opposite the lens.

The technical parameters and functional capabilities of the multiprofile experimental apparatus of the second bench are described in detail in Ref. 24. The apparatus is a 36-channel transmitting device operating with a symmetric horn antenna array and providing in-phase summation of the radiation in the volume of the experimental chamber in the centimeter wavelength range. The main element of each channel is a 100 kW klystron amplifier. Phasing of the radiation of the individual horns of the antenna array occurs within the volume of a horizontally situated cylindrical chamber 1 m in diameter and 3 m long. The radiation is
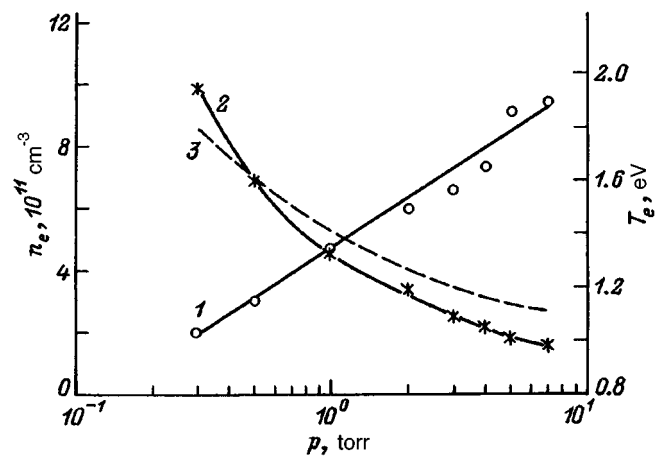

FIG. 1. Pressure dependence of the electron density (1) and temperature (2 experiment, 3 calculation) in the plasma of a freely localized microwave discharge in nitrogen ($\nu \simeq 7$ GHz).

coupled into the chamber through a radio-transparent polyethylene window 1 m in diameter on the side of the chamber facing the antenna array. A traveling wave mode of the radiation in the chamber is provided by shielding the metallic parts of its inner surface by a radio-absorbing device. The klystron oscillators have the following characteristics: variability of the duration of the microwave pulses in the range 1–800 μs, capability of operating either in a single-pulse mode or in a repeated-pulse mode at a repetition rate of 0.5–10 Hz; angle of convergence of the beam 50°; radius of the focal region of the order of wavelength (the frequency of the microwave radiation $\nu \simeq 7$ GHz); energy flux density in the focal region $S \leqslant 10^5$ W/cm^2.

Measurements of the electron density and temperature in the plasma of the ''traveling'' microwave discharge (on the second bench $\nu \simeq 7$ GHz) were made in ''especially pure'' grade nitrogen. The discharge chamber was pumped down to pressures of less than 10^{-3} torr, after which nitrogen was admitted to the required pressure. Figure 1 shows the measured electron density and temperature as functions of the nitrogen pressure at the threshold values of the microwave power delivered to the discharge chamber. It is seen that under the conditions of the experiment in ''especially pure'' nitrogen the electron density reaches the critical value n_{ec}.

Our results are consistent with the data of a numerical calculation[11] in which the evolution of the initial stage of the discharge was described using the model of Ref. 12, in which the field of an axisymmetric wave beam $E(r,z,t)\exp[i(\omega t - kz)]$ is described by a parabolic equation which takes into account refraction by small angles, transverse diffusion, and absorption, and the electron density $n_e(r,z,t)$ is described by an extremely simple balance equation with a specified amplitude dependence of the ionization frequency $\nu_i(|E|) \simeq |E|^\beta$ and a specified effective loss frequency $\nu_a =$ const. In the calculations[11] the radial coordinate r and longitudinal coordinate z were taken in units of the characteristic transverse scale a and longitudinal scale $l_F = ka^2$ of the focal region of the unperturbed beam, the angle of convergence of which was $\Theta = (ka)^{-1}$; $n_{ec} = m(\omega^2 + \nu_e^2)/4\pi e^2$ is the critical density, and ν_e is the
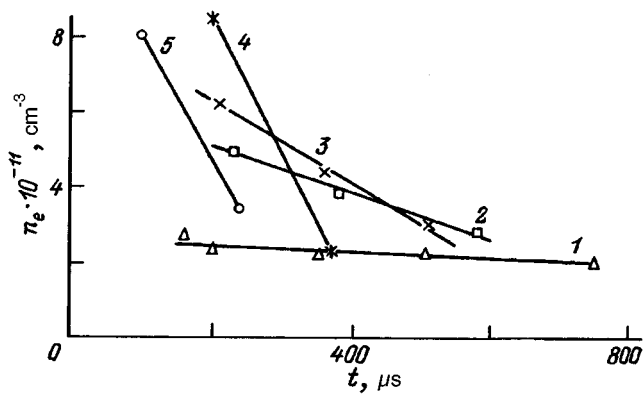
FIG. 2. Time dependence of the electron density in the plasma of a traveling microwave discharge in nitrogen: $\nu = 7$ GHz, $p = 0.3$ (1), 1 (2), 2 (3), 3 (4), 7 torr (5).

FIG. 3. Time dependence of the electron density in the plasma of a traveling microwave discharge in nitrogen: $\nu = 7$ GHz, $p = 0.3$ (1), 1 (2), 2 (3), 3 (4), 7 torr (5).

frequency of collisions between electrons and gas molecules. The initial and boundary conditions were specified in such a way that in the pre-breakdown state ($t = 0$) the density was low, $n_e(r,z,t) = n_0 = $ const, and the beam was Gaussian and focused to a specified distance z_0 from the boundary $z = 0$. Under the conditions of our experiment ($\Theta \simeq 0.3$), estimates according to the data of Ref. 11 give a ratio $n_{max}/n_{ec} \simeq 1$, in good agreement with our results.

When the pressure is increased from 0.3 to 7 torr the electron density increases from $\sim 2 \times 10^{11}$ to $\sim 10^{12}$ cm^{-3}, and the electron temperature T_e decreases from ~ 2 to ~ 1 eV. Since this range of pressures corresponds to the minimum of the breakdown curve, as the pressure is increased the field in the plasma remains practically constant while E_{eff}/n decreases, and this leads to a drop in the average energy of the electrons, as is shown by the experiment (curve 2 in Fig. 1) and by a calculation in the nonstationary kinetic model[17,25] (curve 3).

It was not possible to elucidate how the electron density and temperature vary over the course of the microwave power pulse in some chosen single plasmoid produced on the second bench, since the plasmoid, after being formed at the focus, does not remain in place but moves toward the focusing antenna, leaving the field of view of the spectral apparatus. The optical system that we use projects a one-to-one image (with unit magnification) of the plasmoid onto the input slit of the monochromator. Therefore, only the radiation from regions of the plasmoid with an area equal to the input slit will strike the photocathode of the photomultiplier. As the plasmoids move in the direction counter to the energy propagation in the focused microwave beam, they arrive successively at a given place in the discharge region, so that it is possible to find out how the electron density and temperature behave at different times during the microwave pulse at the same place in the discharge region, but in different plasmoids. The results of the measurements for different nitrogen pressures $p = 0.1$–10 torr are given in Figs. 2 (n_e) and 3 (T_e). It is seen from these figures that the electron density decreases toward the end of the pulse, while the electron temperature increases, and both n_e and T_e change faster in time at higher pressures than at lower pressures.
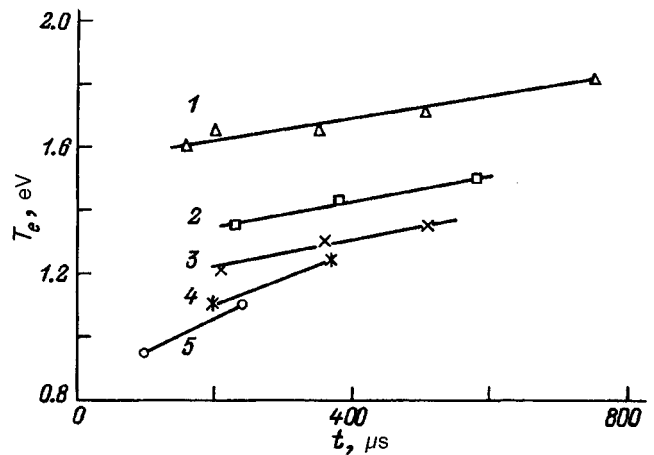
The growth of the electron temperature during the pulse

can be explained as follows. In the initial moments of the microwave pulse the lower vibrational levels of the nitrogen are still weakly populated, and superelastic collisions play only a small role; the electron energy goes mainly to excitation of these levels, and T_e therefore remains low. As time goes on, as the microwave pulse is lengthened, it is in principle possible for the population of the lower vibrational levels to become rather large and for superelastic collisions to begin to play a noticeable role; this can lead to growth of the electron temperature, as the results of a mathematical modeling show.[25] The decrease in the electron density toward the end of the microwave pulse can be attributed to effects of heating of the gas.[16] Over a time of the order of hundred of microseconds the gas can heat up, and the nitrogen molecules begin to escape from the region occupied by the plasma. In this way the region in which the microwave discharge exists becomes depleted of neutral molecules, with the result that the electron density decreases.

Figure 4 gives the experimental dependence of the electron density and temperature in the plasma of a "traveling"
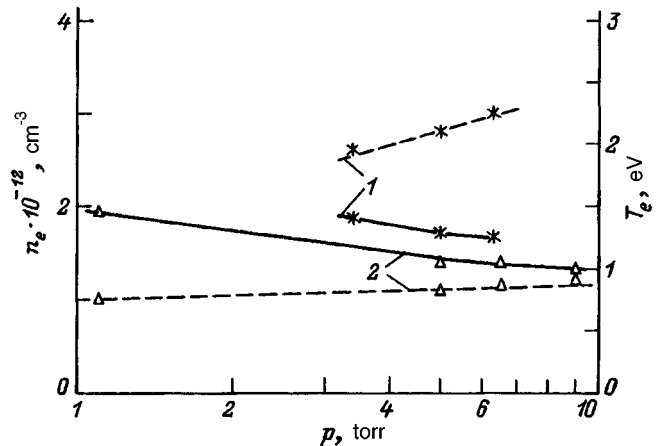

FIG. 4. Electron density (dashed curves) and electron temperature (solid curves) versus the gas pressure in the plasma of an electrodeless microwave discharge in nitrogen (1) and air (2); $\lambda = 2.4$ cm.
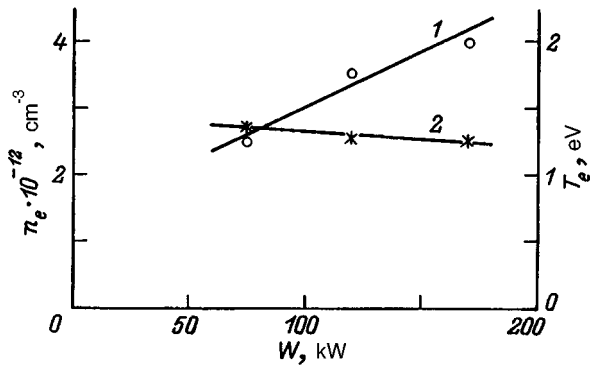
FIG. 5. Electron density (*1*) and electron temperature (*2*) versus the microwave power level in the plasma of a freely localized discharge in nitrogen at $p = 5$ torr ($\lambda = 2.4$ cm).

microwave discharge as obtained on the first bench. The wavelength of the microwave radiation was 2.4 cm, the pulse duration was 5 μs, and the repetition rate was 200 Hz. Measurements were made at the threshold level of the delivered microwave power. It is seen that, as in the previous case ($\nu \simeq 7$ GHz), the electron density increases with increasing pressure, while the electron temperature falls off. The electron density in the plasma of a ''traveling'' microwave discharge in nitrogen is somewhat higher than in air, a fact which can be attributed to the presence of an efficient channel of electron loss in air owing to their attachment in reactions forming negative oxygen ions.

On the first bench we also measured n_e and T_e for various levels of the power delivered to the discharge. Figure 5 shows plots of $n_e(W)$ and $T_e(W)$. It is seen that the electron density increases with increasing power, while the temperature falls off slightly.

From a comparison of the experimental values of n_e with the values of the critical density n_{ec} for the investigated ranges of microwave radiation we see that the measured values of the electron density in the plasma of a freely localized discharge produced by a focused microwave beam is of the order of n_{ec}. The values obtained for the average electron energy $\varepsilon \simeq 1.5$–3 eV are in good agreement with the average energy corresponding to the electron energy distribution functions calculated in Ref. 25 from the time-dependent Boltzmann transport equation for values of the reduced electric field E_{eff}/n corresponding to the experimental conditions.

[1] W. Beust and W. L. Ford, Microwave J. MTT. **10**, 91 (1961).
[2] Yu. P. Raĭzer, Zh. Éksp. Teor. Fiz. **61**, 222 (1971) [Sov. Phys. JETP **34**, 114 (1972)].
[3] V. M. Batenin, I. I. Klimovskiĭ, P. V. Lysov, and V. N. Troitskiĭ, *Microwave Plasma Generators* [in Russian], Énergoatomizdat, Moscow (1988).
[4] P. L. Kapitsa, Zh. Éksp. Teor. Fiz. **57**, 1801 (1969) [Sov. Phys. JETP **30**, 973 (1970)].
[5] A. G. Litvak (ed.), *Radio-Frequency Discharges in Wave Fields* [in Russian], IPF AN SSSR, Gorkiĭ (1988).
[6] A. G. Litvak (ed.), *Strong Microwave in Plasmas*, IAP AS USSR, Nizhny Novgorod (1991).
[7] G. M. Batanov, S. I. Gritsinin, I. A. Kossyĭ *et al.*, Tr. Fiz. Inst. Akad. Nauk SSSR **160**, 174 (1991).
[8] G. A. Askar'yan, G. M. Batanov, I. A. Kossyĭ *et al.*, Fiz. Plazmy **17**, 85 (1991) [Sov. J. Plasma Phys. **17**, 48 (1991)].
[9] V. V. Gil'denburg and S. V. Golubev, Zh. Éksp. Teor. Fiz. **67**, 89 (1974) [Sov. Phys. JETP **40**, 46 (1974)].
[10] V. V. Gil'denburg and V. E. Semenov, *Nonlinear Waves. Structure and Bifurcations* [in Russian], Nauka, Moscow (1987).
[11] A. L. Vikharev, V. B. Gil'denburg, S. V. Golubev *et al.*, Zh. Éksp. Teor. Fiz. **94**(4), 136 (1988) [Sov. Phys. JETP **67**, 724 (1988)].
[12] V. B. Gildenburg, A. G. Litvak, and A. D. Yunakovsky, J. Phys. (Paris) **40**, Colloq. C7, 215 (1979).
[13] V. G. Brovkin and Yu. F. Kolesnichenko, Pis'ma Zh. Tekh. Fiz. **17**(1), 58 (1991) [Sov. Tech. Phys. Lett. **17**, 22 (1991)].
[14] A. S. Zarin, V. N. Kulikov, and V. E. Mitsuk, Pis'ma Zh. Tekh. Fiz. **8**, 1373 (1982) [Sov. Tech. Phys. Lett. **8**, 590 (1982)].
[15] V. V. Zlobin, A. A. Kuzovnikov, and V. M. Shibkov, Vestn. Mosk. Univ. Fiz. Astron. **29**(1), 89 (1988).
[16] A. M. Devyatov, A. A. Kuzovnikov, V. V. Lodinev, and V. M. Shibkov, Vestn. Mosk. Univ. Fiz. Astron. **32**(2), 29 (1991).
[17] V. M. Shibkov, A. F. Aleksandrov, and A. A. Kuzovnikov, *Ball Lightning* [in Russian], Khimiya, Moscow (1994), pp. 136–150.
[18] N. A. Armand, A. S. Zarin, V. E. Mitsuk *et al.*, in *Radiophysics* [in Russian], G. I. Batskikh, ed., MRTI AN SSSR, Moscow (1991), pp. 135–142.
[19] V. M. Shibkov, *Proceedings of the Twentieth International Conference on Phenomena in Ionized Gases (XX ICPIG)*, Riga (1991), Contributed Paper No. 5, pp. 1137–1140.
[20] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **55**, 389 (1985) [Sov. Phys. Tech. Phys. **30**, 228 (1985)].
[21] L. P. Grachev, I. I. Esakov, G. I. Michin, and A. B. Fedotov, Zh. Tekh. Fiz. **59**(10), 149 (1989) [Sov. Phys. Tech. Phys. **34**, 1181 (1989)].
[22] M. N. Andreeva, I. G. Persiantsev, V. D. Pis'mennyĭ *et al.*, Fiz. Plazmy **3**, 1383 (1977) [Sov. J. Plasma Phys. **3**, 770 (1977)].
[23] M. G. Berdichevskiĭ and V. V. Marusin, Zh. Prikl. Spektrosk. **18**, 1055 (1973).
[24] G. I. Batskikh and Yu. I. Khvorostyanoĭ, Radiotekh. Elektron. **37**, 311 (1992).
[25] A. F. Alexandrov, A. A. Kuzovnikov, V. M. Shibkov, and A. S. Zarin, *Strong Microwave in Plasmas*, A. G. Litvak, ed., IAP RAS, Nizhny Novgorod (1994), Vol. 1, pp. 281–286.

Translated by Steve Torstveit

Existence regions of regimes with anomalously high rates of plasma generation in grid plasma switch elements

N. I. Alekseev, V. B. Kaplan, A. M. Martsinovskiĭ, and I. I. Stolyarov

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

Results are presented from a systematic investigation of the conditions under which anomalously high rates of plasma generation are observed in the anode region of a low-voltage Knudsen arc in grid plasma switch elements. The phenomenon develops over a wide range of currents and switched voltages under conditions for which the plasma density in the cathode–grid region is noticeably higher than the density in the anode region. and its onset is characterized by a pronounced pressure threshold of $\sim 2 \times 10^{-2}$ Torr. The results are analyzed from the standpoint of the possible mechanisms for anomalous plasma generation — collisional nonresonance diffusion of electrons in velocity space, leading to enrichment of the distribution function in fast particles, and the collapse of Langmuir waves in the gap at the high energies of the beam produced when the wires of the grid are bridged by the quenching pulse and and the current is blocked. © *1997 American Institute of Physics.*
[S1063-7842(97)00406-6]

INTRODUCTION

Plasma switch elements with total grid control of the current (PSEs) are gas-filled (cesium or cesium–barium) triodes operating in the low-voltage Knudsen arc regime (electrode separation ~ 1 mm, Cs vapor pressure $P_{Cs} \sim (0.5$–2×10^{-2} Torr).[1,2] In these PSEs a negative voltage pulse V_g^1 with a steep rise (~ 0.1 μs) and an amplitude of tens of volts is applied to the grid, causing quenching of the discharge and interruption of the current at plasma densities of up to 10^{13}–10^{14} cm^{-3} and currents of up to ~ 10 A/cm^2. An anomalous non-quenching effect (AN) arises, consisting in the fact that in certain discharge regimes an increase in $|V_g^-|$, which ordinarily promotes quenching, begins to act in the opposite way (hence the name).

As we know,[1] when a negative pulse with a steep rise is applied to the grid of a PSE, at the front of the pulse an expansion of the electrode sheaths around the wires of the grid occurs, its electrical transparency decreases, and, accordingly, the anode current I which is passed decreases by an amount ΔI (which we call the initial blocking) from its initial value I_0 (Fig. 1b). When the anode current decreases, the switched voltage E_a (which is ordinarily much larger than the voltage drop V_a^0 across the switch in the conducting state) (Fig. 1a) is redistributed between the device and the anode load R_a: $E_a = V_a + IR_a$. The excess voltage falls across the plane of the grid and produces an electron beam with a current $I_{res} = I_0 - \Delta I$ (curve *2* in Fig. 1b) and an energy of the order of E_a, which causes rapid growth of the plasma density in the anode–grid region owing to heating of the plasma electrons by plasma oscillations excited by the beam.[3,4] If at a specified value of the pulse amplitude the initial blocking of the current is so large that the rate of generation of ions in the anode region is insufficient to restore the passage of current in the cells of the grid, the current is interrupted and the discharge is quenched; otherwise it recovers. The characteristic recovery time of the current is 1

μs — a time of the order of the time required for a bound electron in an excited cesium atom to diffuse over the spectrum in stepwise ionization.

It has been found that at currents $I_0 \geqslant 1$ A/cm^2 and pressures $P_{Cs} \geqslant 1.2 \times 10^{-2}$ Torr the current–voltage (I–V) characteristic exhibits a region in which an increase in the amplitude of the grid pulse above the value $-V_g^q$ necessary for quenching can bring an end to the quenching, starting at a certain level $|V_g^a| > |V_g^q|$. Here the recovery of the current occurs over a very short time, of the order of the blocking time itself (0.1 μs), and, as is shown by probe measurements, is accompanied by an extraordinarily large and rapid growth of the plasma density in the anode region (from 10^{12} to 10^{13}–10^{14} cm^{-3} in ~ 0.5 μs), which is natural to regard as the cause of the restoration of the passage of current.[3,4] Here, as $-V_g^-$ increases, the ionization increases not gradually but abruptly, upon a very small increase in the pulse, which is indicative of a threshold character of the actuation of the mechanism responsible for it. In Ref. 4 a mechanism was proposed which provides the necessary order of magnitude of the ionization rate: enrichment of the electron energy distribution function (EEDF) in fast electrons owing to the nonresonance diffusion in energy space of plasma electrons on Langmuir plasmons. This effect should be observed when the power expended by the beam on excitation of oscillations and the energy density in these oscillations are rather large ($\geqslant 20$ W), as is ordinarily the case for AN. The rapid blocking of the current in the presence of even a small parasitic inductance in the anode circuit leads to a large overshoot of the anode voltage, comparable to or even greater than E_a, which imparts additional energy to the beam.

The goal of the present study was to investigate in detail the existence region of the AN effect as a function of the parameters (pressure, current, position of the working point on the I–V characteristic, the switched voltage, etc.) and the influence of the external conditions of the discharge on the
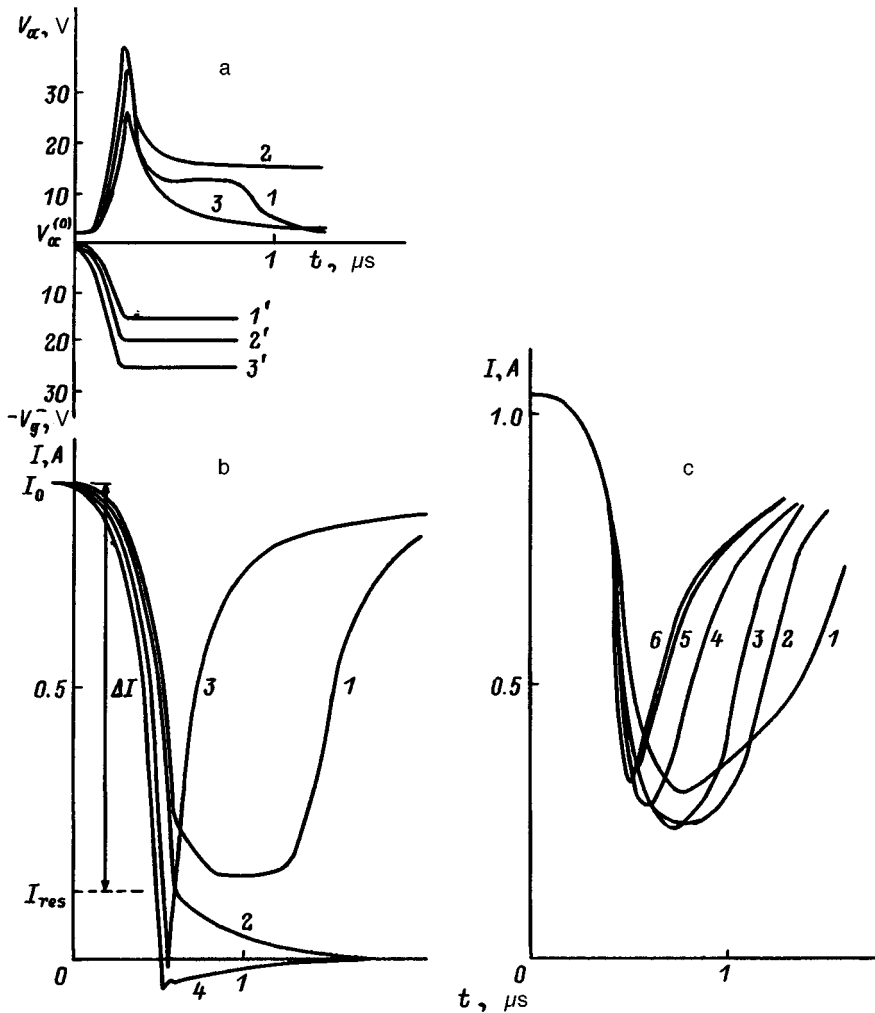
FIG. 1. a: Time dependence of the anode voltage (1–3) when a voltage pulse V_g^- ($1'$–$3'$) is applied to the grid in the state corresponding to point 1 of the I–V characteristic in Fig. 2a; $|V_g^-| = 15.1$ (1), 19.3 (2), 24 V (3), $P_{Cs} = 2.87 \times 10^{-2}$ Torr. b: The corresponding change in the current (curve 4 is for $|V_g^-| = 30$ V). c: The kinetics of the current at point 2 on the I–V characteristic A in Fig. 2a; $|V_g^-| = 15.1$ (1), 22 (2), 29 (3), 35 (4), 41 (5), 52 V (6).

character of the transition to AN, which has not been done before (in the first papers attention was concentrated solely on measuring the plasma parameters during AN and on explaining the observed rates of ion generation). In addition to the need for studying the phenomenology of this effect on purely practical grounds (AN essentially determines the upper limit on the working pressures of a PSE), such research is of physical interest as well.

The point is that in addition to diffusion in energy space on plasmons, there is also another possible way of generating fast electrons in a plasma at high energy densities of the plasma oscillations — modulational instability and collapse of Langmuir waves,[5–7] which give rise to fast particles. However, as previous measurements[3,4] have shown, the procedure of recording the probe characteristics and using them to calculate the EEDF directly at the point of blocking and recovery of the anode current in the AN effect is insufficiently reliable. It is therefore important to know how the AN depends on the external parameters in order to obtain indirect information that might decide in favor of one mechanism or another.

MEASUREMENT PROCEDURES

The experiments were done on an apparatus analogous to that used previously in a study of PSEs.[1] The working

surfaces were the ends of molybdenum cylinders 1 cm in diameter, the (indirectly heated) cathode was coated with a layer of platinum (~ 2 μm) to increase the emission. A grid of molybdenum foil 20 μm thick and having square cells 100×100 μm, with bridges 30 μm wide, was placed a distance of 1.5 mm from the cathode and 1 mm from the anode.

The measurement scheme was the conventional one for PSE studies.[1] An anode voltage in the form of a square pulse 200–300 μs long, with a frequency of ~ 100 Hz and an amplitude of up to 100 V, was applied to the anode through an adjustable (0–200 Ω) low-inductance (~ 0.05 μH) load approximately 1 μs before the positive voltage pulse ($|V_g^+| \sim 5$–20 V, $\tau(V_g^+) \sim 10$ μs) that ignites the discharge. After ignition of the discharge and relaxation to a steady state (100–200 μs) a negative quenching voltage pulse ($0 \leq |V_g^-| \leq 60$ V, $\tau(V_g^-) \sim 20$ μs) was applied to the grid. The grid pulses were applied through a pulse transformer.

For the measurements of the anode and grid currents, inductance-free shunts (tenths of an ohm) were connected between the ground and cathode and between the ground and the grid transformer. The measured quantities (currents and voltages) were recorded by a boxcar integrator (40 ns resolution) or taken off an oscilloscope screen.

EXPERIMENTAL RESULTS

a) Influence of the initial conducting state of the discharge (i.e., dependence on the position of the working point on the I–V characteristic)

We investigated this dependence first because the plasma parameters vary in a most essential way along the I–V characteristic,[8,9] and this, as it turns out, substantially alters the character of the transition to AN.

At the lower point of the vertical segment of the I–V characteristic (curve A in Fig. 2a) the AN effect is ordinarily absent; it begins at a certain critical level of the current (lying somewhat below point 1 on the I–V characteristic in Fig. 2a). In this case as the amplitude of the grid pulse increases, quenching occurs in a certain interval $|V_g^q| < |V_g^-| < |V_g^a|$ (curves 2 and $2'$ in Fig. 1a and 1b), while for $|V_g^q| > |V_g^-|$ we have "ordinary" non-quenching at insufficient pulse amplitudes (curves 1 and $1'$). As the amplitude increases to $|V_g^-| > |V_g^a|$, however, the current, after initially being blocked, begins to recover rapidly (curves 3 and $3'$ in Fig. 1a and 1b). At still higher $|V_g^-|$ quenching reappears (curve 4 in Fig. 1b).

Such behavior can be explained on the basis of the concepts developed in Ref. 3. Our grid pulse generator had a rise time of ~ 0.1 μs independent of the amplitude V_g^-. Therefore, as $|V_g^-|$ increased, the rate of decrease of the electrical transparency of the grid increased, and, along with it, so did the rate of decrease of the current dI/dt, the inductive overshoot of the anode voltage $L_a dI/dt$ (where L_a is the parasitic inductance of the anode circuit), the power of the electron beam, and the energy density of the excited oscillations. If this increase turns out to be sufficient to "turn on" the mechanism of fast ion generation, then the AN effect begins, even if there had been complete blocking of the current and quenching of the discharge at lower amplitudes.[1]

As the anode current increases along the I–V characteristic (point 2 in Fig. 2a) the quenching vanishes, and AN is manifested on increasing $|-V_g^-|$ as a sharp decrease in the recovery time of the current in the case of incomplete initial blocking $\Delta I/I_0 < 1$ (it decreases from 0.5–1 to 0.1–0.05 μs; curves 1–6 in Fig. 1c). Here after the onset of AN the growth in the relative initial blocking $\Delta I/I_0(|V_g^-|)$ slows, and then $\Delta I/I_0$ falls to a certain minimum value $(\Delta I/I_0)^{\min}$, after which, over an appreciable interval of $|V_g^-|$ values, $\Delta I/I_0$ slowly increases (curve 2 in Fig. 2b).

This character of the dependence of $\Delta I/I_0$ on $|V_g^-|$ is maintained as the working point moves further along the I–V characteristic to higher anode currents and voltages (curves 3 and 4 in Fig. 2b); there is a decline in the initial blocking level at which $\Delta I/I_0$ begins to drop and also in $(\Delta I/I_0)^{\min}$. At a certain voltage V_a the dependence of $(\Delta I/I_0)^{\min}$ on V_a, shown by curve B in Fig. 2a, reaches a minimum $(\Delta I/I_0)^{\min}$ and then again increases.[2]

As V_a increases further, the interval of V_g^- in which the initial blocking varies slowly becomes narrower and degenerates into an inflection point (curves 4 and 5 in Fig. 2b). Here at large $|V_g^-|$ complete blocking of the current can again be attained, and quenching is restored. At still larger V_a the AN effect vanishes altogether (curve 6). Quenching
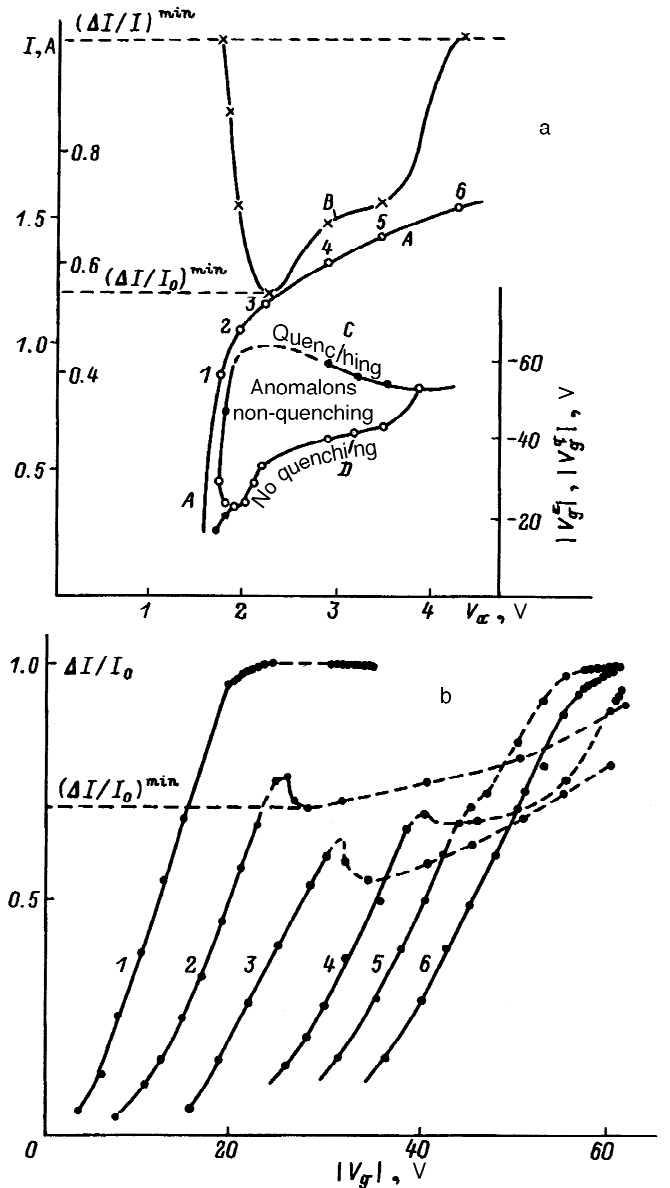

FIG. 2. a: Anomalous non-quenching regions on the I–V characteristic of a discharge ($P_{Cs} = 2.87 \times 10^{-2}$ Torr): A is the I–V characteristic $I(V_a)$ (the current is shown to the left of the left-hand vertical axis); B shows the dependence of the local minimum of the initial blocking curves $(\Delta I/I)^{\min}$ (Fig. 2b) (scale to the right on the left-hand axis) on the point of location along the I–V characteristic; C is the pulse amplitude $|V_g^q|$ corresponding to the transition to the quenching regime ($\bullet$); D is that for transition to anomalous non-quenching ($\bigcirc$); the values of $|V_g^q|$ and $|V_g^a|$ are given on the right-hand axis, $E_a = 15$ V. b: The dependence of the initial blocking of the current ($\Delta I/I_0$) on $|V_g^-|$ for points 1–6 on the I–V characteristic A (Fig. 2a).

and non-quenching regions along the I–V characteristic are illustrated in diagram C–D in Fig. 2a (the $|V_g^-|$ values here are plotted against the scale on the right-hand axis).

Thus there exists a region on the I–V characteristic, ordinarily immediately after the inflection point, in which the processes leading to anomalous non-quenching develop the most rapidly and in which their effect on the passage of current is maximum and corresponds to the minimum initial blocking $(\Delta I/I_0)^{\min}$ (or to the maximum residual current $(I_{res}/I_0)^{\max}$). As we know,[8] this segment corresponds ap-

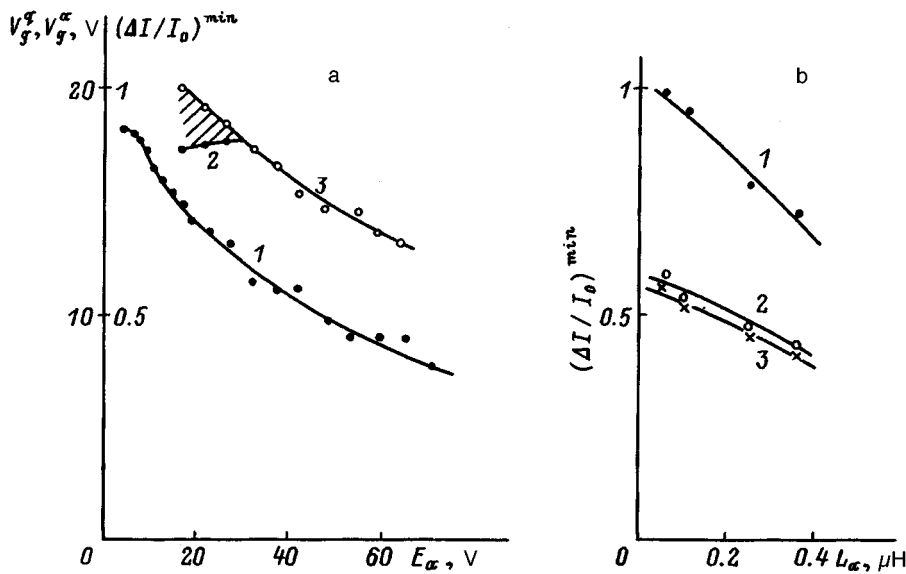FIG. 3. a: Plots of $(\Delta I/I_0)^{\min}$ (1), V_g^q (2), and V_g^a (3) versus E_a; the I–V characteristic corresponds to A in Fig. 2a. b: The function $(\Delta I/I_0)^{\min}(L_a)$ for various E_a (V): 1 — 5.4, 2 — 24, 3 — 49. $P_{Cs}=2.6\times10^{-2}$ Torr, $V_a=22$ V; $j_c^e=1.25$ A/cm^2.

proximately to the maximum plasma density n_1 (close to the limiting degree of ionization) in the cathode–grid region and to essentially the minimum value of the density n_2 in the grid–anode region, i.e., it corresponds to regimes with a maximum difference n_1-n_2, $T_{e1}-T_{e2}$ (where T_e is the electron temperature) and with a maximum magnitude of the negative (retarding for the electrons of the cathode region) jump in potential at the plane of the grid.

b) Influence of the value of the switched voltage

It has been assumed previously[4] that an appreciable value of the switched voltage ($E_a \gtrsim 20$–30 V) is the main factor giving rise to AN, since efficient generation of ions requires a high enough energy of the electron beam injected into the anode region. Studies have shown that although AN actually arises more easily at large E_a, it can appear at any value of the switched voltage — all the way down to 3–4 V. In that case the beam energy is due almost entirely to the inductive overshoot of the anode voltage V_a^{abn}, where V_a^{abn} can be quite small, as low as 15–20 V.

As E_a increases, the relative role of the inductive overshoot decreases, and the conditions for transition to AN become less stringent. AN sets in at lower amplitudes $|V_g^a|$, and the level of the minimum initial blocking decreases (Fig. 3a). For $E_a \gtrsim 20$–30 V anomalous non-quenching can begin at such small values of $-V_g$ that the inductive overshoot immediately prior to the onset of AN is not observed; it is less than E_a and is masked by the growth of the anode voltage due to the redistribution of the voltage between the discharge gap and the load. The beam energy in this case comes mainly from the anode voltage source.

In accordance with the arguments presented above, the influence of the inductance of the anode circuit on AN at large and small E_a is also different: this influence decreases noticeably as E_a increases (Fig. 3b).

c) Influence of the current and pressure

The most convenient parameters for investigation of these influences are again $(I_{res}/I_0)^{\max}$ and $(\Delta I/I_0)^{\min}$, which are "integral" characteristics of the intensity of the processes leading to AN. The typical dependence of $(I_{res}/I^0)^{\max}$ on the anode voltage along the I–V characteristic is given in Fig. 4 for various cathode emission currents j_c^e at a constant cesium pressure. We see that the influence of the current turns out to be rather complicated.

In the region of large currents (curves 3–6 in Fig. 4a and 3–5 in Fig. 4b) an increase in the emission promotes AN, and it begins sooner on the vertical segment of the I–V characteristic, and the maximum on the curve of $(I_{res}/I_0)^{\max}(V_a)$ increases. The point at which the maximum occurs is shifted toward the inflection point of the I–V characteristic. The decrease of the residual current with increasing V_a becomes steeper, so that the interval of anode voltages over which the conditions change from the most favorable for the onset of AN to the complete disappearance of the effect can amount to only 0.2–0.3 V. This is apparently due to the abrupt growth in the plasma density n_2 in the anode region, which can occur upon an increase in the anode voltage V_a in a developed discharge.[9] In fact, as n_2 increases, the conditions for excitation of oscillations worsen on account of the increase in the rate of Coulomb (electron–ion) collisional damping (according to Ref. 9, for $P_{Cs}=1.5\times10^{-2}$ Torr, which corresponds to a gas density $N_{Cs}\cong2\times10^{14}$ cm^{-3} and an electron temperature $T_e=0.5$ eV, the contributions from scattering on ions and atoms become equal at $n\sim3\times10^{12}$ cm^{-3}).

At higher anode voltages quenching of the discharge can occur at a sufficiently large initial blocking of the current, but this requires larger amplitudes V_g^-, since the plasma density in this case approaches 10^{14} cm^{-3} (Ref. 8).

In the region of smaller emission currents ($j_c^e \lesssim 1$ A/cm^2) one observes the opposite picture — the residual current increases with decreasing j_c^e, the maximum shifting
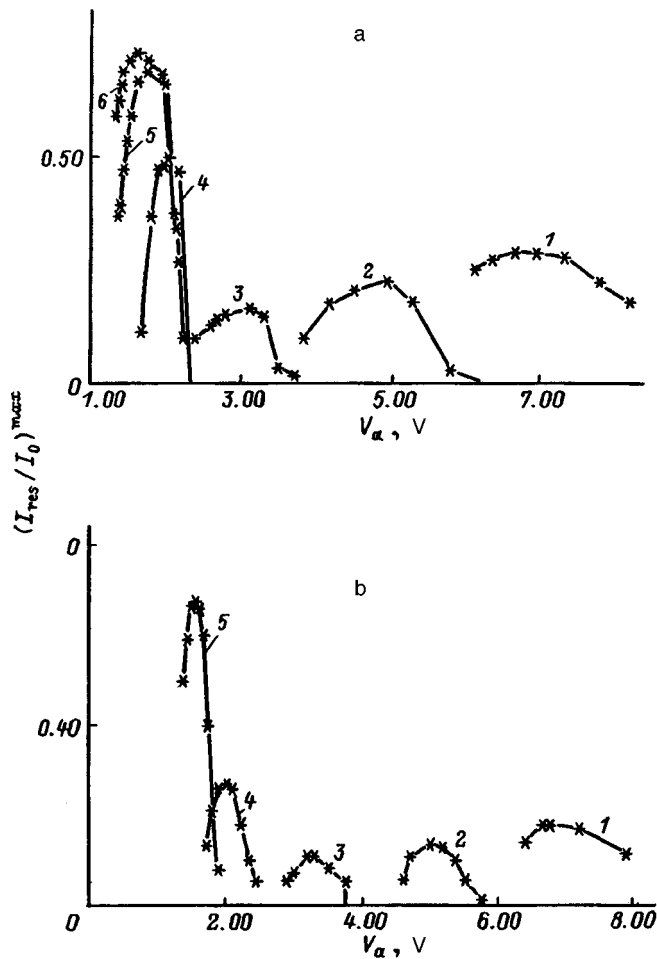
FIG. 4. Dependence of $(I_{res}/I_0)^{max}$ on the point of location along the I–V characteristic (on the voltage across the device) for different cathode emission currents; $E_a = 15$ V. a: $P_{Cs} = 2.90 \times 10^{-2}$ Torr; j_c^e (A/cm^2): 1 — 1.6, 2 — 1.3, 3 — 0.95, 4 — 0.8, 5 — 0.52, 6 — 0.3. b: $P_{Cs} = 2.55 \times 10^{-2}$ Torr, j_c^e (A/cm^2): 1 — 1.62 2 — 1.05, 3 — 1.0, 4 — 0.7, 5 — 0.41.

in the opposite direction, toward larger V_a. Here, as we see from Fig. 4, in the region of the inflection point of the I–V characteristic at small $j_c^e \leq 2.5$ A/cm^2 the AN is completely absent. Since the earlier papers devoted most of their attention to this region as being of the greatest practical interest for PSEs, it was erroneously concluded that AN vanishes at small currents even in the case of large P_{Cs}.
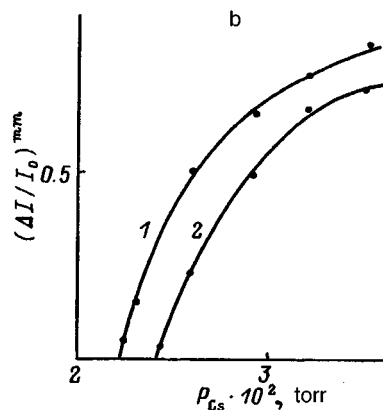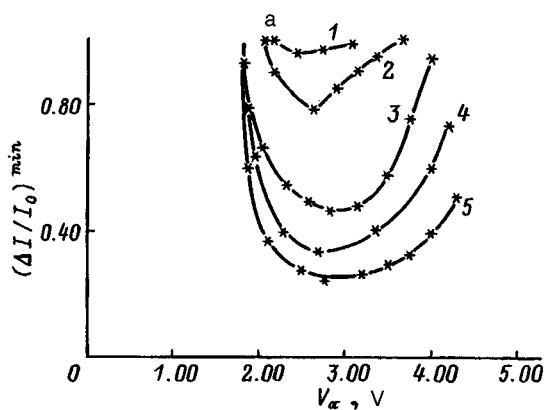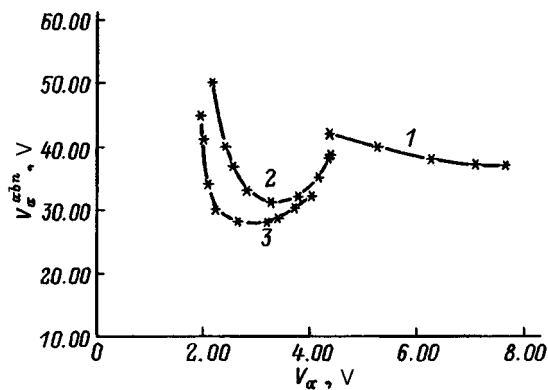

FIG. 5. Dependence of the critical overshoot V_a^{abn} on the point of location along the I–V characteristic for different currents: $P_{Cs} = 2.6 \times 10^{-2}$ Torr; $E_a = 15.5$ V; j_c^e (A/cm^2): 1 — 0.36, 2 — 0.75, 3 — 1.07.

The shift of the maximum of the residual current to larger V_a with decreasing cathode emission can be explained: at smaller j_c^e the transition to regimes with a completely ionized plasma in the cathode region occurs at ever higher anode voltages.[8] Indeed, if one reconstructs the dependence of the residual current on the power $I_0 V_a$ released in the discharge in the conducting state, then all the maxima turn out to correspond to approximately the same power: ~ 2–4 W/cm^2. However, the cause of the increase in the residual current at small j_c^e remains unclear. Nevertheless, this character of the dependence has been observed at all the pressures investigated, while the dependence on the other parameters characterizing the AN, e.g., V_a^{abn} (Fig. 5), have a different character in the regions of large and small currents.

Unlike the situation with the current, the lower boundary of the existence region of AN in terms of the cesium vapor pressure is very sharp. Figure 6a shows the typical behavior of $(\Delta I/I_0)^{min}$ along the I–V characteristic at various pressures near this boundary. Since the range of variation of P_{Cs} is rather large, the small change in the cathode emission with changing pressure was compensated by a small change in the cathode temperature, and the I–V characteristics for all the regimes agree to an accuracy of 1–2%. It is seen that a decrease of only 5–7% in P_{Cs} leads to a transition from well-developed AN (curve 4) to the complete absence of AN. The boundary of the transition is affected only weakly by a

FIG. 6. a: Plots of $(\Delta I/I_0)^{min}$ versus V_a along the I–V characteristic for different pressures P_{Cs} ($j_c^e = 0.9$ A/cm^2): $P_{Cs} \times 10^{-2}$ Torr: 1 — 2.44, 2 — 2.55, 3 — 2.90, 4 — 3.19, 5 — 3.48. b: Plots of $(\Delta I/I_0)^{min}$ versus P_{Cs} for $j_c^e = 0.9$ (1) and 1.45 A/cm^2 (2).

change in the cathode emission current (Fig. 6b).

DISCUSSION OF THE RESULTS AND CONCLUSION

The studies done here have outlined the existence region of the AN regimes — regimes with an anomalously high rate of generation of ions in the anode region after an initial blocking of the current in the PSE. It has turned out that this region is significantly wider than had been assumed previously, so that AN can occur in practically all regimes above a threshold pressure $\sim(2-2.5)\times10^{-2}$ Torr, where this value apparently depends only weakly on the geometry of the discharge gap and control grid. Here AN extends to the region of small cathode emission currents and small E_a, a region which had been considered free of this effect. It has also turned out that the most characteristic regime for AN is not one in which it sets in after total initial blocking of the current (as has been predominantly investigated in earlier studies) but one in which AN arises upon a certain threshold of the initial blocking of the current (which can be much less than complete blocking) and leads to a sharp increase in the current recovery time and to the cessation of further growth in the initial blocking with increasing $|V_g^-|$.

Unfortunately, the results obtained do not give any definite indications as to the cause of the onset of anomalous generation and AN, and the actual mechanism can be elucidated only after an investigation of AN at the level of the plasma parameters. However, some arguments can be offered on the basis of the data obtained.

First of all, the regimes in which AN arises upon incomplete initial blocking (Fig. 1b) show that the processes leading to recovery of the current can begin very rapidly after the start of the blocking — after less than 100–50 ns. This brings an end to the growth of the electrode sheaths around the wires of the grid and to the growth of the initial blocking of the current, in spite of the increase of the negative voltage across the grid.

At threshold pulse amplitudes ($V_g^- \cong V_g^a$), when the processes causing AN are still not too intense, the unblocking of the current can begin with a delay, though earlier than for $|V_g^-| < |V_g^a|$ (curves 2–4 in Fig. 1b). But upon a further increase in $|V_g^-|$ by only a few volts the unblocking of the current, which also takes ~100 ns, begins immediately after its blocking (curves 5 and 6 in Fig. 1c). The fact that the time until the start of the unblocking of the current is so short casts doubt on the attribution of the recovery of the current to the growth in the density in stepwise ionization of the gas by hot plasma electrons. As we have said, stepwise ionization is a fundamentally sluggish process, so that the growth in the density is delayed by several tenths of a microsecond both in experiment (Part II of Ref. 8) and in a calculation (Ref. 9).[3] The "overheating" of the electrons at a large energy density of the Langmuir waves (enrichment of the EEDF in fast particles) does not eliminate the delay of the stepwise generation.

A different picture should appear if the recovery of the current is due to modulational instability and collapse of Langmuir waves. In that case fast electrons and ions are generated in the rarefied plasma regions (cavities) that arise.[6]

These electrons can cause direct delay-free ionization of Cs atoms at the boundary of the electrode sheath of the grid wires. On the other hand, the fast ions can reach the sheath after a time of the order of 100 ns if they are formed at a distance of tenths of a millimeter from it. If there are many fast particles, then they should immediately cause unblocking of the current and a subsequent growth in the density — which is now only an accompaniment of the AN effect. If there are not enough fast particles to "instantly" unblock the current, they can still slow the growth of the electrode sheaths and promote recovery of the current.

If the restoration of the transparency of the cells of the grid is in fact due to a short-wavelength modulational instability of the OTSI type,[10] then it is easy to explain the threshold character of the onset of AN as a function of the amplitude of the negative pulse. The necessary energy level W_e of the electrons for the onset of OTSI, for a diffuse beam with $v_b/v_T \gg (n/n_b)^{1/3}$ is estimated as $W_e/T_e \gtrsim (15 j_T/j_e)^{2/5}$ (Ref. 10); $j_T = n v_{T_e}/4$, where j_e is the "through" current density), and for a narrow beam as $W_e/T_e \gtrsim j_T/j_e$ (Ref. 11). In either case for typical conditions in PSEs in regimes with AN ($T_e \sim 0.5$ eV, $j_e \sim 1$ A/cm², $j_T \sim 5$ A/cm²), the beam energy comes out to be not too large — of the order of 10–15 eV, in complete agreement with the minimum anode voltage overshoot at which AN arises at small E_a. We note that the arguments presented above do not in any way imply a pressure threshold for the onset of AN. An increase in the density of neutral atoms has the direct effect only of decreasing the collisional damping length of the waves and the amplitude of the driven oscillations, i.e., it has an effect in the opposite direction. This suggests that the cause is nontrivial and may be due to effects of a most unforeseen nature, e.g., related to the conditions of absorption and reflection of the waves on the anode Langmuir sheath.

[1)]Of course, in the case of complete blocking of the current the beam always receives all the energy stored in the inductance L_a during the passage of current, but the transfer time decreases in the case of a more rapid drop in current.

[2)]The shift of the initial part of the $\Delta I/I_0(-V_g^-)$ curves, which corresponds to "ordinary" initial blocking, is due initially to growth of the plasma density in the cathode region and of the jump in potential at the grid plane and then to growth of the density in the anode region (see Part II of Ref. 8).

[3)]It should be noted that the main growth in density occurs in the period when the current has already recovered,[3,4] so that the generation of ions occurs mainly on account of the energy stored in the bound electrons.

---

[1] V. B. Kaplan, A. N. Makarov, A. M. Martsinovskiĭ et al., Zh. Tekh. Fiz. 47, 274 (1977) [Sov. Phys. Tech. Phys. 22, 159 (1977)].

[2] V. B. Kaplan, A. M. Martsinovskiĭ, V. F. Mustafaev et al., Zh. Tekh. Fiz. 49, 527 (1979) [Sov. Phys. Tech. Phys. 24, 303 (1979)].

[3] F. G. Baksht, A. A. Kostin, A. M. Martsinovskiĭ et al., Pis'ma Zh. Tekh. Fiz. 5(8), 905 (1979) [Sov. Tech. Phys. Lett. 5, 374 (1979)].

[4] F. G. Baksht, A. A. Bogdanov, V. B. Kaplan et al., Fiz. Plazmy 3, 547 (1991) [Sov. J. Plasma Phys. 3, ***** (1991)].

[5] A. A. Artsimovich and R. Z. Sagdeev, Plasma Physics for Physicists [in Russian], Énergoatomizdat, Moscow (1973).

[6] V. E. Zakharov, Zh. Éksp. Teor. Fiz. 62, 1745 (1972) [Sov. Phys. JETP 35, 908 (1972)].

[7] P. Y. Cheung and A. M. Wong, Phys. Fluids 28, 1538 (1985).

[8] F. G. Baksht, V. B. Kaplan, A. A. Kostin *et al.*, Zh. Tekh. Fiz. **48**, 2279 (1978) [Sov. Phys. Tech. Phys. **23**, ***** (1978)].

[9] F. G. Baksht and A. A. Kostin, *Mathematical Modeling of Processes Occurring in Low-Voltage Beam–Plasma Discharges* [in Russian], Énergoatomizdat, Moscow (1990).

[10] A. A. Galeev, R. Z. Sagdeev, V. A. Shapiro, and V. I. Shevchenko, Zh. Éksp. Teor. Fiz. **72**, 508 (1977) [Sov. Phys. JETP **45**, ***** (1977)].

[11] M. V. Nezlin, *Dynamics of Beams in Plasmas* [in Russian], Énergoatomizdat, Moscow (1982).

Translated by Steve Torstveit

Self-organization of a light-sustained discharge

A. A. Tel'nikhin

Altai State University, 656099 Barnaul, Russia

A theoretical model is proposed to describe a light-sustained discharge sustained by the radiation from a neodymium laser. The model is based on the Navier-Stokes equations. A solution of these equations is found in the form of a quasisimple wave, and it is shown that the evolution of the system exhibits bifurcation, where the point of bifurcation determines the energy deposition into the discharge. A study is made of the transition of the system to the stable state with complicated space-time and functional structures and a high level of organization (the degree of ordering is estimated from the change in the information entropy). Within this model calculations are made of the macroscopic parameters of the discharge, the level of fluctuations and the required power, and they agree with experimental data. © *1997 American Institute of Physics.* [S1063-7842(97)00506-0]

A glow discharge in the Earth's atmosphere sustained by the radiation from a neodymium laser is the object of intense research, and is used for practical purposes in technology.[1,2] A discharge of this type was first obtained by Bunkin *et al.*[3] Subsequent investigations showed that the discharge has a threshold in the radiation intensity w: ($w_c \simeq 10$ MW/cm^2). The power needed for sustaining the discharge is $P_c \simeq 1$ MW and does not depend on the radius of the light beam. As the power is increased, the discharge front moves along the light channel with a velocity of tens of meters per second. The discharge plasma is optically transparent (the absorption coefficient is $\mu \sim 10^{-2}$ cm^{-1}), and its parameters (temperature and density) are on the average constant in time and uniform within the light channel. The average electron (ion) density is $n \simeq 2 \times 10^{17}$ cm^{-3}, the average temperature is $T \sim 1$ eV, and the pressure is equalized in space because of the subsonic mode of propagation. The velocity of the discharge front depends on the intensity of the external source, and above a certain threshold increases severalfold according to the formula $v_0 \sim w^{1/2}$. Effects are seen in the region of the threshold that are associated with fluctuations in the velocity of the front while the profile of the front changes very little. In Refs. 4 and 5 probe, acoustic, and optical methods were used to study the properties of the discharge associated with its nonequilibrium nature. The measurements revealed the nature of the macroscopic fluctuations of the plasma parameters, and it was also observed that the discharge is a source of intense acoustic waves with a frequency of the order of 10 kHz. The first theoretical model for the discharge was proposed by Yu. P. Raizer.[1] In that paper, as in subsequent ones,[2] an analogy was made between the burning of a Bickford fuse and the motion of the discharge. This model, which is described by a one-dimensional nonlinear heat conduction equation, predicts the correct dependence of the velocity of the front on the radiation intensity.

The model constructed in this investigation, which reflects the behavior of the discharge, begins from the equations of hydrodynamics that take into account highly nonhydrodynamic mechanisms for energy transport: heat conduction and radiation. The divergence of the light beam is neglected, the plasma is assumed to be optically transparent, and it is assumed that the discharge channel has cylindrical symmetry with a characteristic transverse dimension (radius) r_0. It is also assumed that the principal mechanism for the loss of energy is the intrinsic emission of the plasma (this is true if the beam radius is $r_0 > 0.1$ cm).[1,2]

PRINCIPAL EQUATIONS OF THE MODEL

In describing the properties of the discharge we shall start from the Navier–Stokes equations for the density field ρ, the velocity field $\mathbf{v}$, the pressure field p, and the temperature field T:

$$\rho_t + \nabla \cdot (\rho \mathbf{v}) = 0,$$

$$\mathbf{v}_t + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\rho^{-1} \nabla p,$$

$$\rho T(s_t + \mathbf{v} \cdot \nabla s) = \varkappa \Delta T + F, \quad F = \mu w - \Phi, \tag{1}$$

where s is the entropy per unit mass of the plasma, $\varkappa$ is the thermal conductivity, Φ is the power of the intrinsic emission of the plasma, and Δ is the Laplacian operator.

Let us consider the non-steady-state solution of Eqs. (1) in the Boussinesq approximation, retaining up to terms of second order in the small parameters

$$\varepsilon = v/c, \quad \varepsilon^{1/2} = r_0/\lambda, \tag{2}$$

where c is the speed of sound in the medium and λ is the characteristic length of a disturbance.

We take the z axis along the direction of motion of the discharge and go over to a system of coordinates moving with the velocity v_0 of the discharge front. The solution of Eqs. (1) with the obvious symmetry of the problem can be written in the form

$$\rho = \rho_0 + \rho'(\mathbf{r}, t), \quad T = T_0 + T'(\mathbf{r}, t), \quad v = \mathbf{e}_z = v(\mathbf{r}, t), \tag{3}$$

where ρ' and T' are the perturbed values of those variables, and the standard equation of state of the system (ρ_0, T_0) is given by the nonlinear equation

$$F(\rho, T) = \mu(\rho, T)w - \Phi(\rho, T) = 0|_{T=T_0, \rho=\rho_0}. \tag{4}$$

Substituting Eqs. (3) into the two last equations (1) and using Eq. (4), we find with the accuracy of the problem

$$\rho(v_t + v v_z) = -p_z, \tag{5}$$

$$\rho_0 T_0 s_t = (\varkappa \Delta + F')T', \tag{6}$$

$$F' = (\partial F / \partial T)|_{T=T_0, \rho=\rho_0}. \tag{7}$$

Equation (6) implies that the change of entropy is a second-order quantity. The equation of continuity is conveniently written in another form, which expresses the density perturbation in terms of the variation of the pressure and entropy

$$\rho' = (\partial \rho / \partial p)_s p' + (\partial \rho / \partial s)_p s.$$

Using this relation and Eq. (6), as well as the relation

$$\left(\frac{\partial p}{\partial \rho}\right)_s \left(\frac{\partial \rho}{\partial s}\right)_p \left(\frac{\partial T}{\partial p}\right)_s = \left(\frac{1}{c_v} - \frac{1}{c_p}\right)T,$$

where c_v and c_p are the specific heats, the equation of continuity can be written in the following way

$$p_t + v p_z + c^2 \rho_0 v_z = \frac{\gamma - 1}{\rho_0 c_p} (\varkappa \Delta + F')p'. \tag{8}$$

Here $c^2 = (\partial p / \partial \rho)_s$ is the adiabatic speed of sound and $\gamma = c_p / c_v$. Equations (5), (6), and (8) constitute a complete system of equations up to and including nonlinear second-order terms, with the linear dissipative terms considered equivalent in order of magnitude to the squares of the nonlinear terms. The solution of these equations will be sought in the form of quasisimple waves,[6] setting

$$p(\mathbf{r}, t) = \pm \rho_0 c + \psi(\mathbf{r}, t). \tag{9}$$

Substituting Eq. (9) into Eqs. (5), (6), and (8), we obtain the starting equation for a quasisimple wave travelling in one direction

$$v_t + \left(c_0 + \frac{\gamma + 1}{2} v\right) v_z = (D\Delta + \nu)v, \tag{10}$$

$$D = \frac{(\gamma - 1)\varkappa}{2\rho_0 c_p}, \quad \nu = \frac{(\gamma - 1)F'}{2\rho_0 c_p}. \tag{11}$$

Here c_0 is the velocity of sound in the unperturbed medium, D is the heat diffusion constant, and ν is the characteristic frequency of energy deposition into the discharge.

STABLILITY OF THE LINEARIZED SYSTEM

If we go over to a coordinate system moving with the velocity c_0 relative to the medium, then Eq. (10) takes on a simpler form

$$v_t = L(\sigma)v + \hbar(v; \sigma), \tag{12}$$

where $t(t \to t - z/c_0)$ is the "slow" time, $L = D\Delta + \nu$ is a linear operator acting in the space of definition of the function $v(\mathbf{r}, t)$ and $\hbar(v; \sigma)$, which allows for the nonlinear nature of Eq. (10) and σ reflects the dependence of the solution on the parameters of the problem (ν, D).

Since the system of equations (12) is autonomous, the linear auxiliary problem

$$v_t = L(\sigma)v \tag{13}$$

admits of a solution of the form

$$v = u(\mathbf{r})\exp(\sigma t). \tag{14}$$

Substituting Eq. (14) into (13) and using the explicit form of the operator L we obtain in the linearized coordinate system

$$D\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2}\right)u + (\nu - \sigma)u = 0. \tag{15}$$

This equation, plus the boundary conditions, defines an eigenvalue problem For example, for the case of axial symmetry with the boundary condition $u(r = r_0) = 0$, we find from Eq. (15) the following eigenvalues and eigenfunctions of the problem:

$$\sigma = \nu - (k_\perp^2 + k^2)D, \quad k_\perp r_0 \approx 2.40;$$

$$u = J_0(k_\perp r)\exp(ikz), \tag{16}$$

where $J_0(k_\perp r)$ are the Bessel functions.

The solution of the linearized problem (13) is thus

$$v(\mathbf{r}, t) = a u(\mathbf{r})\exp(\sigma t), \tag{17}$$

where u and σ are defined by Eq. (16) and a is a constant.

It is easy to see that the nature of the solution depends to a large extent on the eigenvalues of the parameter σ (or equivalently, on the parameter ν). If $\nu \geq k_c^2 D$ an instability arises in the system; the critical point

$$\sigma_c = 0, \quad \nu_c = k_c^2 D, \quad k_c^2 = (k_\perp^2 + k^2)_c \tag{18}$$

corresponds to marginal stability, and as follows form Eqs. (7) and (11), determines the threshold condition for the existence of a discharge. It is important to note that Eqs. (16) and (18) imply that the unstable mode is entirely determined by the system of parameters and is characterized by the spatial length of perturbation of the steady-state solution. It might be expected that for $\sigma > \sigma_c$ this perturbation will determined the main properties of the system.

EVOLUTION OF THE NONLINEAR SYSTEM

We turn now to an analysis of the dynamics of the system described by the nonlinear equation (12). Henceforth we shall limit the discussion to bifurcation solution near the critical point σ_c. Then, taking account of the explicit form of the nonlinearity and the results of analyses of the linear problem, we will seek the solution of the problem as a series in the small parameter

$$v(\mathbf{r}, t) = a(t)(\varepsilon e^{ikz} + \varepsilon^2 e^{2ikz} + \ldots$$

$$+ \text{c.c.}) \sum_{n=1}^{\infty} b_n J_0(k_{\perp n} r), \tag{19}$$

where, according to Eqs. (16) and (17), $a(t)$ is an unknown function and $b_1 = 1$, and $J_0[(k_\perp)r_0] = 1$.

Substituting the solution in the form (19) into Eq. (12), averaging over the high-frequency oscillations, and using the orthogonality of the Bessel functions

$$\int_0^{r_0} J_0(k_n r)J_0(k_p r)r\,dr = \frac{1}{2}\delta_{np}J_1^2(k_p r_0),$$

we find the solution that determines the function $a(t;R)$,

$$a_t + \nu_c(R-1) - \mu a^3 = 0. \tag{20}$$

Here we have used the notation

$$R = \frac{\nu}{\nu_c}, \quad \mu = \frac{(\gamma+1)^2 J_1^2(k_\perp r_0)}{12D}. \tag{21}$$

The evolution equation (20) with the parameters (21) have a well-known solution and describes the supercritical bifurcation of the system.[7] For $R>1$ this equation admits of the following two solutions

$$a_s(R) = \pm\sqrt{(\nu_c/\mu)(R-1)}. \tag{22}$$

These solutions are asymptotically stable and come in in a characteristic time τ, $\tau = 1/2\nu_c(R-1)$. The mathematical reflection of the qualitative change in the behavior of the system caused by bifurcation at the point $R_c = 1$ is the singularity, making the solution nonanalytic near the critical point.

Now, going back to the laboratory coordinate system and taking into account Eqs. (16) and (22), we can write the equation for the traveling wave in the form $(t>\tau, \tau = 1/2\nu_c(R-1))$

$$v_i = a(R)J_0(k_\perp r_0)\exp(-i\omega t + ik(z-z_0)), \tag{23}$$

where $\omega = kc_0$ is the vibrational frequency and $z_0 = v_0 t$ is the coordinate of the front.

Let us turn now to the conditions at the discharge front, allowing for the interaction of the incident wave (p_i', v_i), the reflected wave (p_r', v_r), and the transmitted wave (p_{tr}', v_{tr}). In the boundary condition $p'=0$ at the discharge front we must in the approximation neglect the emitted wave (we shall see that the intensity of the acoustic emission is low). Then we have the condition $p_i' = p_r'$; for the velocity we shall have, correspondingly, $v_r = -v_i$, so that the total velocity behind the front is $v = 2v_i$, and before the front it is $v_{tr} = 2(\rho_0/\rho_\infty)v_i$ (we have taken into account the law of conservation of mass; ρ_∞) is the gas density ahead of the discharge front). From these conditions we find the final expression that describes the space-time structure of the field in the discharge plasma

$$v(\mathbf{r},t) = 2a_s(R)J_0(k_\perp r)\cos k(z-z_0)\exp(-i\omega t). \tag{24}$$

The measured velocity of the discharge front is given by the expression $v_0 = \sqrt{1/2\langle\langle|v|^2\rangle\rangle}$, where the angle brackets denote averaging over the time and over the cross section of the plasma channel. Now substituting Eq. (24) and using the conditions (18), (12), and (22), we obtain an expression for the velocity of the front

$$v_0 = \sqrt{2}a_s(R)J_1(k_\perp r_0) = \frac{2\sqrt{6}}{\gamma+1}k_\perp D\sqrt{R-1}. \tag{25}$$

Let us turn now to another aspect of the problem — the intensity of the acoustic emission. The energy flux in the incident wave is $I_i = (1/2)\rho_0 c_0 |v_i|^2$ and the total intensity of the emitted sound is $I_{tr} = \pi r_0^2\rho_\infty|\dot{v}_{tr}|^2/8\pi c_\infty$, where c_∞ is the

velocity of sound in air (ahead of the discharge front).[8] Using Eqs. (24) and (25) and the relations $\rho_0/\rho_\infty = T_\infty/T_0$ and $c_0/c_\infty = (T_0/T_\infty)^{1/2}$, we obtain the ratio of the emitted energy to the flux in the incident wave

$$\frac{I_{tr}}{I_i} = \left(\frac{r_0\omega}{c_0}\right)^2\cdot\left(\frac{T_\infty}{T_0}\right)^{1/2}, \quad I_i = \frac{1}{4}\rho_0 c_0 v_0^2. \tag{26}$$

Since $r_0 \leq c_0/\omega$ and $T_\infty \leq T_0$, we obtain $I_{tr}/I_i \leq I$. Let us study the application of the mathematical structure of the model in a specific physical situation. We begin from the definition (18) for the calculation of the threshold power of the laser radiation, $P_c = \pi r_0^2 w_c$. Using conditions (2) and and expressions (4), (11), (16), we obtain from (18) $P_c = (2.4)^2\pi\varkappa T_0/\mu B$. Here we note that the power required for sustaining the discharge does not depend on the radius of the light beam, and for typical values ($T_0 \simeq 1.6\times 10^4$ K, $\varkappa = 2.5\times 10^{-2}$ W/cm·K, $r_0 \simeq 0.2$ cm, $\mu \simeq 10^{-2}$ cm^{-1}, and $B \simeq 0.5$ (Refs. 1 and 2)), is equal to $\simeq 1$ MW. Similarly it may be shown that the velocity profile described by the function $J_0(k_\perp r)$ and the discharge velocity itself, Eq. (25), near the threshold ($P \gtrsim P_c$) are of the order of 10 m/s, the level of the fluctuations in the plasma is $\rho'/\rho_0 \sim v_0/c_0 \sim 10^{-2}$, and the intensity (26) of the sound wave is $I_{tr} \sim 1-10$ W/m^2. We also note the dependence of these quantities on the degree of supercriticality $R = (P/P_c)$

$$v_0 \sim \sqrt{R-1}, \quad \rho'/\rho_0 \sim \sqrt{R-1}, \quad I_{tr} \sim (R-1).$$

It is easy to see that these results agree with experimental data.

Self-organization effects

Let us go now to Eq. (20), which describes the evolution of the system. This equation has a singularity at the critical point $R_c = 1$, associated with a qualitative change and transition of the system from one state to another. To describe correctly the dynamics at the bifurcation point we must take into account the fluctuation forces. Let us assume that the system has a Langenvin source with a power $Q\delta(t-t') = \langle F(t)F(t')\rangle$, where $\delta(t-t')$ is the Dirac delta function. Then the evolution equation assumes the form

$$a_t = \nu_c(R-1)a - \mu a^3 + F(t), \tag{27}$$

where $F(t)$ is a random force.

We note that Eq. (27) has the form of a paradigmatic equation describing a nonequilibrium phase transition.[9] We introduce the distribution function $f(a,t)$ in the space of the "coordinate" $a(R)$. The Fokker-Planck equation describing the change with time of the function $f(a,t)$, and correspondingly, Eq. (27), have the form

$$\frac{\partial f}{\partial a} = Q\frac{\partial^2 f}{\partial a^2} + \frac{\partial}{\partial a}[(\nu_c(R-1)a^2 - \mu a^4)f];$$

$$\int f(a)\,da = 1. \tag{28}$$

The steady-state solution of this equation is

$$f(a) = N\exp(\alpha a^2 - \beta a^4), \tag{29}$$

$$\alpha = \nu_c(R-1)/2Q, \quad \beta = \mu/4Q. \tag{30}$$

There are two most probable values for the distribution (29), called the order parameter,

$$a_s = \pm (\alpha/2\beta)^{1/2}, \tag{31}$$

which merge for $R=1$ and coincide with expression (22). By definition the quantity

$$I = - \int da f(a) \ln f(a), \tag{32}$$

is called the information entropy, and the effectiveness of the system is defined by the formula

$$W = \frac{d\langle a^2 \rangle}{d\alpha}, \tag{33}$$

where

$$\langle a^n \rangle = \int_{-\infty}^{\infty} f(a) a^n da \tag{34}$$

is the nth moment of the distribution function $f(a)$.

Using Eq. (29) and calculating I and W from the formulas (32) and (33), we obtain

$$I = -\ln N - \alpha \langle a^2 \rangle + \beta \langle a^4 \rangle,$$
$$W = \langle a^4 \rangle_\alpha - \langle a^2 \rangle_\alpha^2. \tag{35}$$

A simple application of the standard integral

$$\int_0^{\infty} x^n \exp(\alpha x^2 - \beta x^4) dx$$
$$= \frac{1}{2}(2\beta)^{-(n+1)/4} \Gamma\left(\frac{n+1}{2}\right) D_{-(n+1)/2}(\chi) \exp(\chi^2/4),$$

where $\chi = \alpha/\sqrt{2\beta}$, Γ is the gamma function and D is the parabolic cylinder function, to the integral $\int f(a) da = 1$ gives the following normalization constant

$$N = 2 (2\beta)^{1/4} \pi^{-1/2} \exp(-\chi^2/4)[D_{-1/2}(\chi)]^{-1}. \tag{36}$$

In a similar way, formula (34) allows us readily to calculate the second and fourth moments

$$\langle a^2 \rangle = (2\beta)^{-1/2} D_{-3/2}(\chi)/D_{-1/2}(\chi),$$
$$\langle a^4 \rangle = 3 (4\beta)^{-1} D_{-5/2}/D_{-1/2}. \tag{37}$$

We shall study the properties of the system at the bifurcation point $(R=1)$ and in the region of stability $(R \gg 1)$ using the relation between the parabolic cylinder function and the Weber function, $D_{-p-1/2}(\chi) = U(p, \chi)$. At the point of instability $(R_c = 1, \chi_c = 0)$, we have used the properties of the Weber function

$$U(p,0) = \frac{\sqrt{\pi}}{2^{(2p+1)/4}}\left[\Gamma\left(\frac{p}{2} + \frac{3}{4}\right)\right]^{-1},$$

to find from Eqs. (36) and (37)

$$N_c = 1.1032 \beta^{1/4}, \quad \langle a^4 \rangle_c = 0.510 \beta^{-1}, \tag{38}$$

$$\langle a^2 \rangle_c = 0.6760 \beta^{-1/2}. \tag{39}$$

In the other limit $(\chi \gg \chi_c)$ the corresponding expansion for the Weber function has the form of $U(p,\chi) \simeq \exp(-\chi^2/4) \chi^{p+1/2}$. Using this approximate representation in Eqs. (36) and (37), we obtain

$$N = (2/\sqrt{\pi})(2\beta/\alpha)^{1/2}, \quad \langle a^4 \rangle = (3/2)(\alpha/2\beta)^2, \tag{40}$$

$$\langle a^2 \rangle = \alpha/2\beta. \tag{41}$$

Equation (41) is the same as the square of the order parameter, Eq. (31). Substituting Eqs (40) and (41) in the second of Eqs. (35), and using the definition (30), we find that the efficiency of the system increases with the threshold as $W \propto (R-1)^2$. The information entropy was calculated using the recipe of the S-theorem,[10] i.e., the relative degree of ordering, will be estimated from the value of the entropy for a given value of the mean effectiveness of the energy state of the open system. In other words, we require that the following inequality hold

$$\langle a^2 \rangle_{c,r} = \langle a^2 \rangle_r. \tag{42}$$

Applying Eqs. (39) and (41) to the conditions (16), we find the relation $(\alpha/\sqrt{2}\beta = \sqrt{2} \cdot 0.676$. Using the value found for the quantity $(\alpha/\sqrt{2}\beta)_r$ in formulas (38)–(41) it is not difficult to infer from the definition (35) that the increase in the entropy is $\Delta I = (I - I_c) \simeq -2.33$. This implies that the information entropy decreases as the system makes the transition into a more ordered (less symmetric) state, i.e., self-organization occurs in the system.

In conclusion, let us assess the behavior of the system at the critical point, assuming that fluctuations $\delta w(t)$ in the laser radiation serve as the source of random forces. The term $\mu w(\delta w/w)$, of the next order of smallness, appears in the third of the basic equations (1). By making a transformation similar to that used in the derivation of Eq. (10), we obtain

$$Q = (c_0/4J_1(k_\perp r_0))^2 (\mu w/\rho_0 c_p T_0)^2 K^2 t_c, \tag{43}$$

where t_c is the characteristic correlation time of the fluctuations, and $K = \sqrt{\langle \delta w^2 \rangle / w^2}$ is their relative level.

Substituting Eqs. (21) and (43) into (39), we estimate the fluctuations in the velocity of the front at the critical point $R=1$

$$\langle a^2 \rangle_c \simeq 0.676\sqrt{12}(c_0/(\gamma+1)J_1)(\mu w/\rho_0 c_p T_0) \sqrt{K^2 t_c D}. \tag{44}$$

For a given noise level (43) we determine from the condition $\langle a^2 \rangle / \langle a^2 \rangle_c \gg 1$ the region of stability of the system

$$R - 1 \gg 0.676(\gamma+1)(c_0/k_\perp D)(\mu w/\rho_0 c_p T_0) \sqrt{K^2 t_c/12D}. \tag{45}$$

For typical parameter values of the discharge and $K \sim 10^{-5}$, $t_c \sim 10^{-5}$ s, formulas (44) and (45) give the following estimates: $\sqrt{\langle a^2 \rangle_c} \sim 1$ m/s, $R - 1 \gg 10^{-3}$.

CONCLUSIONS

The formation of a light-sustained discharge has been studied within the framework of the hydrodynamic equations and the Boussinesq approximation. It has been shown that a hydrodynamic instability is probable in the system, with the

Rayleigh number played by the ratio $R = \nu(w)/k_c^2 D$, where $\nu(\omega)$ is the characteristic frequency of energy deposition into the system, which depends parametrically on the external field; D is the thermal diffusion coefficient, and k_c is the wave number, which is determined by the internal parameters of the system. The bifurcation point $R_c = 1$ of the system reflects the qualitative changes in the evolution and dictates the threshold nature of the development of the discharge. For $R > 1$ the irreversible processes in the nonequilibrium discharge plasma initiate self-organization of the discharge and the formation of coherent space-time structures with axial symmetry, described by Bessel functions.

Demonstrations of these internal processes are found in macroscopic effects: the directional motion of the discharge and the magnitude and profile of the front velocity. In addition, the discharge takes on a functional structure — induced noise. It also should be noted that the presence of Langenvin sources in the dynamics shows up in the dynamics of the system near the threshold and in the equal probability of the appearance of leading and trailing discharge fronts. A comparison of the results obtained in the framework of this model with experimentally determined data, brings us to the conclusion that the phenomenon of self-organization plays an important role in a discharge of this type.

[1] Yu.P. Raĭzer, *Fundamentals of the Modern Physics of Gas-Discharge Processes* [in Russian] (Nauka, Moscow, 1980).
[2] I. A. Bufetov, A. M. Prokhorov, V. B. Fedorov, and V. K. Fomin, Tr. IOFAN **10**, 3 (1988).
[3] F. V. Bunkin, V. I. Konov, A. M. Prokhorov *et al.*, JETP Lett. **9**, 371 (1969).
[4] V. I. Bukatyĭ, A. A. Kobolov, and A. A. Tel'nikhin, Zh. Tekh. Fiz **55**, 312 (1985) [Sov. Phys. Tech. Phys. **30**, 182 (1985)].
[5] V. I. Bukatyĭ, K. I. Deĭnes, and A. A. Tel'nikhin, Zh. Obshch. Akust. **4**, 753 (1991).
[6] V. I. Karpman, *Nonlinear Waves in Disperse Media* [in Russian] (Novosibirsk, 1968).
[7] I. Prigogine, *From Being to Becoming: Time and Complexity in the Physical Sciences* (Freeman, San Francisco, 1980) [Russ. transl., Nauka, Moscow, 1985)].
[8] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, Oxford, 1987) [Gidrodinamika, Nauka, Moscow (1986)].
[9] H. Haken, *Synergetics: an Introduction* (Springer-Verlag, Berlin, 1977) [Mir, Moscow, 1980].
[10] Yu. L. Klimontovich, *Turbulent Motion and the Structure of Chaos. A New Approach to the Statistical Theory of Open Systems* [in Russian] (Nauka, Moscow, 1990).

Translated by J. R. Anderson

Non-self-sustained hollow-cathode glow discharge for large-aperture ion sources

A. V. Vizir', E. M. Oks, P. M. Shchanin, G. Yu. Yushkov

Institute of High-Current Electronics, Siberian Branch of the Russian Academy of Sciences, 634055 Tomsk, Russia

A discharge system is proposed in which an auxiliary gas discharge is used to inject electrons into the cathode cavity of a hollow-cathode glow discharge. A study is made of the region of stable existence of a non-self-sustaining hollow-cathode discharge. It is shown that the injection of electrons permits a reduction to $<10^{-2}$ Pa in the minimum pressure at which a discharge can exist. It is shown experimentally that this discharge can be used to generate wide-aperture ion beams. © *1997 American Institute of Physics.* [S1063-7842(97)00606-5]

INTRODUCTION

Stable parameters, a high current density, and simplicity of design make a glow discharge attractive for application in plasma sources of charged particles.[1,2] It is known[3] that for a glow discharge to operate in the high-current low-voltage mode it is necessary to maintain a pressure of about 1 Pa in the discharge gap. However, to maintain the dielectric strength of the accelerating gap and to ensure transport of the accelerated beam over large distances the pressure in these regions must be lower than 10^{-2} Pa.[4] In sources of narrowly focused[5] or tubular[6] beams based on a glow discharge, this problem has been solved by making a pressure drop in the narrow emission channel that separates the region of charged particle acceleration from the region of plasma generation.

Because of the large area of the emission surface, the residual gas pressure in the discharge region and the accelerating region will be nearly the same in plasma-emitter systems for making wide-aperture ion beams.[7] A reduction of the pressure in the discharge gap results in a fast increase in the discharge voltage, and the discharge is converted to a high-voltage low-current form or is interrupted if the voltage supply is inadequate.

In order to create conditions for electron oscillation, the pressure must be reduced, as it is in Penning discharges[5] or magnetron discharges,[8] and also in a hollow-cathode discharge.[2] However, in the first case the presence of a magnetic field in the discharge gap degrades the uniformity of the plasma parameters, and in the second case, even though the minimum pressure at which a low-voltage discharge can burn is somewhat lower than in an ordinary glow discharge, it is still above 10^{-1} Pa.[9]

In addition, it is clear that the low-voltage form of a glow discharge, when the ratio of the electron component of the current from the cathode to the ion component is small, external injection of a small fraction of the electrons would be able to provide a stable glow discharge at lower pressures. For this purpose a thermionic cathode could be used, but the high temperature of the emitter and its rapid poisoning in the active gaseous media limit the scope of its application.

In this paper we present the results of investigations on the conditions of a stable hollow-cathode glow discharge in the range of low pressures, in which a plasma electron emitter with a cold cathode is used for auxiliary injection of particles.

EXPERIMENTAL APPARATUS

The discharge system used in these experiments, which is shown in Fig. 1, consists of two main units: a discharge gap for the main discharge and a plasma electron emitter.

The main discharge is ignited between a cylindrical hollow cathode *1* (length 40 cm and diameter 22 cm) and an anode *2* located within, in the form of a metal rod with $d=$ 8 mm. The anode could be drawn vertically into the insulator *3* so as to change the operating length of the anode and consequently vary the area S_a of its working surface between 3 and 45 cm^2. The plasma ions from the main discharge were sampled through a rectangular emission window *4*, 2.5 cm wide and 30 cm long, which was covered with a fine stainless steel grid with a mesh size of 1 mm, placed on the side surface of the hollow cathode *1* to stabilize the discharge.

Supplementary electrons injected into the cathode cavity of the main discharge were withdrawn from the plasma of the auxiliary discharge which was ignited between the hollow cathode *5* (length and diameter 12.5 cm) and the electrode *1*, which acted as the anode of the auxiliary discharge. The auxiliary discharge was initiated by a discharge over the surface of the ceramic ring *6*, which started at the instant that a voltage was applied between the electrodes of the auxiliary discharge. An aperture 5 mm in diameter was provided in the center of the end of the hollow cathode and was situated coaxially with an aperture 7 mm in diameter in electrode *1*. This size of aperture in the electrodes *1* and *5* helped maintain the pressure drop necessary for stable ignition and burning of the auxiliary discharge, and also provided effective electron emission from the plasma of the auxiliary discharge and the passage of the electron flux into the cathode cavity of the main discharge with minimum losses (the current of the injected electrons was 0.8–0.9 times the current of the auxiliary discharge). The electrons emitted by the plasma of the auxiliary discharge were accelerated in the cathode fall of the main discharge and consequently had an energy equal to the energy of the electrons emitted from the walls of the hollow cathode of the main discharge as a result of γ processes. To reduce the penetration of the potential into the cathode cavity

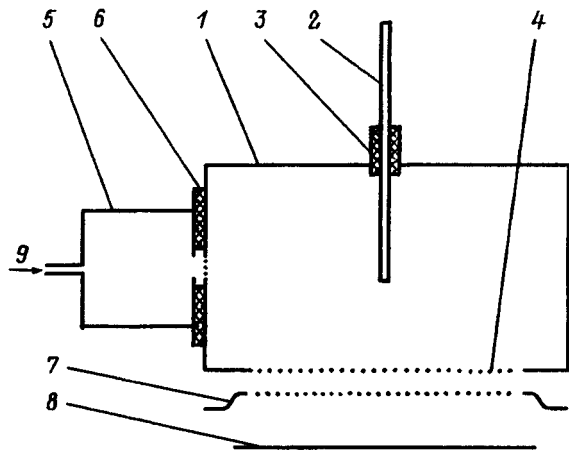FIG. 1. Diagram of the discharge system: *1* — Hollow cathode of the main discharge; *2* — anode of the main discharge; *3, 6* — ceramic insulators; *4* — emission window; *5* — hollow-cathode auxiliary discharge; *7* — accelerating electrode; *8* — collector; *9* — gas inlet.

FIG. 2. Discharge ignition voltage as a function of the pressure: *1* — without auxiliary discharge, *2* — auxiliary discharge current 50 mA.

of the auxiliary discharge, the aperture in the end electrode was covered with a tungsten grid with a mesh size of 0.2 mm. It was possible to inject electrons simultaneously from both ends of the cathode cavity of the main discharge by providing two identical auxiliary discharge cells.

The electrical supply of the main discharge came from a 2 kV voltage rectifier, isolated by an amount equal to the accelerating voltage, which provided a discharge current up to 1 A. The current of the main discharge was measured in the cathode circuit. The auxiliary discharge with a current up to 0.3 A was supplied from a 3 kV voltage rectifier. The voltage to accelerate the ions, up to 5 kV, was applied between the cathode *1* and the accelerating electrode *7*, which was held at a negative potential relative to the collector *8*. The negative potential is required to keep the γ electrons formed by the accelerated ion beam from entering the accelerating gap from the collector. The ion current density distribution from the beam was measured by a Faraday cup placed above the collector.

The pressure in the discharge system was regulated by varying the flow of the working gas admitted into the hollow cathode of the auxiliary discharge. The gas pressure in the cathode region of the main discharge and in the accelerating gap varied between 5×10^{-1} and 5×10^{-3} Pa. Because of the pressure drop incurred in the passage of the gas through the apertures in electrodes *1* and *5* the pressure in the cathode region of the auxiliary discharge was higher than 1 Pa. The working gases were N_2, He, Ar, and O_2.

EXPERIMENTAL RESULTS

When the electrons emitted by the plasma of the auxiliary discharge were not injected into the cathode *1*, the pressure dependence of the ignition voltage of the main discharge (curve *1* in Fig. 2) follows the left-hand branch of the Paschen curve.[10] The sharp falloff of the ignition voltage is obtained at a gas pressure of 2 Pa, which corresponds to an interelectrode distance of 20 cm, as determined from the Pashen curve, which was close to the experimentally deter-
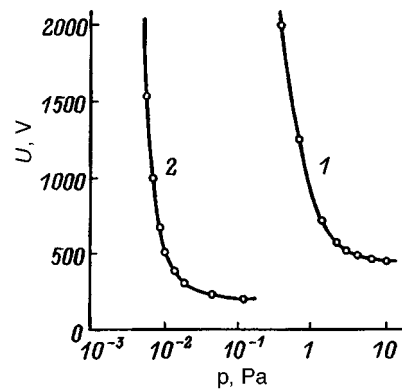
mined value. As one can see in Fig. 2, the ignition voltage increases rapidly when the pressure is reduced, and the discharge cannot start at pressures <0.1 Pa without electron injection. If electrons are injected, the ignition curve of the main discharge is displaced to lower pressures (curve *2* in Fig. 2). For example, at an electron current of 50 mA the main discharge is ignited stably at pressures to 5×10^{-3} Pa. The auxiliary discharge is ignited stably at a working gas flow $>3 \times 10^{-3}$ Pa·m³/s. A voltage of 1–2 kV is sufficient to ignite the auxiliary discharge over the entire pressure range.

A change of the anode surface area S_a has a large influence on the voltage U_d of the main discharge (Fig. 3). It can be seen from this figure that there exists an optimum anode area for which the discharge voltage is minimum. All subsequent measurements were made with the optimum ratio between the anode and cathode areas, S_a/S_c, equal to 4.4×10^{-3}.

The voltage of the main discharge, which was measured in the range from 200–800 V, depends on the pressure, the discharge current I, and the discharge non-self-sustainment coefficient, (Fig. 4), defined as the ratio I_a/I of the auxiliary discharge current to the main discharge current. The pressure at which the main discharge burns, as a function of the discharge non-self-sustainment coefficient, is shown in Fig. 5. This figure shows that an auxiliary discharge with a current no higher than 1/3 of the main discharge current makes it
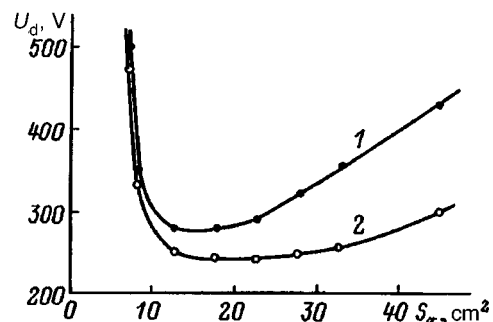

FIG. 3. Discharge voltage as a function of the anode area: $I = 0.5$ A; $I_a = 0.1$ A; p, Pa: *1* — 2.5×10^{-2}, *2* — 3×10^{-2}.
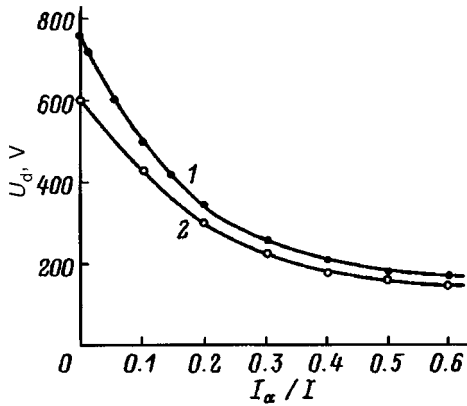
FIG. 4. Discharge voltage as a function of the discharge non-self-sustainment coefficient: $p=2\times10^{-2}$ Pa; I, A: 1 — 0.5, 2 — 0.2.

FIG. 6. Distribution of ion beam current density over the length of the emission window. Injection of electrons from one ($\bullet$) and two ($\circ$) auxiliary discharge cells; $p=2.2\times10^{-2}$ Pa; $I=1$ A; the curent (the sum of the currents) of the auxiliary discharge is 0.2 A.

possible to reduce the pressure by a factor of 5–10 and also reduce the main discharge voltage U_d to 200 V.

The current density distribution of the ion beam extracted from the plasma of the main discharge was very uniform over the length of the emission window (Fig. 6) and was essentially the same whether the electrons were injected from only one end of the cathode cavity or both.

DISCUSSION OF EXPERIMENTAL RESULTS

The effect of electron injection on the displacement of the ignition curve towards lower pressures is in our view due both to the entry into the cathode cavity of fast free electrons that can ionize the gas and to their oscillation in the cathode region. The effect of the oscillation is greatly enhanced by the creation by the injected electrons of a primary plasma, whose existence localizes the electrode voltage drop across the narrow cathode sheath. Since the absorption of fast non-colliding electrons by the cathode surface is improbable,[11] and is impossible for electrons that have lost some of their energy in collisions, the electrons that have been accelerated in the cathode sheath will oscillate in the cathode cavity, reflecting from the cathode sheath until they depart at the anode. The mean range of the electrons in the cathode cavity, which is determined by the probability of their trajectory
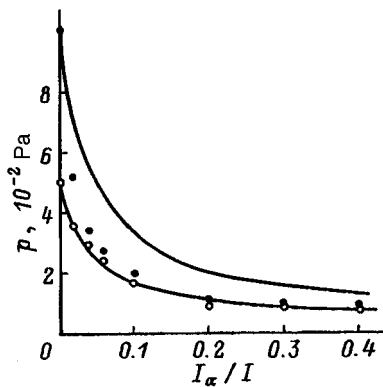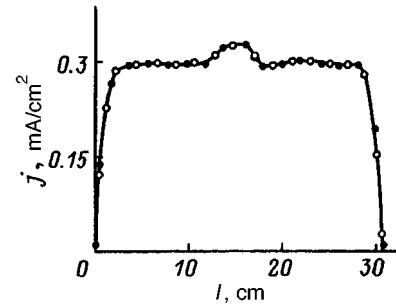

FIG. 5. Operating pressure as a function of the discharge non-self-sustainment coefficient : $U_d=420$ V; I, A: $\bullet$ — 0.5, $\circ$ — 0.25; curves — theory for $\gamma = 0.05$.

intersecting the anode, can be found, according to Ref. 9, by the relation $L=4V/S_a$, where V is the volume of the cathode cavity. Under experimental conditions L was 4×10^3 cm. An increase in the electron mean free path is equivalent to an increase in the electrode gap and causes the discharge ignition curve to be displaced towards lower pressures. The pressure at which the discharge can be ignited, which is determined by the minimum of the Paschen curve ($pd\approx40$ Pa $\cdot$cm), should be 10^{-2} Pa, which is close to the experimentally observed values (curve 2 in Fig. 2).

Since the electrons injected into the cathode cavity are accelerated in the same cathode fall as the γ electrons, the characteristics and properties of a discharge with injected electrons should not be very different from the corresponding characteristics of a hollow-cathode self-sustained discharge. As was the case in Ref. 9, the discharge voltage increases both with increased anode area S_a, because of the departure of the high-energy electrons at the anode (the right-hand branch of the curves in Fig. 3), and with decreased area S_a less than $S_c(2m_e/M_i)^{1/2}$ because of the formation of an anode sheath (the left-hand branch in the curves in Fig. 3).

When the pressure is so low that the primary electrons strike the anode before losing all their energy, the condition for sustaining the discharge without injected electrons has the form[10]

$$n\sigma_i L=1/\gamma, \qquad (1)$$

where n is the gas concentration in the cathode cavity, σ_i is the cross section for ionization of the gas by electron impact, averaged over the electron energy, L is the mean range of the electrons accelerated in the cathode sheath, and γ is the ion-electron secondary emission coefficient.

The condition (1) implies that at a low pressure the secondary emission coefficient γ will increase in proportion to the decrease in pressure. For a self-sustaining discharge the increase in γ can occur only by way of kinetic ion-electron emission. For this to happen requires a sharp increase in the cathode fall and consequently a sharp increase in the discharge voltage. Since the electron component is not more than 10% of the cathode current in a glow discharge, supplementary injection of electrons into the cathode cavity with a current that comprises, for example, the same 10% of the discharge current, is equivalent to doubling the coefficient

γ without any increase in the discharge voltage, and a further increase in the number of electrons injected into the cathode cavity can even reduce the discharge voltage (Fig. 4). With electron injection, the condition for sustaining the discharge becomes

$$n\sigma_i L = (\gamma + (I_e/I_{ic}))^{-1}, \qquad (2)$$

where I_e/I_{ic} is the ratio of the electron current emitted by the auxiliary discharge to the ion current to the cathode of the main discharge.

Since $I_e \approx I_a$ under experimental conditions and $I_{ic} = I/(1+\gamma) \approx I$, and since $\gamma \ll 1$, the dependence of the pressure at which a discharge can exist on the discharge non-self-sustainment coefficient I_a/I can be written as

$$p = \frac{kT}{\sigma_i L}(\gamma + I_a/I)^{-1}, \qquad (3)$$

where k is the Boltzmann constant and T is the temperature of the gas in the cathode cavity.

For a self-sustaining discharge ($I_a = 0$) expression (3) becomes

$$p_0 = kT/\sigma_i L \gamma. \qquad (4)$$

The value of p_0 can be determined from Fig. 5 as the point where the curves intersect the vertical axis. Despite the approximations made in the derivation of expression (3), we find satisfactory agreement between the dependence of the operating pressure of a non-self-sustained discharge on the ratio I_a/I (the line) and the experimental data (the points). The increase in p_0 with increased discharge current can be accounted for by an increase in the gas temperature.

The high degree of uniformity of the ion emission current density (Fig. 6) is caused by the properties of a low-pressure hollow-cathode discharge. The high degree of uniformity of the plasma parameters is due to the fact that the primary ionizing electrons travel a distance many times longer than the average dimensions of the cathode cavity because of multiple reflections from the cathode sheath, and therefore the probability of an ionization event is about the same at any point in the plasma-filled cathode cavity. The high degree of uniformity of the plasma is responsible for the uniform distribution of the ion emission current density. The distribution is not distorted even if the electrons are injected from only one end of the cavity.

CONCLUSIONS

In summary, the injection into the cathode cavity of electrons that are accelerated in the cathode fall permits a reduction to 5×10^{-3} Pa in the pressure at which stable ignition and stable burning are possible in a low-voltage hollow-cathode glow discharge, while the conditions for multiple oscillations of the high-energy electrons lead to efficient ionization of the gas at low pressure and a high degree of uniformity of the plasma discharge. Therefore a non-self-sustained hollow-cathode glow discharge can be regarded as a promising means for generating wide-aperture ion beams.

[1] B. Wolf, *Handbook of Ion Sources* (CRC Press, New York, 1995).
[2] A. S. Metel', *First All-Union Conference on Plasma Emission Electronics* [in Russian] Ulan-Ude (1991), pp. 77–81.
[3] B. I. Moskalev, *Hollow-Cathode Discharges* [in Russian] (Énergiya, Moscow, 1969).
[4] I. G. Brown *The Physics and Technology of Ion Sources* (Wiley and Sons, New York, 1989).
[5] M. A. Zav'yalov, Yu. E. Kreĭndel, A. A. Novikov *et al.*, *Plasma Processes in Technological Electron Beams* [in Russian] (Énergoatomizdat, Moscow, 1989).
[6] *Charged Particle Sources with a Plasma Emitter* [in Russian] edited by P. M. Shchanin (Nauka, Ekaterinberg, 1993).
[7] A. T. Forrester, *Large Ion Beams* (Wiley and Sons, New York, 1988) [Russ. transl., Mir, Moscow, 1991].
[8] E. M. Oks and A. A. Chagin, Zh. Tekh. Fiz **58**, 1191 (1988) [Sov. Phys. Tech. Phys. **33**, 697 (1988)].
[9] A. S. Metel', Zh. Tekh. Fiz **54**, 241 (1984) [Sov. Phys. Tech. Phys. **29**, 141 (1984)].
[10] Yu. P. Raĭzer, *The Physics of a Gas Discharge* [in Russian] (Nauka, Moscow, 1992).
[11] V. N. Glazunov and A. S. Metel', Fiz. Plazmy **8**, 1099 (1982) [Sov. J. Plasma Phys. **8**, 625 (1982)].

Translated by J. R. Anderson

Anisotropy of quasistatic magnetization-reversal processes in (210)-oriented iron-garnet films

E. N. Il'icheva, A. V. Klushina, N. B. Shirokova, N. N. Ustanov, and A. G. Shishkov[1]

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

An experimental study is made of the effect of an in-plane field H_p of various orientations on the domain structure and shape of the hysteresis loops of epitaxial iron garnet films with the (210) orientation. The characteristic of the magnetization reversal process (in fields somewhat lower than the anisotropy field) is taken to be the critical fields H_{p1}, H_{p2}, and H_{p3}, for which the magnetization reversal processes is interrupted at distinct stages. A method is proposed for constructing the phase diagram of the magnetic states of films, using measurements of the critical fields H_p for different amplitudes of the magnetization reversal field H_z. Two directions in the plane of the film are determined with an accuracy of a fraction of a degree from the hysteresis loops, where in the corresponding field H_p the transition from a single-domain state to a multidomain state occurs as a second-order phase transition. The characteristic changes in the shape of the hysteresis loop are consistent with the features in the reorganization of the domain structure of the (210) film. The preferential orientations of the stripe domain structure of the samples are determined relative to the crystal axes as determined by x-ray methods. © *1997 American Institute of Physics.* [S1063-7842(97)00706-X]

INTRODUCTION

The induced anisotropy of (210)-oriented epitaxial iron garnet films contains a large orthorhombic component,[1] so that these films can also be used as the working media for magnetic-bubble memory devices with a faster response speed.[2,3]

The complex anisotropy permits the coexistence of domain structures as magnetic bubbles and stripe domains in (210) films. Control can be exercised by a bias field H_z and an in-plane field H_p independently. These possibilities are important for developing memory based on vertical Bloch lines and for other devices.

The effect that a field H_x along the direction of orientation of the domain walls in planar films exerts on the properties of the domain walls and on the quasistatic magnetization has been studied in Refs. 4 and 5.

In this paper we study the effect of the orientation of the field H_p on the domain walls and the magnetization reversal of (210) films by a low frequency field H_z.

EXPERIMENTAL SAMPLES

In the work reported here we studied iron-garnet thin-film samples of composition $(BiLu)_3(FeGa)_5O_{12}$ deposited by liquid- phase epitaxy on nonmagnetic $Cd_3Ga_5O_{12}$ substrates with the (210) orientation.

In zero magnetic field the films had a stripe domain structure through the film, oriented along the X axis. From x-ray analysis data this is the $[\bar{1}20]$ crystallographic axis. The direction in the plane of the film perpendicular to the domain walls (the y axis) is an axis of the type [001], while the normal **n** to the plane of the film is essentially along the [210] direction.

The easy axis deviates from the normal **n** by an angle $\theta_0 \approx 30°$ and lies in a crystallographic plane of type (001) (the zx plane). The principal parameters of some of the films studied were determined in previous investigations[4,5] and are listed in Table I.

EXPERIMENTAL METHOD

To measure the parameters of the quasistatic magnetization reversal and study the domain structure of the (210) films we used the magnetooptial method described in Ref. 4. The magnetic film was placed in a special holder, which could rotate the film by an arbitrary angle about the normal **n** to the sample surface. The magnetic field $H_p(\phi)$ acting in the plane of the sample was produced by an electromagnet and reached 3 kOe. The bias field H_z, parallel to the normal **n**, did not exceed 100 Oe.

The magnetization reversal signal $m_z(H_p, H_z)$ was measured in dc current with amplitude modulation of the illuminating beam from an LG-79 laser and a U2-8 tuned amplifier. Hysteresis loops were observed in the oscilloscope screen during the magnetization reversal of the sample by means of an ac field H_z with an amplitude in excess of the saturation field of the sample, H_{zs}.

The measurement method was the following. The first direction of the field H_p (i. e., $\phi = 0$) was along the x axis. When the field H_p was increased the hysteresis loops were observed to change on the oscilloscope screen. Readings were made of the critical fields H_p at which the characteristic shape changes occurred in the hysteresis loops. The measurements were made for all the orientations of the field H_p; that is, near the x axis the angle ϕ was varied from 0 to 2π in $\approx 5°$ steps. Near the y axis the steps were changed to 0.5°. For each angle ϕ the changes in the domain structure were observed during the magnetization reversal and the visual observations were compared with the measured amplitudes of the magnetization reversal signal and with the changes in

Sample	h, μm	P_0, μm	θ_F, deg/cm	$\Delta a/a$	θ_0, deg	χ_{zz}	χ_{zx}	M_s, G
1	18.7	26.2	10400	$2 \cdot 10^{-3}$	30	0.116	0.066	5.4
2	13.0	29.5	11900	$0.63 \cdot 10^{-3}$	30	0.161	0.090	4.8

[a]*Note*. h — film thickness, P_0 — equilibrium period, θ_F — Faraday rotation of the plane of polarization of light, $\Delta a/a$ — film-substrate lattice misfit, θ_0 — angle of inclination of hte easy axis from the sample normal, χ_{zz} and χ_{zx} — susceptibility of the domain wall motion, M_s — spontaneous magnetization.

the shape of the hysteresis loop. For $\phi = \pi/2$ and $\pi/2 \pm 10°$ the dependence of the z component of the vector $\mathbf{M}_s$ was measured as a function of the field H_p.

HYSTERESIS LOOPS AND PHASE DIAGRAMS OF (210) FILMS

The hysteresis loops $m_z(H_p, H_z)$ for one of the samples, measured in a field with an amplitude $H_{z0} = 44.5$ Oe ($H_{zs} = 24$ Oe) with variation of the field H_p for $\phi = 0$, are shown in Fig. 1. The field H_p causes a shift of the hysteresis loop along the field H_z. In fields $H_p \gtrsim H_{p1}$ (Fig. 1b) a transition occurs in the parts of the magnetization reversal cycles
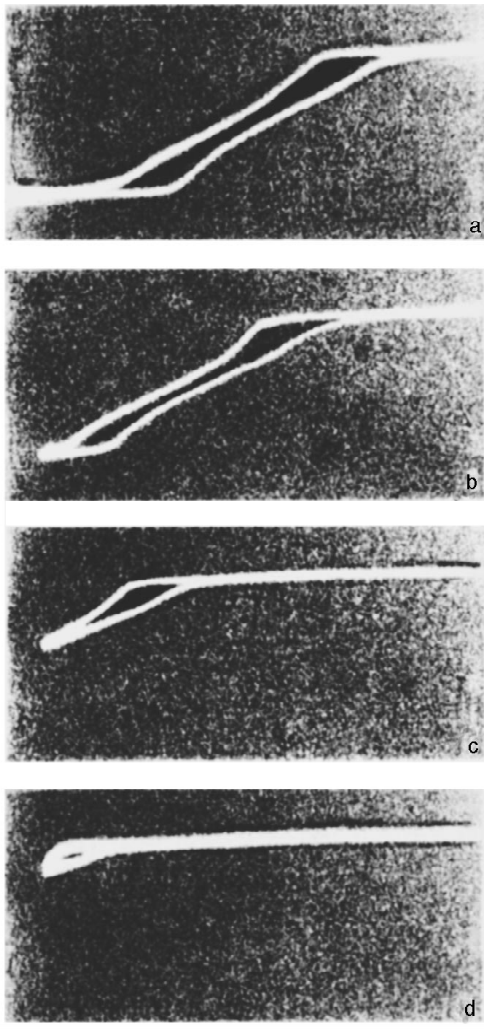

FIG. 1. Hysteresis loops. H_p, Oe: a — 0, b — 56, c — 87, d — 123.

in the corresponding polarity of the field H_z. For $H_p = H_{p2}$ (Fig. 1c) the magnetization reversal signal decreases by half, while in a field H_{p3} it vanishes (Fig. 1d).

The critical fields H_{p1}, H_{p2}, and H_{p3} are indicated on the magnetic state phase diagram for the (210) film in Ref. 5 constructed for the same sample from observations of the domain structure (Fig. 2). In fields $H_p > H_{p1}$ the transition to the saturated state of the film for a single polarity becomes complete, and in fields $H_p > H_{p3}$ nucleus formation cannot occur. Therefore a dc field H_p interrupts the magnetization reversal of the (210) film at definite stages.

To construct the phase diagrams (for an arbitrary fixed angle ϕ) we can use the fields H_{zs}, H_{p1}, H_{p2}, and H_{p3}, measured from the hysteresis loops for various amplitudes of the magnetization reversal field H_{z0}. The observations showed that in low fields the phase diagrams of these samples remain as straight lines over a wide range of ϕ.

When the angle ϕ is changed the shape of the hysteresis loop is preserved, but the critical fields H_{p1}, H_{p2}, and H_{p3} are changed. The slope of the phase curves θ' has a specific value for each angle ϕ. To characterize the slope we can use the magnitude of the field H_{ps} that corresponds to the saturation field of the sample for $H_z = 0$; then $\tan \theta' = H_{zs}/H_{ps}$. The dependence of H_{ps} on the angle ϕ is shown in Fig. 3. Near the x axis the value of H_{ps} and consequently the slope of the phase diagram are essentially unchanged.

There are two directions in the plane of the film, $\phi = 90°$ and $260°$, where the field H_p does not cause the hysteresis loop to shift along the field H_z, while with increasing H_p the amplitude of the magnetization reversal sig-
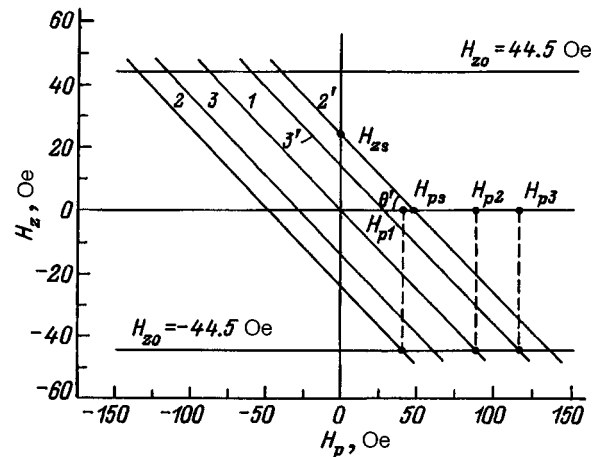

FIG. 2. Phase diagram of the magnetic states of a (210) film for $\phi = 0$. *1* — demagnetized state of the film, *2* and *2'* — saturation field, *3* and *3'* — nucleation field.
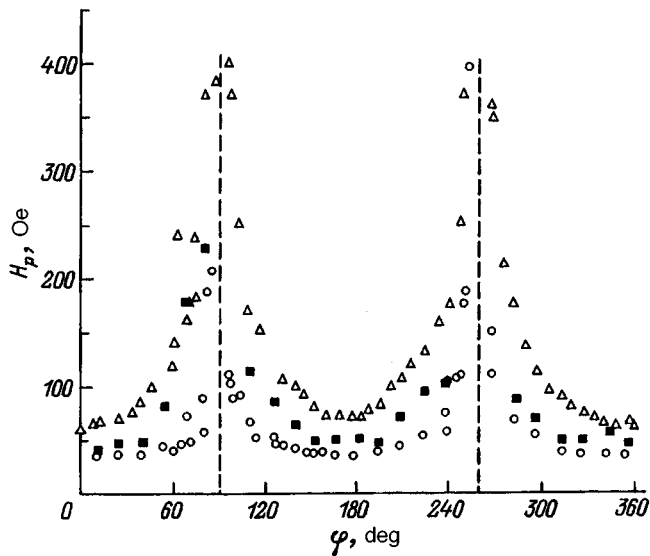
FIG. 3. Angular dependence of the critical in-plane fields H_p. $\triangle$ — collapse field H_{pcol}, $\blacksquare$ — H_{ps}, $\bigcirc$ — elliptical instability field H_{pe}.

nal steadily decreases to zero. The proposed method for measuring has allowed us to measure these angles to an accuracy of a fraction of a degree.

In Fig. 3 we show the angular dependence of the collapse field H_{pcol} and the elliptical instability field H_{pe}, measured on the basis of the domain structure for $H_z = 0$. The behavior of all the critical fields, H_{pcol}, H_{pe}, and H_{ps}, as a function of the angle ϕ is the same . The angular dependences of the critical fields are consistent with the observed changes in the domain structure and the means of its nucleation.

DOMAIN STRUCTURE IN (210) FILMS

The principal types of domain structure in (210) films created by magnetic fields H_z and H_x have been studied in Ref. 5. The following domain structures have been observed: regular stripe domain structure through the thickness of the film, a hexagonal lattice of inclined magnetic bubbles, and a mixed domain structure.

The equilibrium structure for the samples studied here is the regular stripe domain structure, oriented along the x axis. The remaining domain structures are also stable, but the previous history is an important factor in nucleating them, and special conditions are necessary (i. e., a combination of dc fields H_z and H_x with periodically repeated pulses of the field H_z of a certain amplitude).

When the driving fields are oriented in the zx plane the sample is magnetized mainly through the translational mo-

tion of domain walls that are nearly parallel to the x axis with an invariant period P_0 up to $m_z \approx 0.4$. Then the irreversible motion of magnetic dislocations results in an abrupt increase in the period. The saturated state is obtained by collapse of the free ends of the stripe domains or of the magnetic bubbles in the field H_{col}.

The nature of the quasistatic magnetization, the orientation of the stripe domain walls, and the elongation of the magnetic bubbles along the x direction are maintained when the in-plane field $H_p(\phi)$ acts in the neighborhood of the x axis ($\Delta \phi = \pm 40°$).

For $+40° \leqslant \phi \leqslant +140°$ and $+220° \leqslant \phi \leqslant +320°$, a change is observed with increasing dc field H_p in the orientation of the stripe domain structure to the direction of the bisector of the angle between the x and the $-y$ axes, which is close to the crystallographic $[\bar{1}2\bar{1}]$ axis. The reorganization of the domain structure from the equilibrium state caused by the field H_p for $\phi = 90°$ is illustrated in Fig. 4. The new structure gradually develops passing through several stages of instability of the domain wall form. In weak fields sinusoidal distortions of the domain wall profile occur, and then the symmetry of the distortions is broken up and the domain walls are built up and become elongated in one preferred direction at the apices of the zig-zags. The period P_1 of the new stripe structure corresponds to the period of the sinusoidal distortions in the shape of the domain walls, which in turn depends only weakly on the external parameters (the orientation and magnitude of the field H_p). For the films studied here, $P_1 \approx P_0/1.7$.

Domain structures with two orientations of the stripes at an angle of $45°$ have been observed by Dikshteĭn et al.[6] in (110) films ($\theta_0 \approx 10°$) with an in-plane field and with steady reduction of the temperature.

When the magnetic field H_p is directed along the directions $\phi = 90°$ and $260°$ the domain structure is oriented along the $[\bar{1}2\bar{1}]$ direction and vanishes when $\mathbf{M}_x$ is rotated nearly into the plane of the film. For these two directions, when the field H_p is removed, an amorphous lattice of magnetic bubbles of both polarities is formed out of the saturated state ($H_z = 0$), which indicates a second-order phase transition. The magnitude of this field H_p (for sample 2, $H_p = 980$ Oe) can be considered as equal to the magnetic anisotropy field H_k. The corresponding anisotropy constant, $k = 330$ erg/cm³, is in good agreement with the uniaxial anisotropy constant, $k_u = 312$ erg/cm³ obtained for this sample in Ref. 1. For their data, the orthorhombic and cubic anisotropy constants are, respectively, $k_p = 3340$ erg/cm³ and $k_c = 1555$ erg/cm³.

In a field H_p oriented close to the y direction the quasistatic magnetization occurs not only by motion of the domain
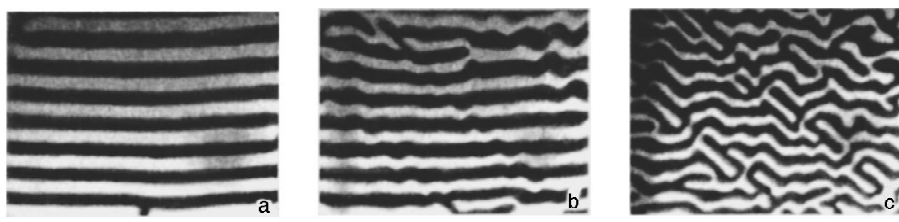

FIG. 4. Restructuring of the domain structure in fields $H_z = 0$, and $H_p = 0$ (a) 210 (b), and 532 (c) Oe.

walls, but also by the rotation of the magnetization, evidence for which is the decrease in the domain structure contrast and in the z component of $\mathbf{M}_s$, obtained from measurements of the Faraday rotation.

By observing the domain structure during magnetization reversal in a magnetic film by an ac field H_z in a dc field H_p oriented at an angle of 40° to the x axis one can see that the new domain structure forms in the interior of the old stripe domain structure and coexists with the latter in some interval of angles ($\Delta\phi \approx 10°$). For other orientations of the field H_p one of the two preferred directions of orientation of the stripe domains was followed: either along the x axis or at an angle of −45° to it.

The principle mechanism for the formation of nuclei of reverse magnetization in these films in the absence of ''shaking'' is the ingrowth of elongated domains, as a rule, from the edge of the sample. In addition, for each angle ϕ there exists a range of values of dc fields H_z and H_p where magnetic bubbles of a certain polarity are formed from the single-domain state. The process starts with the formation of elongated domains in locations of local nonuniformity of $\mathbf{M}_s$. With increasing magnetization reversal field H_p nuclei in the form of magnetic bubbles are formed near the ends of the domains and propagate frontally in the region of uniform magnetization of the film, forming chains oriented along the x axis if $\phi \lesssim \pm 40°$ or at an angle of −45° to the x axis for other orientations of H_p. In a field of elliptical instability, H_{pe} the magnetic bubbles either merge together into stripe domains or a domain structure is formed with a high density of magnetic dislocations (mainly this refers to the free ends of the elongated domains).

These investigations have established that the angular dependence of the critical fields H_{p1}, H_{p2}, and H_{p3} and the changes in the hysteresis loop indicate that the orthorhombic symmetry is anisotropic in epitaxial iron garnet films with the (210) orientation.

We have observed two preferred directions of orientation of the stripe domain structure, which nearly coincide with the $[\bar{1}20]$ and $[\bar{1}2\bar{1}]$ crystallographic orientations.

For the two orientations the field $H_p = H_k$ (for $\phi = 90$ and 260°) and in zero field H_z the domain structure is nucleated in a second-order phase transition.

These results may be important for explaining the experimental data on the alternating-sign slope of the domain walls in a field $H_p = H_x$ in (210) films.[7]

[1)]Deceased

[1] A. M. Balbashov, F. V. Lisovskiĭ, and E. G. Mansvetova, Preprint No. 25 (500), IRE, Academy of Sciences of the USSR, Moscow, (1988).
[2] V. A. Bokov, V. V. Volkov, N. L. Petrichenko et al., Fiz. Tverd. Tela (Leningrad) 31(11), 310 (1989) [Sov. Phys. Solid State 31, 2017 (1989)].
[3] A. M. Balbashov, I. E. Dikshteĭn, F. V. Lisovskiĭ et al., Mikroélektronika 19(1), 45 (1990).
[4] E. N. Il'icheva A. G. Shishkov, A. M. Balbashov et al., Zh. Tekh. Fiz. 63(11) 143 (1993) [Tech. Phys. Lett. 38, 1005 (1993)].
[5] E. N. Il'icheva, A. V. Klushina, N. N. Ustanov et al., Vest. Mosk. Univ. Fiz. Astronomiya 35(2), 59 (1994).
[6] I. E. Dikshteĭn, F. V. Lisovskiĭ, E. G. Mansvetova et al., Fiz. Tverd. Tela (Leningrad) 28, 2494 (1986) [Sov. Phys. Solid State 28, 1394 (1986)].
[7] E. N. Ilyicheva, A. V. Klushina, V. K. Peterson et al., J. Magn. Magn. Mater. 148, 251 (1995).

Translated by J. R. Anderson

Size dependence of the properties of boron nitride crystals

G. P. Bogatyreva, E. R. Zusmanov, N. V. Kotova, V. M. Maevskiĭ, and A. B. Roĭtsin

Institute of Semiconductor Physics, Ukrainian Academy of Sciences, 252650 Kiev, Ukraine

The physical properties of powder samples of high-strength cubic boron nitride are studied for particle sizes from 2 to 200 μm. The studies include the impurity (elemental) composition of the bulk and the surface composition, the magnetic, electrophysical, microwave-spectroscopic characteristics, as well as the adsorption-structural characteristics, the density and the physical-chemical properties of the material. It is shown that the physical properties depend to a large degree on the particle size of the powder. The nature of the observed effects is discussed. © *1997 American Institute of Physics.* [S1063-7842(97)00806-4]

1. Many studies have been made of the properties of disperse ultrahard materials that are finding widespread use in machine construction. Among these materials is cubic boron nitride (CBN). It suffices to say that CBN is the most thermally resistant material, and is generally used where diamond cannot be used (for large loads and high temperatures).

In addition, powder samples for various reasons are widely used for studying the physical properties of single crystals, in particular for studying their defect structure.[1,2] It is usually assumed that grinding up the crystals has no great effect on their bulk properties. For example, the bulk centers (defects) contained in single crystals do not change in structure or symmetry. This is a necessary condition for the use of powders for studying the bulk properties of crystals.

Moreover, it has been demonstrated that in a number of cases the properties of the ground samples are quite different from those of coarser samples. For example, it has been shown[3] that when alkali halide crystals with manganese impurities are ground up, the axial centers based on the Mn^{2+} ion turn into cubic centers: the vacancy that compensates the excess charge of the Mn^{2+} ion is separated from the latter. Another effect has been observed when diamond samples are ground up:[4,5] the intensity of the EPR spectrum changes in intensity but not in its parameters. It has also been observed[6–8] that the intensity of the EPR spectrum of the dangling bonds on a silicon surface depends on the grain size of the diamond powder used to polish it. Finally, we note investigations in which the width of the EPR line of Cr^{3+} and Fe^{3+} impurity ions in corundum depends on the corundum sample size.[9]

These size effects show that the surface influences the properties of the material and this influence must necessarily be taken into account in investigations of these properties. On the other hand, with powder samples one can study the structure and properties of the surface itself. The present investigation is focused on the properties of powder samples of high-strength CBN and how they depend on the particle size. We studied the elemental impurity content, including the surface composition, the adsorption-structural, the physical-chemical, the magnetic, electrophysical, and microwave resonance properties.

2. The samples were type KR grinding powder and KM micropowder with various particle sizes (Table I). The samples were prepared at the Institute of Ultrahard Materials, Ukrainian National Academy of Sciences, from the hexagonal modification of boron nitride at high pressures and temperatures. Manganese was added as a process stimulator. The impurity content of the samples was determined by the method of spark mass spectroscopy. The results of this investigation are given in Table I . The surface composition was determined by Auger electron spectroscopy. The adsorption-structural characteristics were investigated by the BET (Brunauer, Emmet, and Teller) method using the isotherms obtained by low-temperature adsorption of nitrogen on the Akusorb device.[10]

Table II lists the specific surface area S_{BET} the total pore

TABLE I. Impurity composition and surface composition of CBN powder.

Sample	Sample name, Particle size, μm	Si	Fe	Mg	Al	Ca	Ti	B	N	C	O
		Impurity concentration, %						Surface composition, at. %			
a	KP 200/160 (batch 1)	0.09	0.02	0.8	0.01	0.05	0.01	20.5	16.8	42.9	19.8
b	KP 100/80	0.10	0.02	0.56	0.01	0.04	0.01	27.6	19.5	41.6	11.2
c	KM 20/14	0.07	0.01	0.23	0.01	0.05	0.01	37.9	25.7	30.9	5.5
d	KM 7/5	0.09	0.01	0.26	0.01	0.05	0.01	31.5	25.5	36.3	6.8
e	KM 5/3	0.03	0.01	0.23	0.01	0.05	0.01	47.3	29.5	17.1	6.1
f	KM 3/2	0.02	0.04	0.11	0.01	0.04	0.01	47.0	30.9	15.8	6.3
g	KP 200/160 (batch 2)	0.09	0.02	0.8	0.01	0.05	0.01	20.5	16.8	42.9	19.8
*h**	"*g*," ground	–	–	–	–	–	–	–	–	–	–

*Per cent content of various particle sizes: $d \approx 200$ μm — 24%, $d \approx 100$ μm — 24%, $d \approx 60$ μm — 48%, $d \approx 20$ μm — 4%.

TABLE II. Absorption-structural, physical-chemical, Magnetic, and electrophysical characteristics of CBN powder.

Sample	ρ, g/cm^3	S_{BET}, m^2/g	V_S, 10^{-5}ml/g	ΔG_S, mJ/mole·g	$-\Delta G_S$ sp, mJ/mole m^2	χ, 10^{-8} m^3/kg	ρ', 10^{13} Ω	tan δ, 10^{-4}
a	3.52	0.014	47.21	-2.96	197.33	8.2	6.6	1066
b	3.51	0.018	49.23	-3.13	173.89	3.3	2.2	1307
c	3.50	0.167	62.99	-0.65	3.89	0.9	3.6	443
d	3.49	1.127	645.1	-6.42	5.7	2.5	2.8	224
e	3.53	1.431	435.6	-9.64	9.64	2.5	3.3	341
f	3.54	1.816	728.6	-4.7	4.7	1.9	2.6	363
g	3.52	0.014	47.21	-2.96	197.33	8.2	6.6	1066
h	–	–	–	–	–	–	–	–

volume V_S, and the density ρ of the material. This table also gives the results of measurements of the physical-chemical, magnetic, and electrophysical parameters: the free energy of saturation by water vapor, ΔG_S, the specific energy of saturation by water vapor ΔG_S sp, the characteristic hydrophylic degree, the specific magnetic susceptibility χ, the resistivity ρ', and the electrical loss tangent, tan δ.

3. In addition to the methods enumerated above, electron paramagnetic resonance was used as a structurally sensitive method by which one can study on the microscopic level the defect structure in the bulk and on the surface.[11] In particular, by observing the changes in the EPR spectra for various external influences on the sample (temperature, gases, radiation, etc.) one can draw conclusions about the model of surface defects and about the changes that occur in their structure as a result of this influence. Similar studies were made on, for example, silicon for the purpose of determining the nature of the defect (with a g factor equal to 2.0055), called a "dangling bond." For a number of reasons[11] it is easier to identify bulk defects than surface defects.

The measurements were carried out on an RÉ-1306 three-centimeter EPR spectrometer at room temperature. The whole EPR spectrum is in general a superposition of three types of lines. Line L_1 with a g factor of about 2 and a width[1)] $\Delta H = 35$–130 Oe; line L_2, a narrow line ($\Delta H = 5$–7 Oe) with $g \approx 2$; and line L_3, which was observed in a broad range of resonance magnetic fields (from 0 to 6000 Oe) and has a ΔH of several thousand oersteds. Some of the L_1 and L_3 lines have a multicomponent structure. The widths of the L_1 and L_2 lines in the EPR spectrum (hundreds of oersteds) are small compared to the region occupied by the L_3 lines (thousands of oersteds), and the peak intensities of the L_1 and L_2 lines are an order of magnitude higher than that of line L_3. Therefore the spectra corresponding to these lines are given separately below with different scales along the axes.

The EPR spectra of the polishing-powder sample and the micropowder sample are qualitatively different, whereas the shape of the spectrum for the powders within each type for various particle size are similar and differ only by their widths and intensities. For example, the spectra from the polishing powder contains the lines L_1 and L_2, with L_1 having a poorly defined structure. For the micropowder sample, the line L_2 is not observed, and line L_1 is structureless. A typical shape of the spectrum for the polishing powder is shown in Fig. 1A (sample a), and that for the micropowder sample is shown in Fig. 1B (sample c) (the notation for designating the samples is given in Table I). The parameters of the lines L_1 and L_2 for the polishing powder and the micropowder samples are given in Table III. The shape of line L_1 is close to Lorentzian, and as the particle size decrease it is easily approximated by the sum of two Lorentzians with different widths. In Fig. 2 we show the dependence of the parameters of the line L_1 on the sample dimension d; the parameter d was taken to be the arithmetical mean of the extreme particle size (for example in the case of KR 200/160, $d = 180$ μm). Finally, in Fig. 3 we show the line L_3, whose parameters are listed in Table IV.

4. In discussing the results we note first of all that previously the EPR method has been used to study both hexagonal BN (Ref. 12)[2)] and cubic BN (Refs. 13–19). The EPR spectra presented in Refs. 13–19 or their main parameters (g factor, ΔH) on the whole are similar to ours for the L_1 and L_2 lines (Table III). The sometimes large differences (for
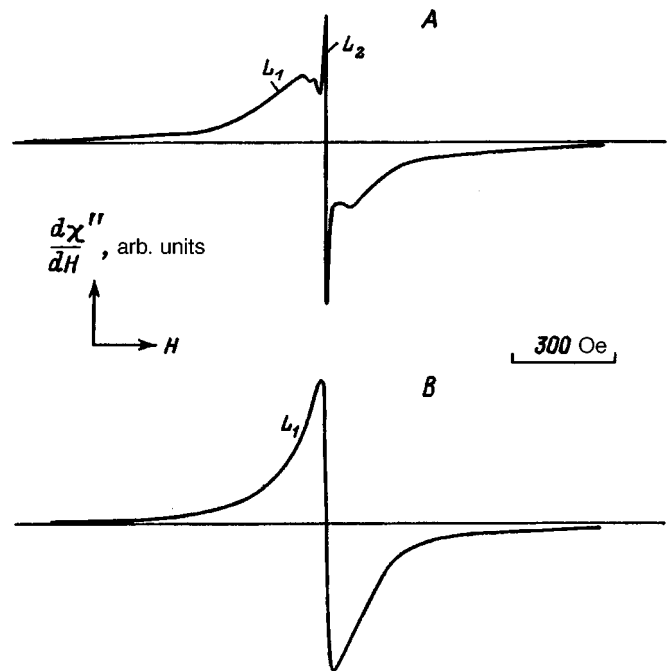

FIG. 1. Typical shape of the L_1 and L_2 lines of the EPR spectrum of cubic boron nitride.

TABLE III. Parameters of the L_1 and L_2 lines of the EPR spectra.

Sample	ΔH_1, Oe	ΔH_2, Oe	$I_1(S_1)$, arb. units	I_2, arb. units	g_1	g_2
a	108 ± 15	5.6 ± 0.5	0.33(0.76)	0.66	2.0026 ± 0.0015	2.0030 ± 0.0005
b	114 ± 5	5.6 ± 0.5	0.36(0.74)	0.39	2.0040 ± 0.0015	2.0032 ± 0.0005
c	45 ± 5	–	0.72(0.92)	–	2.0061 ± 0.0008	–
d	41 ± 5	–	1.00(0.95)	–	2.0062 ± 0.0008	–
e	40 ± 5	–	0.91(0.68)	–	2.0062 ± 0.0008	–
f	42 ± 5	–	0.58(0.44)	–	2.0057 ± 0.0008	–
g	116 ± 15	5.6 ± 0.5	0.31(1.00)	0.61	2.0040 ± 0.0015	2.0029 ± 0.0005
h	41 ± 5	6.5 ± 0.5	0.38(0.72)	0.68	2.0052 ± 0.0015	2.0038 ± 0.0005

Note. The values I_1 and I_2 correspond to the peak intensities of the lines, and the values of S_1 in parentheses to the integrated intensities.

example, the differences in ΔH for the L_1 line are as much as a factor of three), are evidently related to the difference in the powder size, for which data are frequently not given. The L_3 lines we have observed have no analog in the literature in the spectra of any modification of BN.

The most characteristic defects in hexagonal BN are the so-called one- and three-boron centers, in which an unpaired electron enters into a hyperfine interaction with either one or three boron nuclei (^{10}B, ^{11}B). The nature of only the three-boron center has been established — a positively charged nitrogen vacancy that has trapped an electron (the analog of an F center in alkali halide crystals). By analogy with this model, a four-boron model has been proposed for cubic BN (Ref. 15): a positively charged nitrogen vacancy located in the tetrahedral surroundings of four boron atoms, which has trapped an electron. This center, as it is assumed, is responsible for the lines analogous to our narrow line L_2. Although its hyperfine structure is not split (unlike the centers in hexagonal BN), the model is convincing in that in cubic BN the nitrogen atom also carries a negative charge.[20] It might be that the model could be established finally by means of electron-nuclear double-resonance experiments.[21] A model for the center responsible for the wide line (the analog of L_1) is not yet known. It is only assumed that it may be related to complex defects (groups of F centers in combina-

tion with different impurities), whose nature depends on the means of preparing the samples and their subsequent interaction with various external agents.

We have observed the L_3 line in BN for the first time. Similar lines, however, have been observed previously in studies of synthetic diamond.[22,23] They were attributed to ferromagnetic inclusions added to the starting mixture as an aid to crystallization in synthesizing the diamond. Considering the external similarity of the spectra shown in Fig. 3 to the spectra published in Refs. 22 and 23, and the means of preparing the BN and diamond samples and the analysis of their elemental composition (Table I), we can conclude that the L_3 lines we have observed are connected with metal inclusions.

One can see from Figs. 1–3 and Tables III and IV, that the EPR spectra from samples of various sizes are generally quite different from one another. For example, the intensity of the L_1 line has a maximum near $d \approx 6$ μm, whereas its width remains nearly constant for small values of d and increases rapidly with increasing d. Also quite different are the EPR spectra of ferromagnetic inclusions from samples of different sizes (Fig. 3). Here we can distinguish three groups of lines: from samples a, b, and g; from samples c, d, e, and h; and from sample f. Within each group, which is determined by the particle size, the spectra are qualitatively simi-
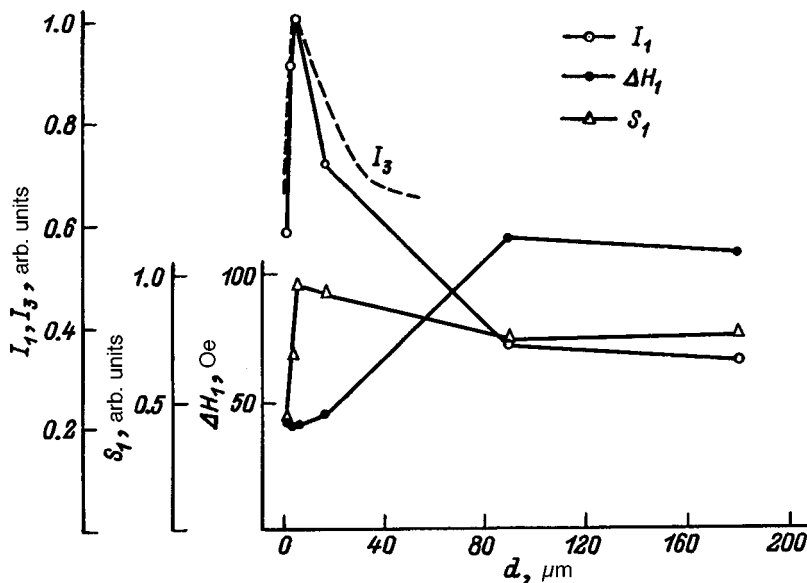

FIG. 2. Dependence of the parameters of the L_1 line on the particle size of powder cubic boron nitride (the notation for the parameters is given in Table II).
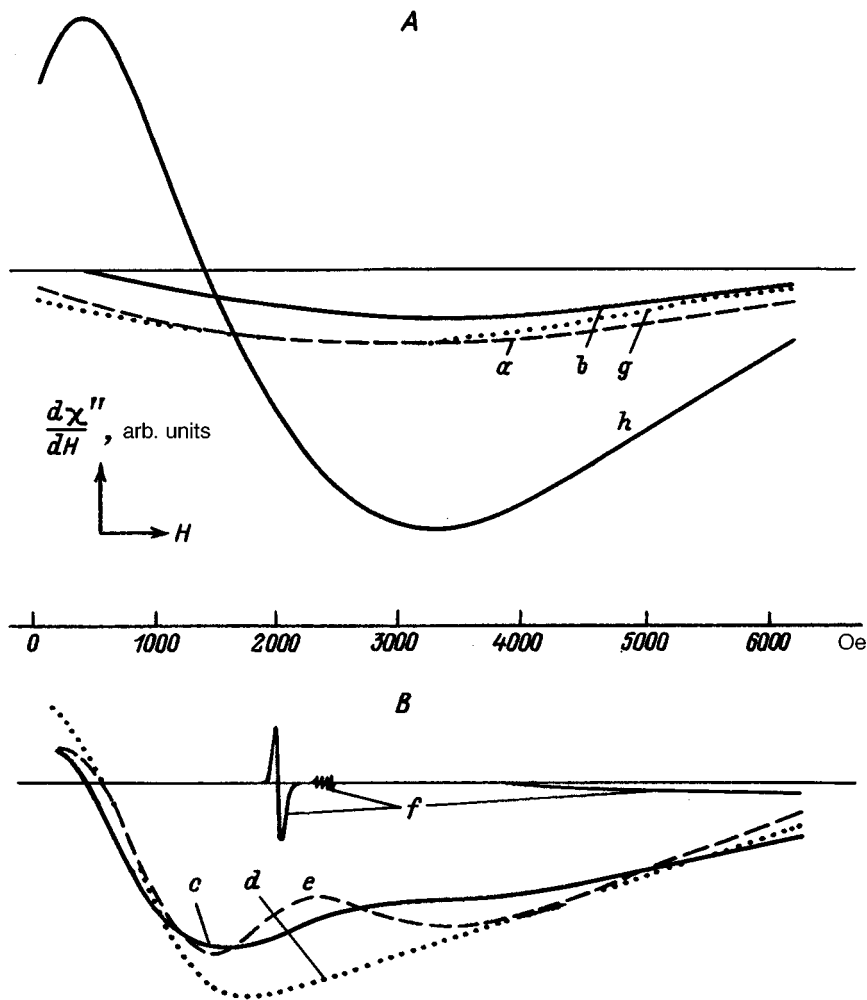
FIG. 3. Lines L_3 of the EPR spectrum of cubic boron nitride: A — polishing powder, B — micropowder (the notation for the samples is given in Table I).

lar. However, the parameters of the L_2 line depend only slightly on the sample size. The fact that line L_2 is scarcely seen at all for small particles is probably related to the fact that they are "buried" in the L_1 lines, whose intensities increase sharply with decreasing particle size. It can thus be concluded that the L_2 lines are associated with paramagnetic centers located mainly in the bulk, whereas the L_1 and L_3 lines are associated with centers located near the surface.

The EPR data correlate with data from other methods. For example, the large loss tangents indicate a high defect concentration in the crystals. In the EPR spectra this corresponds to broad lines and low intensities. It should be noted that the maximum intensity of the L_1 line corresponds to the minimum in tan δ. The large specific surface area of BN for small values of d is a favorable factor for the increase in the number of free bonds, which causes an increase in the number of unpaired electrons and as a consequence an increase in the intensity of the L_1 line, which is observed with decreasing d. An increase in the number and volume of the pores also brings about the same effect in the EPR spectra.

In addition to the parameters of line L_1, Fig. 2 also contains a plot of I_3, which shows how the intensities of the EPR lines of the "dangling-bond" paramagnetic centers on the silicon surface depend on the diamond powder grain size used for polishing the Si surface so as to create these centers

on it. The curve if I_3 was measured in Ref. 6 and confirmed in Ref. 8. The maxima of the curves of I_1 and I_3 are displaced in the vertical direction for clarity. It can be seen that the curves of I_1 and I_3 are similar in form. Their extrema fall at the same size $d \approx 6$ μm, which indicates the generality of the phenomenon. In general terms this can be interpreted in the following way. As the particle size decreases, the total surface area increases and the EPR signal also increases.

TABLE IV. Parameters of the L_3 line of the EPR spectra.

Sample	H_0, Oe	g_{eff}	ΔH, Oe	Notes
a	–	–	2700*	
b	650	10.4	3400*	
c	500	13.5	1750*	has structure
d	600	11.2	1500*	" "
e	630	10.7	1800	
f	2120	3.18	60	Still a four-
	4000	1.685	6500*	component line near $H = 2430$ Oe
g	–	–	2650*	
h	1500	4.49	2600	

Note. H_0 is the magnetic field for zero intensity of the derivative of hte absorption line; the values ΔH, indicated by the asterisk correspond to the position of the lower extremum.

However, for particle sizes (a few μm) corresponding to the surface region,[24] the paramagnetic centers disappear because of the increase in the area of contact of the particles, which leads to a decrease in the open surface, and the partial damage to the surface structure, which reduces the EPR signal.

[1)]The distance between the maximum slopes.
[2)]For references of other work on this modification of BN see Ref. 12.

[1]G. M. Zhidomirov, Ya. S. Lebedev, S. N. Dobryakov *et al. The Interpretation of Complex EPR Spectra* [in Russian] (Nauka, Moscow, 1975).

[2]M. V. Vlasova and N. G. Kakazeĭ, *EPR in Mechanically Damaged Solids* [in Russian] (Naukova Dumka, Kiev, 1979).

[3]V. M. Maevskiĭ, N. I. Vitrikhovskiĭ, E. R. Zusmanov *et al.* Ukr. Fiz. Zh. **38**, 442 (1993); Pis'ma Zh. Tekh. Fiz. **19**(8), 46 (1993) [Tech. Phys. Lett. **19**, 206 (1993)]; Zh. Tekh. Fiz **64**(5), 107 (1994) [Tech. Phys. **39**, 514 (1994)].

[4]M. J. A. Smith and B. R. Angel, Philos. Mag. **15**, 783 (1967).

[5]J. H. N. Loubser, Solid State Commun. **22**, 767 (1977).

[6]A. Taloni and W. J. Rogers, Surf. Sci. **19**, 371 (1970).

[7]N. N. Gerasimenko, A. V. Dvurechenskiĭ, and L. S. Smirnov, Fiz. Tekh. Poluprovodn. **6**, 987 (1972) [Sov. Phys. Semicond. **6**, 862 (1972)].

[8]B. P. Lemke and D. Haneman, Phys. Rev. B **17**, 1893 (1978).

[9]R. S. De Biasi and D. C. S. Rodriguez, J. Mater. Sci. **16**, 968 (1981); J. Mater. Sci. Lett. **2**, 210 (1983).

[10]V. G. Aleshin, A. S. Smekhnov, G. P. Bogatyreva, and V. B. Kruk, in *The Chemistry of the Diamond Surface* [in Russian] (Naukova Dumka, Kiev, 1990).

[11]A. B. Roĭtsin and V. M. Maevskiĭ, Usp. Fiz. Nauk **159**, 297 (1989) [Sov. Phys. Usp. **32**, 891 (1989)]; *Radio-Frequency Spectroscopy of Solid Surfaces* [in Russian] (Naukova Dumka, Kiev, 1992).

[12]A. V. Kabyshev, V. M. Ketskalo, Yu. P. Surov *et al.* Porosh. Metall. No. 8, 27 (1992).

[13]V. S. Neshpor, G. V. Samsonov, L. I. Fel'dgun *et al.*, Izv. Akad. Nauk SSSR Neorg. Mater. **7**, 173 (1971).

[14]M. V. Vlasova, N. G. Kakazeĭ, and G. I. Savvakin, Izv. Akad. Nauk SSSR Neorg. Mater. **16**, 2160 (1980).

[15]L. A. Shul'man, K. A. Tikhonenko, and V. K. Gerasiminko, *The Effect of High Pressure on the Structure and Properties of Materials* [in Russian] ISM, Akad. Nauk UkrSSR, Kiev, 1993, pp. 29–34; Ukr. Fiz. Zh. **30**, 578 (1985).

[16]V. B. Shipilo, A. E. Rud' G. V. Leushkina *et al.*, Izv. Akad. Nauk SSSR Neorg. Mater. **23**, 1993 (1987) Fiz. Tekh. Vys. Davlen. No. 28, 29 (1988).

[17]V. L. Solozhenko, V. V. Chernyshev, G. V. Fetisov *et al.*, J. Phys. Chem. Solids **51**, 1011 (1990).

[18]V. B. Shipilo, A. E. Rud', N. G. Anichenko *et al.*, Izv. Akad. Nauk SSSR Neorg. Mater. **26**, 1643 (1990); Fiz. Tekh. Vys. Davlen. No. 33, 43 (1990).

[19]G. P. Bogatyreva, N. V. Kotova, V. L. Gvyazdovskaya *et al.*, Dokl. Akad. Nauk Ukrany Ser. Mat. Estestv. Tekh. Nauki No. 10, 116 (1994).

[20]C. A. Coulson, L.B. Redei, and D. Stocker, Proc. R. Soc. London Ser. A **270**, 357 (1962).

[21]M. D. Glinchuk, V. G. Grachev, M. V. Deĭgen *et al.*, in *Electrical Effects in Microwave Spectroscopy* [in Russian] (Nauka, Moscow, 1981).

[22]G. P. Bogatyreva, L. Yu. Vladimirova, E. R. Zusmanov *et al.*, Poverkhn. Fiz. Khim. Mekh. No. 1, 95 (1995); *ibid*, No. 5, 7.

[23]G. P. Bogatyreva, G. A. Bazaliĭ, E. R. Zusmonov *et al.*, Zh. Tekh. Fiz **65**(2), 55 (1995) [Tech. Phys. **40**, 146 (1995)].

[24]V. M. Maevskiĭ, N. I. Vitrikovskiĭ, E. V. Mozdor *et al.*, Poverkhn. Fiz. Khim. Mekh. No. 1, 128 (1989).

Translated by J. R. Anderson

Characteristic electron thermalization time in dielectric media

S. V. Ivliev, V. K. Lyapidevskiĭ, and M. I. Ryazanov

Moscow Engineering Physics Institute, 115409 Moscow, Russia

An approximate solution is obtained for the thermalization transport equation for electrons with an energy below the atom ionization potential. The results are used to estimate the thermalization time and the time of removal of the ionization electron from its atom in the gaseous and liquid states of inert gases. © *1997 American Institute of Physics.*
[S1063-7842(97)00906-9]

INTRODUCTION

Primary ionization electrons ejected from atoms by fast particles move far from their ion of origin, so that their effect on the ion is very weak. Therefore local charge neutrality is disrupted for a certain time in the immediate vicinity of these ions in the dielectric, and this disruption plays a role in the formation of the particle track. The time during which the local charge neutrality is disrupted depends on the thermalization time of the primary electrons. Therefore it is of interest to estimate the thermalization time and discuss the various mechanisms for the thermalization of hot electrons with energies lower than the ionization potential of the atom ($I = 10$ eV), but higher than the thermal energy ($kT = 0.025$ eV). Numerical calculations of the thermalization time of fast electrons in gases have been carried out in Refs. 1–3; simplifying assumptions made in these calculations were carried out by the method of moments in Ref. 4 and it was shown that for the hard-sphere model the discrepancy is 25%, but in other cases it is less.

An exact theory of thermalization of electrons requires knowledge of the exact cross sections for various interactions of the electrons with matter over the entire energy range covered by the investigation. However, these cross sections are quite difficult to calculate, and in many cases they have not been calcuated. Experimental measurements of the cross sections also cannot give a complete picture of their energy and angular dependences over the entire range of energy of interest. In this situation exact results of theory must either be supplemented with exact data on the cross sections, or interpolations must be made over the available data on the cross sections.

On the other hand, the thermalization time of particular interest is clearly some averaged characteristic of the material, and to calculate it does not require knowledge of all the details of the behavior of the cross sections. From this point of view, the item of particular interest is the development of approximate methods of analyzing the situation, where one does not need to know the detailed behavior of the cross sections, but only some general integrated characteristics of the material, for which the calculations or measurements are not so difficult.

TRANSPORT EQUATIONS FOR THERMALIZATION OF HOT ELECTRONS

For electrons with an energy lower than the ionization potential of the atom, their wavelength is greater than atomic dimensions, and therefore scattering from atoms is nearly isotropic, and the energy loss comes about by the recoil of atoms in elastic collisions, the excitation of molecules in inelastic collisions, and the generation of low-energy collective excitations of the material, in particular, phonons. To describe the behavior of the electrons that are being slowed, we introduce the electron distribution function over the position and momentum, $f(\mathbf{r},\mathbf{p},t)$. If electron trapping in traps in the material is neglected, then the transport equation obeyed by the electron distribution function obeys will have the form[5]

$$\{d/dt + \mathbf{v} \cdot d/d\mathbf{p} + \mathbf{F} \cdot d/d\mathbf{r}\} f(\mathbf{r},\mathbf{p},t)$$

$$= \int d^3 p' \{w(\mathbf{p},\mathbf{p}') \delta[E - E' + \delta E(\mathbf{p},\mathbf{p}')] f(\mathbf{r},\mathbf{p}',t)$$

$$- w(\mathbf{p},\mathbf{p}) \delta[E - E' + \delta E(\mathbf{p},\mathbf{p})] f(\mathbf{r},\mathbf{p},t)\}, \tag{1}$$

where $E = p^2/2m$; $E' = p'^2/2m$;, $\mathbf{v} = \mathbf{p}/m$; m is the electron mass, $w(\mathbf{p},\mathbf{p}') \delta[E - E' + \delta E(\mathbf{p},\mathbf{p}')]$ is the probability of a transition of an electron from a state with momentum p to a state with momentum $\mathbf{p}'$ in an interval of momentum $d^3 p'$ per unit time; $\delta E(\mathbf{p},\mathbf{p}')$ is the energy transferred to the atom in a collision, and $\mathbf{F}$ is the force acting on the electron from the ionized atom.

In the initial stage the kinetic energy of an electron is large compared to the potential energy, and so the action of this force can be neglected. The role of this force later depends on which falls off faster — the kinetic energy of the electron because of its stopping, or the potential energy of the electron because it gets farther away from the ionized atom. Therefore in the first approximation the force $\mathbf{F}$ can be dropped. Using the solution of the first approximation, we shall later estimate the correction due to the Coulomb interaction of the electron with the ion.

In the range of electron energy considered, the energy loss in one collision is relatively small; i.e., we can assume that $\delta E(\mathbf{p},\mathbf{p}') \ll E$. We can therefore transform the transport equation (1) to a Fokker–Planck equation, which describes diffusion in energy space, by transforming from the function

$f(\mathbf{r},\mathbf{p},t)$ to a function that describes the energy distribution of the electrons

$$n(\mathbf{r},E,T)=(1/4\pi)\int d\Omega_{\mathbf{p}}f(\mathbf{r},\mathbf{p},t). \tag{2}$$

From Eq. (1) we obtain an equation for $n(r,E,t)$

$$[d/dt-D(E)\nabla^2]n(\mathbf{r},E,t)$$
$$=\int d^3p'\,\delta E(\mathbf{p},\mathbf{p}')\,\delta(E-E')\{w(\mathbf{p},\mathbf{p}')n(\mathbf{r},E',t)$$
$$-w(\mathbf{p},\mathbf{p})n(\mathbf{r},E,t)\}. \tag{3}$$

Here $(pp'=pp'\cos\theta)$

$$D(E)=(v^2/3)\int d^3p'\,w(\mathbf{p},\mathbf{p}')\,\delta(E-E')(1-\cos\theta). \tag{4}$$

We emphasize that the dependence of the diffusion coefficient on the electron energy is a functional, since in the thermalization the electron energy varies by three orders of magnitude. Since for low energies the cross section for the interaction of an electron with the atoms and molecules is essentially isotropic, we can neglect the cosine under the integral sign in Eq. (4). Then the diffusion coefficient can be expressed in terms of the electron mean free path $l(E)$ by the relation

$$D(E)=\{v/3n\sigma(E)\}=vl(E)/3, \tag{5}$$

where n is the number of atoms per unit volume of the material and $\sigma(E)$ is the total cross section for scattering of an electron with an energy E by a single atom.

We now introduce the energy loss function

$$g(E)=\int d^3p'\,w(\mathbf{p},\mathbf{p}')\,\delta E(\mathbf{p},\mathbf{p}')\,\delta(E-E') \tag{6}$$

and a new variable with the dimensions of length squared, analogous to the ''age'' in the theory of neutron slowing[6]

$$T=T(E)=\int_{E_0}^{E}dE'\,D(E')/g(E'). \tag{7}$$

From Eq. (7) we can also find the inverse dependence $E(T)$, by which it is easy to transform any function of E into a function of the variable T. We go over to a description of the electron distribution by means of the function

$$N(\mathbf{r},T,t)=E(T)^{1/2}g(E(T))n(\mathbf{r},E(T),t), \tag{8}$$

for which, using Eq. (3), we obtain the equation

$$[1/D(T)](dN(\mathbf{r},T,t)/dt)=\{\nabla^2-d/dT\}N(\mathbf{r},T,t). \tag{9}$$

If at the origin of the coordinate system there is a source of particles, which emits at time $t=0$ a certain number of particles with a given energy, then this can be taken as a source density added to Eq. (9), which gives instead of Eq. (9) the equation

$$\{D(T)^{-1}d/dt+d/dT-\nabla^2\}N(\mathbf{r},T,t)=q\,\delta(\mathbf{r})\,\delta(T)\,\delta(t). \tag{10}$$

This equation is solved exactly by means of Fourier and Laplace transforms, and its solution is the function $(T>0)$

$$N(\mathbf{r},T,t)=qr^{-3/2}\exp[-r^2/4T]\delta\{t-T_0dT'D(T')^{-1}\}. \tag{11}$$

From this result it follows that in this approximation the spatial distribution of the thermalizing electrons is Gaussian with a width determined by the age T of the electrons, while the time t and energy E are mutually related by

$$t=\int_0^{T(E)}dT'\,D(T')^{-1}. \tag{12}$$

This kind of relation between the stopping time and the particle energy is a characteristic feature of the ''continuous slowing'' approximation used in the above derivation, in which fluctuations in the electron energy loss in a single collision are neglected. The mean distance the electron moves away from its atom is given by the expression

$$\langle r\rangle=\langle r^2\rangle^{1/2}=(6T(E))^{1/2}, \tag{13}$$

i.e., it is a function of the particle energy, or, invoking Eq. (12), a function of the time t. The advantage of this approximation must be considered the possibility of using Eqs. (12) and (13) derived above for estimating and comparing the thermalization time and the distance traveled by an electron from its atom for various materials that differ from one another by various energy dissipation mechanisms. The specific properties of the stopping medium comes in only in the determination of the diffusion coefficient $D(E)$ and the age $T(E)$.

THERMALIZATION TIME IN INERT GASES

In the electron energy range we are considering, inelastic collisions with inert gas atoms play a small role because of the large excitation energy of these atoms. Therefore in monatomic gases the main mechanism for electron energy loss involves elastic collisions, with energy transfer to the recoil of the atom. For estimates, we shall assume that the cross section for elastic collisions depends only weakly on the energy and angle, and then in the light of this discussion it is easy to obtain (M is the mass of the atom)

$$g(E)=(2Em/M)(v/l);\quad E=E_0\exp\{-6Tm/Ml^2\}. \tag{14}$$

The thermalization time t_{therm} is called the electron time of slowing to a final energy E_T equal to the thermal energy. It follows from Eq. (14) that the thermalization time in an inert gas depends on the initial electron energy $E_0=(1/2)mv_0^2$ through

$$t_{\text{therm}}=(lM/mv_0)\{[E_0/E_T]^{1/2}-1\}. \tag{15}$$

The mean distance the primary ionization electron travels from its atom in an inert gas, according to Eq.(13) is

$$\langle r\rangle=l[(M/m)\ln(E_0/E_t)]^{1/2}. \tag{16}$$

The assumption made regarding the weak energy dependence of the scattering cross section is valid only for the light inert gas atoms, helium and neon, to which relations (15) and (16) are applicable. For heavier inert gas atoms (Ar, Kr, Xe), a sharp minimum exists in the scattering cross section at electron energies of about 0.4 eV (the Ramsauer-Townsend

effect), so that in the integration in Eq. (7) one cannot neglect the energy dependence of the scattering cross section. The existence of a sharp minimum in the cross section means that when an electron reaches an energy of about 0.4 eV it will nearly cease to slow down. This greatly increases the thermalization time of the electron. It should be pointed out that expressions (15) and (16) were derived for the particular case of a cross section that depends only weakly on the angle and the energy, and therefore is inapplicable to the case of a cross section that varies rapidly with the energy. For a crude estimate of the electron thermalization time in heavy inert gases we can use the following reasoning. After a time dt an electron with an energy E undergoes $(v\,dt/n\sigma(E))$ collisions, losing in this time an energy of about

$$dE = -\delta E(\mathbf{p},\mathbf{p}')(v/n\sigma(E))dt$$
$$= -2(m/M)E[2E/m]^{1/2}n\sigma(E)dt. \tag{17}$$

Integrating this expression from the initial energy to the thermal energy E_T it is easy to obtain the following expression for the thermalization time

$$t_{\text{therm}} = (M/2n)(1/2m)^{1/2}\int_{E_T}^{E_0}dE\,E^{-3/2}(1/\sigma(E)). \tag{18}$$

In the particular case of a cross section that does not vary with the energy, the result coincides with Eq. (15). Equation (18) shows that the main contribution to the integral comes from energies near to the lower limit. In other words, the thermalization time is accumulated mainly towards the end of the process, and the existence of a minimum in the energy dependence of the scattering cross section can greatly change the result.

The important stage in determining the distance the electron travels away from its atom is not the final stage in the thermalization, but the initial stage, where the electron can move comparatively long distances in a short time. Therefore a minimum in the energy dependence of the cross section in the final stage of thermalization does not contribute much to the distance from the atom. We can thus infer that for estimating the distance the electron moves away from the atom the approximation of an energy- independent cross section is applicable over a wider range than it is to the electron thermalization time.

EXCITATION OF VIBRATIONAL AND ROTATIONAL DEGREES OF FREEDOM IN DIATOMIC GASES

In diatomic gases energy can be lost to excitation of vibrational and rotational degrees of freedom of the molecule in addition to energy loss to elastic collisions. In the collision with a molecule of this gas the potential energy lost by the electron is discrete, and the interaction cross section is the same in order of magnitude as the cross section for elastic scattering. If we assume as before that the cross section depends only weakly on the energy and angle, and denote by ξ the mean energy loss per collision, we can write instead of Eq. (6)

$$g(E) = \xi(v/l),$$

which, when Eq. (12) is substituted into (13), gives the thermalization time in the form

$$t_{\text{therm}} = (ml/\xi)(v_0 - v_T) \tag{19}$$

and for the mean separation of the electron from its atom we have

$$\langle r \rangle = l[(E_0 - E_T)(2/\xi)]^{1/2}. \tag{20}$$

The loss into excitation of vibrational and rotational degrees of freedom is found to be more important, and this mechanism governs the stopping process as long as the electron energy is sufficient to excite these discrete energy levels. If the energy is lower, then the thermalization proceeds via elastic collisions. The condition that the Coulomb field has little effect in the region of energy loss by excitation of vibrational and rotational levels is satisfied: at lower energies the situation is the same as for the loss to elastic collision.

EXCITATION OF ACOUSTIC PHONONS IN LIQUIDS

Let us now consider the thermalization of electrons in liquid dielectrics through excitation of acoustic phonons. The energy lost in one event has the form

$$\delta E(\mathbf{p},\mathbf{p}') = c|\mathbf{p}-\mathbf{p}'|,$$

where c is the speed of sound.

Substituting this expression into Eqs. (5) and (6) gives relations for the energy loss function

$$g(E) = (4/3)mv^2(c/l) \tag{21}$$

and for the age variable

$$T = (v_0 - v)(l^2/4c). \tag{22}$$

Using Eqs. (21) and (22), one can readily derive an expression for the thermalization time through energy loss to the formation of acoustic phonons

$$t_{\text{therm}} = (3l/4c)\ln(v_0/v_T). \tag{23}$$

The mean separation of the electron from its atom is then

$$\langle r \rangle = l[(3/2c)(v_0 - v_T)]^{1/2}. \tag{24}$$

THE EFFECT OF THE COULOMB FIELD

Let us now estimate the correction associated with the inclusion of the effect of the Coulomb field. If one takes into account the Coulomb field, then the kinetic energy of the particle is supplemented by the negative potential energy $U(\mathbf{r}) = -(e^2/r)$. We shall consider the correction obtained in the region where $|U|$ is small relative to E. We assume that we know the exact energy dependence of the electron separation with allowance for the Coulomb field $E' = E + U$, $\langle r(E') \rangle$. For $|U| \ll E$ we can write

$$\langle r(E+U) \rangle = \langle r(E) \rangle + \langle U \rangle(d\langle r(E) \rangle/dE),$$

where $\langle r(E) \rangle$ and $(d\langle r(E) \rangle/dE)$ are taken for $U=0$, i.e., without the Coulomb field, as in Eq. (13). Then from Eqs. (13) and (7) we have

$$(d\langle r(E) \rangle/dE) = [3/2T(E)]^{1/2}\{D(E)/g(E)\}.$$

We can now write for the average separation of the electron from the atom with allowance for the Coulomb field

$$\langle r(E+U)\rangle = \langle r(E)\rangle\{1 - e^2\langle r(E)\rangle^{-2}[d\langle R(E)\rangle/dE]\}. \quad (25)$$

The condition for the applicability of this relation is that the second term on the right-hand side be small compared to the first term, i.e., the inequality

$$h = e^2\langle r(E)\rangle^{-2}[d\langle r(E)\rangle/dE] \ll 1. \quad (26)$$

According to Eq. (25) the Coulomb attraction of the electron to the ion decreases the distance between the ion and the electron, as one would expect from elementary considerations. It is helpful to note that Eq. (25) refers to the case where there are no other charged particles situated between the ion and the electron. The presence of other charges changes the situation qualitatively, since this then becomes a many-particle problem. Nevertheless, Eq. (25) can be used to estimate the error incurred by neglecting the Coulomb field. For a monatomic inert gas the relative error according to Eq. (14) is

$$h = (e^2/2E_0 l)(m/M)^{1/2}[\ln(E_0/E_T)]^{-3/2}. \quad (27)$$

For electron thermalization in a diatomic gas

$$h = (e^2/l)\xi^{1/2}(2E_0)^{-3/2}. \quad (28)$$

Finally, for electron thermalization in a liquid through scattering by acoustic phonons the relative error incurred by ignoring the Coulomb field is

$$h = (e^2/4l)(c/3)^{1/2}(2/m)^{1/4}E_0^{-5/4}. \quad (29)$$

DISCUSSION OF THE RESULTS

It has been shown that the electron thermalization time depends strongly on the electron energy loss mechanism, and accordingly depends on the initial electron energy differently in different materials. Different energy dissipation mecha-nisms in different materials will dictate that the separation of the electron from its atom will also be quite different in those materials.

The assumption that the energy dependence of the interaction cross section is weak imposes limits the accuracy of the results. The error in this approximation can be assessed by using approximate formulas similar to Eqs. (17) and (18) as the thermalization time. The loss to excitation of vibrational and rotational degrees of freedom of diatomic molecules can be written as in Eq. (18)

$$t_{\text{therm}} = (1/n\xi)(m/2)^{1/2}\int_{E_T}^{E_0} dE/[E^{1/2}\sigma(E)]. \quad (30)$$

For the loss to excitation of phonons in a liquid the thermalization time can be obtained as in Eq. (18) in the form

$$t_{\text{therm}} = (3/4nc)\int_{E_T}^{E_0} dE/[E\sigma(E)]. \quad (31)$$

From a comparison with experimental data for the electron thermalization time in liquid and solid Ar, Kr, and Xe one can conclude that the estimates obtained here are in satisfactory agreement with published experimental data.[7,8]

[1] A. Mozumber, J. Chem. Phys. **72**, 1657 (1980).
[2] A. Mozumber, J. Chem. Phys. **72**, 6289 (1980).
[3] A. Mozumber, J. Chem. Phys. **74**, 6911 (1981).
[4] K. D. Knierim, M. Waldman, and E. A. Mason, J. Chem. Phys. **77**, 943 (1982).
[5] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* [in Russian] (Nauka, Moscow, 1979).
[6] A. I. Akhiezer and S. V. Peletminskiĭ, *Methods of Statistical Physics* [in Russian] (Nauka, Moscow, 1977).
[7] U. Sawada, Phys. Rev. B **25**, 3434 (1982).
[8] S. Kubota, Phys. Rev. B **20**, 3486 (1978).

Translated by J. R. Anderson

Use of a generalized method of eigenmodes in the theory of lasers

V. Yu. Petrun'kin and B. V. L'vov

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

An analysis is made of the application a generalized method of eigenmodes, which is used in steady-state diffraction problems to the theory of lasers. In the study of open laser cavities the use of this method makes it possible to avoid difficulties associated with the exponential growth of the field at infinity. The generalized method developed previously for solving steady-state diffraction problems is extended to the case of non-steady-state narrow-band processes to construct in the ordinary way the equations describing the operation of the laser. The only difference is that the field is expanded in the eigenfunctions that are orthogonal in the region occupied by the active medium. This eliminates the difficulties associated with the behavior of the mode field at infinity. In addition, in a number of cases the generalized eigenfunctions (modes) correspond considerably better to the field distribution in the operating laser. © *1997 American Institute of Physics.* [S1063-7842(97)01006-4]

In the theory of lasers the fields are expanded in the eigenmodes.[1] The application of this method to lasers with open cavities encounters certain difficulties including the fact that the field of the eigenmodes of an open system increase exponentially at infinity. In the present work it is shown that this difficulty can be overcome by using the generalized method of eigenmodes.[2] The direct use of this method is difficult because it has been developed for the solution of steady-state diffraction problems, while the theory of lasers is formulated in the form of differential equations. It is necessary to go from the steady-state problem to the non-steady-state problem. The essence of the generalized method can be elucidated with the example of the scalar wave equation. We consider two regions v_1 and v_2. The region v_1 is external enclosed by a sphere whose radius tends to infinity, and the region v_2 is inside the sphere filled with a medium with a permittivity ε. In region v_1 the field satisfies the equation

$$\nabla^2 u^{(1)} + k^2 \varepsilon u^{(1)} = f, \quad k^2 = \frac{\omega^2}{c^2}, \tag{1}$$

and in region v_2 the equation

$$\nabla^2 u^{(2)} + k^2 u^{(2)} = 0. \tag{2}$$

The boundary conditions

$$(u^{(1)} - u^{(2)})|_{s_1} = 0; \quad \left(\frac{\partial u^{(1)}}{\partial N} - \frac{\partial u^{(2)}}{\partial N} \right)_{s_1} = 0, \tag{3}$$

are satisfied on the interface s_1, where N is the normal to the surface.

The radiation conditions are satisfied on s_2. The eigenfunctions in which the required fields are expanded are found from a solution of the following homogeneous auxiliary problem

$$\nabla^2 u_n^{(1)} + k^2 \varepsilon_n u_n^{(1)} = 0 \quad \text{within } v_1,$$

$$\nabla^2 u_n^{(2)} + k^2 u_n^{(2)} = 0 \quad \text{within } v_2 \tag{4}$$

with the same boundary conditions as before. Here the eigenvalues are ε_n (a parameter of the medium), and $k = \omega/c$ is given by the frequency of the external forces. As shown in Ref. 2, the functions u_n are orthogonal in the volume v_1:

$$\int_{v_1} u_n^{(1)} u_m^{(1)} dv = 0, \tag{5}$$

if $\varepsilon_n \neq \varepsilon_m$.

The solution of the initial problem (1)–(3) is given by an expansion in these functions

$$u = \sum_{(n)} b_n u_n^{(1)} \tag{6}$$

inside of v_1, where

$$b_n = \frac{1}{\varepsilon - \varepsilon_n} \frac{1}{k^2} \frac{\int_{v_1} u_n^{(1)} f dv}{\int_{v_1} (u_n^{(1)})^2 dv}. \tag{7}$$

The theory of lasers is formulated as a set of nonlinear differential equations. To apply the generalized method to laser problems one must extend this theory to the non-steady-state case. Equations (1) and (2) are Fourier-transformed wave equations

$$\nabla^2 u^{(1)}(t) - \frac{\varepsilon}{c^2} \frac{\partial^2 u^{(1)}}{\partial t^2} = f(t); \quad \nabla^2 u^{(2)}(t) - \frac{1}{c^2} \frac{\partial^2 u^{(2)}}{\partial t^2} = 0, \tag{8}$$

where

$$u(t,r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(\omega, r) e^{i\omega t} d\omega;$$

$$f(t,r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega, r) e^{i\omega t} d\omega. \tag{9}$$

The boundary conditions are the same as before. It then follows that the solution of Eqs. (8) can be obtained from the solutions of (1) and (2) by the inverse Fourier transformation

$$u(t,r) = \sum_{(n)} \frac{1}{2\pi} \int_{-\infty}^{\infty} b_n(\omega) u_n(\omega, r) e^{i\omega t} d\omega. \tag{10}$$

In laser problems $b_n(\omega)$ has an extremely narrow spectrum near the lasing frequency ω_0, and $u_n(\omega, r)$ depends only weakly on the frequency. Then, taking the slowly varying factor out of the integral sign, we can write

$$u(t,r) \cong \sum_{(n)} u_n(\omega_0, r) \frac{1}{2\pi} \int_{-\infty}^{\infty} b_n(\omega) e^{i\omega t} d\omega$$

$$= \sum_{(n)} u_n(\omega_0, r) b_n(t), \tag{11}$$

and $b_n(t)$ is obtained from the inverse Fourier transformation of expression (7)

$$b_n(\omega) = \frac{1}{\varepsilon - \varepsilon_n(\omega)} \frac{1}{k^2} \int_{v_1} u_n(\omega, r) f(\omega, r) dv, \tag{12}$$

where u_n is assumed to be normalized

$$\int_{v_1} (u_n)^2 dv = 1.$$

Let us examine the behavior of the factor $1/(k^2(\varepsilon = \varepsilon_n(\omega)))$ as a function of the frequency. In Ref. 2 it was shown that as $\omega \to 0$ this expression tends to a constant value, while for $\omega \to \infty$ it falls off as $1/k^2$. For $\varepsilon - \varepsilon_n(\bar{\omega}_n) = 0$ the expression has a singularity. The frequency $\bar{\omega}_n$ for open systems is complex and the root of the equation is unique. Let us multiply the left- and right-hand sides of Eq. (12) by $\omega - \bar{\omega}_n$

$$(\omega - \bar{\omega}_n) b_n(\omega) = \frac{\omega - \bar{\omega}_n}{\varepsilon - \varepsilon_n(\omega)} \frac{1}{k^2} \int_{v_1} u_n(\omega, r) f(\omega, r) dv \tag{13}$$

and carry out the inverse Fourier transformation. We assume that $f(t)$ is a narrow-band function with a spectrum near ω_0, while the rest of the factors depend only weakly on the frequency and can be taken out of the integral sign. In this way we obtain

$$\frac{db_n}{dt} - i\bar{\omega}_n b_n = i \frac{\omega_0 - \bar{\omega}_n}{\varepsilon - \varepsilon_n(\omega_0)} \frac{1}{k^2} \int_{v_1} u_n(\omega_0, r) f(t, r) dv. \tag{14}$$

We assume further that $f(t)$ can be represented as a complex signal $f(t) = F(t) e^{i\omega_0 t}$ and correspondingly $b_n(t) = B_n(t) e^{i\omega_0 t}$ where $F(t)$ and $B_n(t)$ are slowly varying amplitudes. Substituting these definitions into Eq. (14), we obtain an equation for the complex amplitude

$$\frac{dB_n}{dt} + i(\omega_0 - \bar{\omega}_n) B_n$$

$$= i \frac{\omega_0 - \bar{\omega}_n}{\varepsilon - \varepsilon_n(\omega_0)} \frac{1}{k^2} \int_{v_1} u_n(\omega_0, r) F(t, r) dv. \tag{15}$$

Equations (14) and (15) permit us to find the expansion coefficients for the non-steady-state problem (with the stipu-

lated limitations), and using Eq. (10), to write down the solution. For a constant amplitude, where $dB_n/dt = 0$, we obtain the previous formula (7).

All of this mathematical development for the scalar wave equation also can be carried over to the system of Maxwell's equations.[2] As before we consider two regions. Region v_1 is bounded by the surface s_1 and filled with a dielectric with a permittivity ε. The second region, region v_2 is filled with free space. It is required to find the electromagnetic field created by a specified source $\mathbf{j}$. Inside the regions the field satisfies the system of Maxwell's equations

$$\nabla \times \mathbf{H}^{(1)} - ik\varepsilon \mathbf{E}^{(1)} = \frac{4\pi}{c} \mathbf{j},$$

$$\nabla \times \mathbf{E}^{(1)} + ik\mathbf{H}^{(1)} = 0 \quad \text{within} \ v_1;$$

$$\nabla \times \mathbf{H}^{(2)} - ik\mathbf{E}^{(2)} = 0,$$

$$\nabla \times \mathbf{E}^{(2)} + ik\mathbf{H}^{(2)} = 0 \quad \text{within} \ v_2. \tag{16}$$

At the interface the tangential components of the fields must be continuous

$$(\mathbf{E}_\tau^{(1)} - \mathbf{E}_\tau^{(2)})|_{s_1} = 0, \quad (\mathbf{H}_\tau^{(1)} - \mathbf{H}_\tau^{(2)})|_{s_1} = 0. \tag{17}$$

The radiation conditions must hold at infinity. If there exists any surface, then on that surface null or impedance conditions must be satisfied. The eigenfunctions $\mathbf{e}_n$ and $\mathbf{h}_n$ are obtained from the homogeneous system of equations

$$\nabla \times \mathbf{h}_n^{(1)} - ik\varepsilon_n \mathbf{e}_n^{(1)} = 0, \quad \nabla \times \mathbf{e}_n^{(1)} + ik\mathbf{h}_n^{(1)} = 0 \quad \text{within} \ v_1; \tag{18}$$

$$\nabla \times \mathbf{h}_n^{(2)} - ik\mathbf{e}_n^{(2)} = 0, \quad \nabla \times \mathbf{e}_n^{(2)} + ik\mathbf{h}_n^{(2)} = 0 \quad \text{within} \ v_2. \tag{19}$$

The boundary conditions on the interface are the same as previously. The parameter ε_n is an eigenvalue of the problem. The solution of the system of Eqs. (18) and (19) must give a set of eigenvalues and eigenfunctions. As shown in Ref. 2, the functions $\mathbf{e}_n$ and $\mathbf{h}_n$ are orthogonal in the volume v_1

$$\int_{v_1} (\mathbf{e}_n \cdot \mathbf{e}_m) dv = 0, \quad \int_{v_1} (\mathbf{h}_n \cdot \mathbf{h}_m) dv = 0, \quad \text{if} \ \varepsilon_n \neq \varepsilon_m. \tag{20}$$

The solution of the problem gives the equality
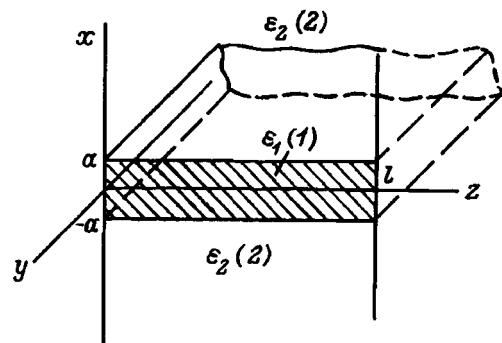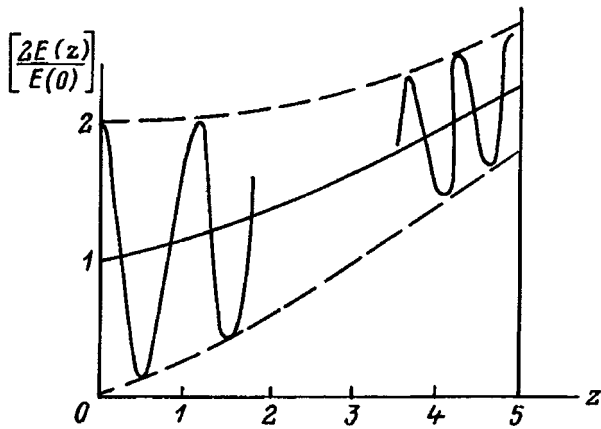

FIG. 1.

FIG. 2.

FIG. 3.

$$E = \sum_n b_n \mathbf{e}_n, \qquad b_n = \frac{4\pi}{ikc} \frac{1}{\varepsilon - \varepsilon_n} \frac{\int_{v_1} (\mathbf{j} \cdot \mathbf{e}_n) dv}{\int_{v_1} (\mathbf{e}_n)^2 dv}. \qquad (21)$$

If all this reasoning and the calculations we have done for the scalar wave equation are repeated we will obtain analogous results, which in the present case will have the form

$$\mathbf{E}(t,r) = \sum_{(n)} \mathbf{e}_n^{(n)}(\omega_0, r) B_n(t) e^{i\omega_0 t},$$

$$\frac{dB_n}{dt} + i(\omega_0 - \bar\omega_n) B_n = \frac{\omega_0 - \bar\omega_n}{\varepsilon - \varepsilon_n(\omega_0)} \frac{4\pi}{kc} \int_{v_1} (\mathbf{J}(t,r) \mathbf{e}_n^{(n)}$$

$$\times (\omega_0, r) dv. \qquad (22)$$

Here the function $\mathbf{e}_n$ is normalized: $\int_{v_1} (\mathbf{e}_n)^2 dv = 1$.

Using Eq. (22) we can in the usual way formulate a system of equations describing the operation of the laser. The only difference is that the expansion is in functions orthogonal in the region occupied by the active medium. In order that Eq. (22) describe the operation of a laser it is necessary to express $\mathbf{J}(t,r)$ in the usual way in terms of the parameters of the active medium and the field in the cavity.

The application of the generalized method to the theory of lasers is in some sense more natural, since the effect of the active medium on the field distribution of the mode is taken into account in the linear approximation. In some cases the effect of the gain on the containment of the field in the active medium may be decisive. It is also more natural to analyze unstable cavities taking into account the amplifying medium in them. The application of this method for solving various laser problems will be published later. For the present we will only use an example that shows the large effect of the active medium, and present the form of the main eigenfunctions of the generalized method for the problem illustrated in Fig. 1. The cavity is a planar waveguide clamped between two parallel mirrors. The left mirror is the low-transmission
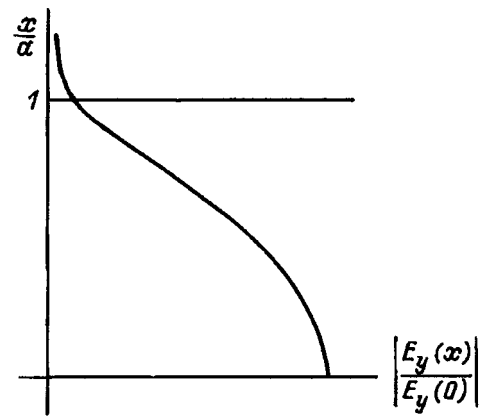
one, and the right one is the impedance mirror, providing a reflection coefficient Γ. The problem is a planar one, i.e., the field does not depend on y. The eigenfunctions for this system can be found by using the results published in Ref. 3, which considered a planar waveguide with a complex permittivity. The calculations will be carried out for the following cavity parameters: $a = 0.1$ cm, $l = 5$ cm, $\varepsilon_2 = 1$, $|\Gamma| = 0.2$, ε_1 is variable, and $\lambda_0 = 10^{-4}$ cm.

In Figures 2 and 3 we show the results of a calculation for the most illustrative case: $\varepsilon_1 \to \varepsilon_2 = 1$, where the field concentration in the active medium is determined only by the amplification of the medium. To be specific, we calculated the case of TE waves. The case of TM waves can be analyzed in a similar way. In Fig. 2 we show the dependence of the field strength on the longitudinal z coordinate. It can be seen that the average field strength increases in the direction of the exit mirror, and the oscillations due to interference between the incident and reflected waves decrease away from the low-transmission mirror. The oscillation period has been magnified for clarity. Figure 3 shows the field distribution in the transverse direction (the x coordinate). One can see that despite the absence of a discontinuity in the permittivity the field is concentrated in the active medium (because the eigenvalue is complex: $\bar\varepsilon_n = 1 + i\varepsilon_n''$). In our example, $\varepsilon'' = 0.5 \times 10^{-5}$. From the figure it is evident that in some cases the eigenfunctions of the generalized method correspond much better to the field distribution in an operating laser than is the case for the conventional method. In our example, the frequency $\bar\omega_n$ entering into Eq. (22) is $\bar\omega_n = 2\pi(3 \times 10^{14} + i0.5 \times 10^9)$ s^{-1}.

[1] W. Lamb, *Quantum Optics and Quantum Radiophysics. Theory of Optical Lasers* [Russian translation] (Mir, Moscow, 1966).
[2] N. N. Voĭtovich, B. É, Katsenelenbaum, and A. N. Sivov, *Generalized method of Eigenmodes in the Theory of Diffraction* (Nauka, Moscow, 1977).
[3] N. S. Karany and J. J. Burk, *Optical Waveguides*, (Academic Press, New York, 1972).

Translated by J. R. Anderson

Broadband acoustooptical filter

I. I. Zubrinov, V. K. Sapozhnikov, and D. V. Sheloput

Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

This paper reports an investigation of the pass band of a noncollinear acoustooptical filter based on TeO$_2$ with spectral drift compensation, implemented in a prototype of a video frequency spectrometer. The tunable acoustooptical filter is a new kind of solid-state device with electrically controlled spectral characteristics, and has the important property that it can filter light beams with a wide angular aperture.[1] In uniaxial crystals the wide-aperture interaction geometry is obtained if the plane of the acoustooptical interaction passes through the optical axis and the wave vector of the ultrasound makes an angle $\alpha \neq 0$ with the perpendicular to the optical axis.[2] This geometry is obtained when the tangents to the wave vector surfaces of the incident and diffracted light on the wave vector diagram are parallel (Fig. 1a). This indicates that for a given ultrasound frequency and light wavelength the condition of Bragg synchronism is automatically satisfied in some range of angles of incidence of the optical beam on the ultrasound, i.e., Bragg diffraction becomes insensitive to the angle of incidence of the light.

The parallelness of the tangents to the wave vector surfaces also means that the group velocity vector of the incident and diffracted light are collinear, and consequently two beams with orthogonal polarizations are not separated before the exit from the crystal.

The diagram of the wide-aperture tunable acoustooptical filter based on a paratelluride single crystal is shown in Fig. 1b. The entering optical radiation passes through the polarizer *1* and impinges on the single crystal *2* perpendicular to its entrance face.

When the electrical oscillations are fed to the piezoelectric transducer *3* ultrasonic traveling waves *4* are excited and then absorbed in the absorber *5*. The analyzer *6* suppresses the undiffracted radiation.

The spectral transmission characteristic of the tunable acoustooptical filter at a single acoustic frequency is described by the relations[2,3]

$$T(\lambda) = T_0 \sin^2[L(\lambda_0 - \lambda)/\lambda_0^2], \qquad (1)$$

where L is the length of the acoustooptical interaction, T_0 is the transmission of the filter at the wavelength λ_0, equal to

$$T_0 = T_n \sin^2(\pi^2 M_2 P_a L^2/2\lambda_0)^{1/2}, \qquad (2)$$

T_n is the total transmission of the polarizer and the analyzer, P_a is the power density of the acoustic oscillations, and M_2 is the acoustooptical quality factor of the crystal.

The total wavelength pass band at the 0.5 level is given by[2,3]

$$\Delta\lambda = 1.8\pi\lambda_0^2/bL \, \sin^2\Theta, \qquad (3)$$

where b is the dispersion constant.

In the first approximation it can be assumed that the angular aperture of the filter is inversely proportional to its wavelength resolution

$$\Delta\Theta \sim n^2(\Delta\lambda/\lambda). \qquad (4)$$

For this acoustooptical interaction geometry the radiation wavelength λ at which the diffraction intensity is a maximum, the frequency f, and the angle of incidence Θ of the radiation on the ultrasonic grating are related by[2]

$$\lambda = (v/f)[n_i \sin\Theta - (n_0^2 - n_i^2 \cos^2\Theta)^{1/2}], \qquad (5)$$

where $n_i = n_0 n_e[n_e^2 \cos^2(\Theta + \alpha) + n_0^2 \sin^2(\Theta + \alpha)]^{-1/2}$; and n_0 and n_e are the refractive indices for the ordinary and extraordinary rays, v is the velocity of slow ultrasonic shear waves in the (110) plane, which is given by

$$v = \left[\left(\frac{c_{11} - c_{12}}{2} \cos^2\alpha + c_{44} \sin^2\alpha\right) \Big/ \rho\right]^{1/2}, \qquad (6)$$

where ρ and c_{ij} are the density and elastic constants of TeO$_2$.

In actual fact, because of the dispersion of the refractive index as a function of λ and f, there is some deviation from inverse proportionality. The tuning curve on a logarithmic scale for the entire spectral range of transparency of TeO$_2$ and for $\alpha = 10°$ is shown in Fig. 2.

We can use Eq. (5) to calculate the frequency dependence of the angle of incidence on the acoustooptical wave front (the Bragg angle). In the calculations we used $n_0 = 2.26$, $n_e = 2.41$, $v = 7.1 \times 10^3$ m/s for $\alpha = 9.5°$, 10°, and 10.5°. These curves are shown in Fig. 3 (curves *1–3*). The extremum of the function corresponds to the position on the vector diagram for which the tangents to the wave surfaces of the optical beams, passed through the ends of the vectors $\mathbf{k}_i$ and $\mathbf{k}_d$, are parallel and satisfy the condition $\partial f/\partial\Theta = 0$. Thus, with a small deviation of the ray from the angle of incidence, the length of the vector $\mathbf{K}$ (Fig. 1a), and consequently the ultrasound frequency for which the condition of Bragg synchronism is satisfied, are not changed. The dots on the curves in Fig. 3 show the experimentally measured frequency dependence of the Bragg angle. As one can see, expression (5) describes the experimental results with good accuracy.

In this investigation we have studied experimentally the pass band of a tunable acoustooptical filter in a model of a video frequency spectrometer. The design of the experimen-
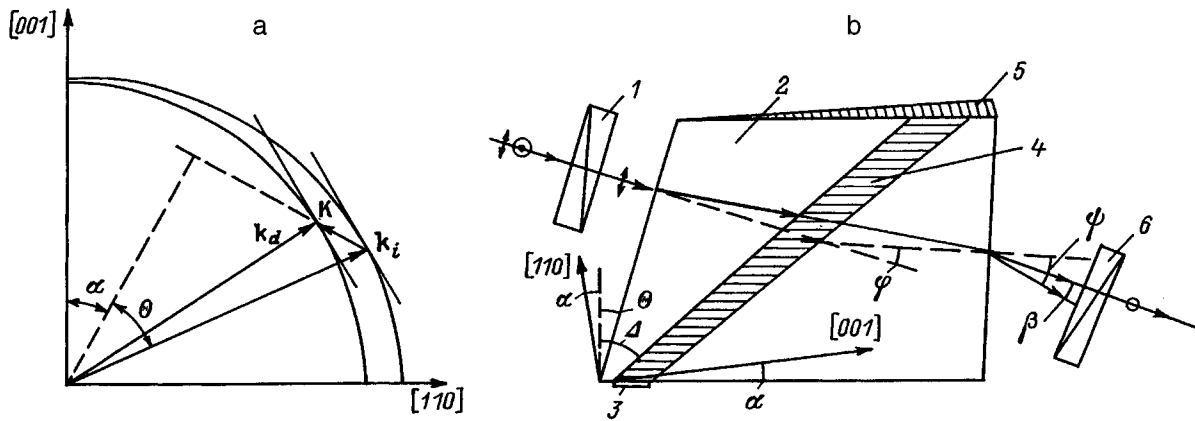
FIG. 1. a — Wave vector diagram for the incident and diffracted light for a given frequency f of the sound and wavelength λ of the light, when the Bragg synchronization condition is satisfied in some interval of divergence of the optical beam; b — diagram of the tunable acoustooptical filter based on TeO$_2$.

tal apparatus is shown in Fig. 4. The source of a continuous spectrum in the visible and near infrared was a small incandescent lamp *1*. The radiation was collimated by a small objective *2*, and with objective *3* produced the required angular aperture. The polarization of the incident radiation was set by the plane of transmission of the polarizer *4*. The incident light beam propagated in the filter as an extraordinary ray. The electrical control signal was fed from the generator *7* to the piezoelectric transducer of the filter *5* with a frequency that could be monitored by a frequency meter *8*. An analyzer *6*, crossed with the polarizer was placed at the exit from the filter. A light pipe *9* brought the diffracted radiation to a monochromator *10* and the radiation was recorded with a photodetector *11*, a tuned amplifier *12*, and an oscilloscope *13*.

To measure the pass band of the filter a light pipe *9* was moved along the cross section of the beam in the measurement plane. The diameter of the diffraction spot was 25 mm and the diameter of the light pipe was 3 mm. By moving the light pipe over the cross section of the diffraction spot we could measure the pass band for various angles of incidence of the radiation within the angular aperture.

The experimental samples of the tunable acoustooptical filter were prepared from TeO$_2$ single crystals in which the plane of the acoustooptical interaction was the (110) plane, as shown in Fig. 1b. The wave vector of the ultrasound in the crystal made an angle α with the (110) plane, with the angle α for various sample set at 9.5° or 10° with an accuracy of $\pm 5'$. The linear dimensions of the optical and acoustic waveguide were 10 mm in the (110) direction and 14 mm in the perpendicular direction. The length of the optical and acoustic waveguide was selected with allowance for the drift of the acoustic energy by an angle of 56° relative to the wave vector of the ultrasound, and was 25 mm. The input optical face made an angle of 23.5° with the (001) plane. A piezoelectric transducer made of X-cut lithium niobate was cold-welded to the acoustic waveguide by indium compression in vacuum The dimension of the transducer was 4 mm along the direction of propagate of the light and 8 mm in the [110] direction.

The electrical impedance of the output circuit of the rf generator was matched to the impedance of the transducer by sectioning. The frequency band was 60 MHz (60–120 MHz) and 80 MHz (90–170 MHz). With an electrical control signal of $\simeq 0.7$ W the diffraction efficiency was measured at a
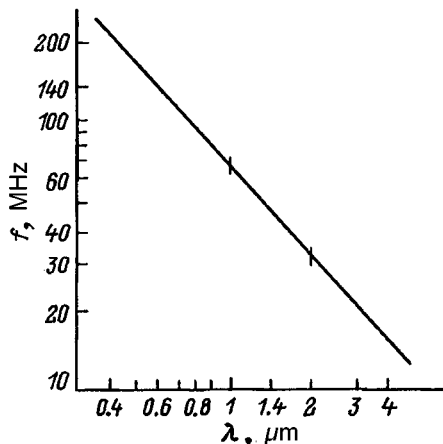

FIG. 2. Dependence of the wavelength λ of light on the frequency f of the sound (on a logarithmic plot) for TeO$_2$ at $\alpha = 10°$.
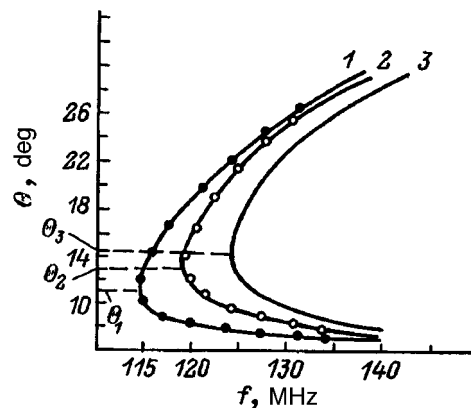

FIG. 3. Frequency dependence of the Bragg angle for λ =3D 0.63 μm. α =3D 9.5° (*1*), 10° (*2*), and 10.5° (*3*). The points show the experimental data.

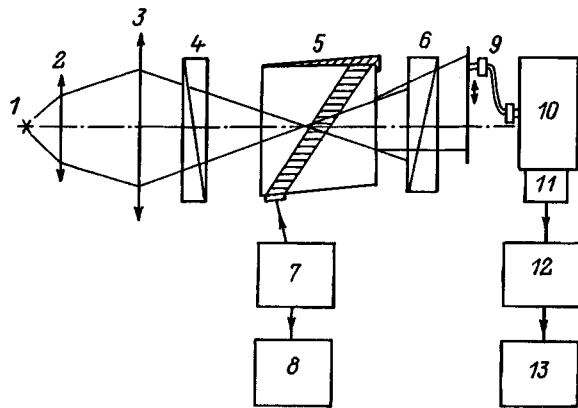FIG. 4. Model of spectrometer for studying the profile of the pass band of the tunable acoustooptical filter.

FIG. 5. Calculated broadening of the pass band of the tunable acoustooptical filter for an angular aperture of 10° (*1*) and 8° (*2*) for conditions such that the ultrasonic wave in the interaction region is a plane wave. The points indicate the experimentally determined dependence of the pass band for a filter with an angular aperture of 10°.

radiation wavelength of 633 nm, and was found to be 80% and 90%. The total pass band of the filter was determined from the frequency width of Bragg synchronism. Because of the angular aperture of the light $\Delta\Theta$ and the divergence $\Delta\alpha$ of the ultrasound the pass band becomes wider.

The frequency width of Bragg synchronism is mainly governed by the length of the acoustooptical interaction. To calculate it one must take into account that the interaction length differs from the actual dimensions of the piezoelectric transducer by an amount $l[1-\cos\Delta/\cos(\Delta-\Theta)]$, where l is the dimension of the transducer in the plane of interaction, Δ is the drift angle equal to

$$\Delta = \tan^{-1}\left\{\frac{1}{v}\frac{\partial v}{\partial \alpha}\right\},$$

and from Eq. (6)

$$\Delta = \tan^{-1}\{[c_{44}-(c_{11}-c_{12})/2]\sin\alpha$$
$$\times\cos\alpha/[(c_{11}-c_{12})/2]\cos^2\alpha + c_{44}\sin^2\alpha\}. \quad (7)$$

In the calculation of widening of the pass band of the tunable acoustooptical filter with increasing angular aperture, we assume that the ultrasonic beam is a plane wave in the interaction plane. For the calculation we use the method described in Ref. 4. Figure 5 shows the calculated widening of the pass band of the acoustooptical filter for an angular aperture of 10° (curve *1*) and 8° (curve *2*). The points indicate the experimentally measured pass band for the acoustooptical filter with an angular aperture of 10°, measured from the half-width of the diffraction efficiency. A comparison of the experimental results and the calculated curve *1* shows that the broadening of the pass band due to the angular aperture is substantial.

The quantitative discrepancy obviously comes about because the acoustic wave is not really a plane wave, and the large acoustic anisotropy in the plane of interaction results in an increase in the spatial divergence of the ultrasound. To narrow the pass band of the filter for a given angular aperture it is necessary both to increase the acoustooptical interaction length and increase the angle between the wave vector of the sound and the perpendicular to the optical axis. However the
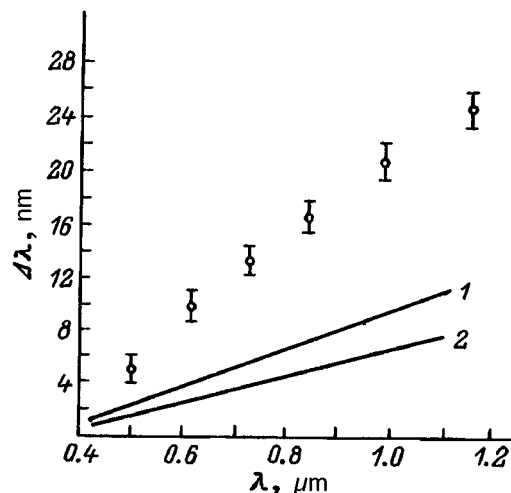
latter tactic results in a decrease in the acoustooptical figure of merit of this optical-acoustic waveguide cut, which reduces the efficiency of the filter.

The dispersion of the refractive index of paratelluride causes a certain mutual transverse displacement of the images in different wavelength intervals. In the operation of the filter in a wide spectral region this may cause a ''creep'' of the beam outside the sensing area of the photodetector.

It is possible to compensate for this displacement by tilting the exit face by an angle φ relative to the entrance face.[5]
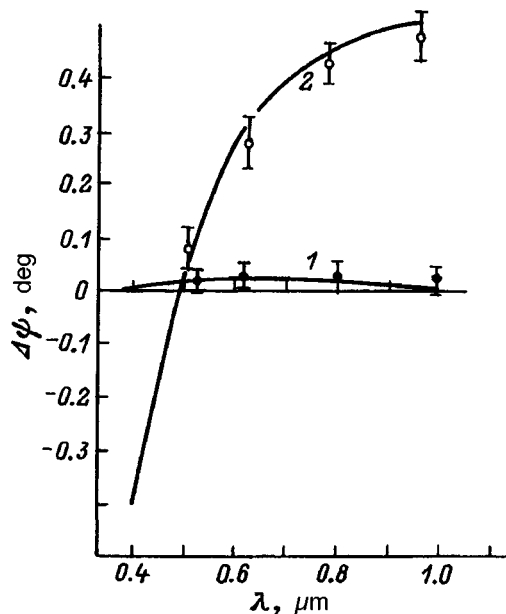

FIG. 6. Calculated wavelength dependence of the variation of the antular coordinate $\Delta\psi$ for a tunable acoustooptical filter with drift compensation in the range 0.4–1 μm (*1*) and the analogous filter without drift compensation (*2*). The points indicate the experimental data.

Upon leaving the crystal the diffracted beam is refracted by an angle to the normal to the exit face of the filter. The condition that the angle ψ be invariant in the wavelength interval $\lambda_1-\lambda_2$ is

$$\sin \psi = n_0 \sin \psi_k = n_0 \sin$$
$$\times [\varphi - \Theta - \alpha + \tan^{-1}(\eta^2 \tan(\Theta + \alpha))] = \text{const}, \quad (8)$$

where $\eta = n_0(\lambda)/n_e(\lambda)$.

Assuming that the angle $\Theta + \alpha$ does not depend on λ and differentiating both sides of Eq. (8) with respect to λ, we find $\varphi = \delta + \psi_k$, where δ is the angle of inclination of the exit face for the standard case where it is orthogonal to the diffracted beam, and ψ_k is the additional angle of inclination for compensating the spectral drift of the diffraction angle.

According to Ref. 5,

$$\psi_k = -\tan^{-1}\left\{ \frac{2n_0 \tan(\Theta + \alpha) \eta d\eta/d\lambda}{dn_0/d\lambda[1 + \eta^4 \tan^2(\Theta + \alpha)]} \right\}. \quad (9)$$

Equation (9) shows that in the case of normal dispersion, as occurs for TeO$_2$, $dn_0/d\lambda < 0$, $d\eta/d\lambda > 0$, and $\psi_k > 0$. In Fig. 6 we display the calculated wavelength dependence of the change in the angular coordinate for the tunable acoustooptical filter with compensation for the spectral drift in the interval $0.4-1$ μm ($\psi(0.4 \ \mu\text{m}) = \psi(1 \ \mu\text{m})$) (curve 1) and the standard filter similar to the tunable acoustooptical filter, where the exit face is perpendicular to the diffracted wave at the wavelength 0.5 μm (curve 2).

The points on the graph show the experimentally measured spectral drift of the diffraction angle for the two kinds of filter. As is evident from a comparison of the data, the experimental results are in good agreement with the calculations.

CONCLUSIONS

Wide-aperture tunable acoustooptical filters based on TeO$_2$ combine such important parameters as a high diffraction efficiency with a comparatively low power of the controlling electrical signals, a fairly high wavelength resolution, and simplicity of design, and they provide high reliability in operation and low cost.

The relatively low chromatism of these tunable acoustooptical filters permit their use in spectrometers in the visible and infrared range for spectral image filters.

Having a large angular aperture and a wide pass band compared to monochromators based on quartz, these filters have several hundred times higher luminosity, which increases the sensitivity of the spectrometer and reduces the variance in the measurements.

[1] V. M. Epikhin, F. L. Vizen, N. V. Nikitin, and Yu. K. Kalinnikov, Zh. Tekh. Fiz **52**, 2405 (1982) [Sov. Phys. Tech. Phys. **27**, 1482 (1982)].
[2] I. C. Chang, Opt. Eng. **20**, 806 (1981).
[3] L. N. Magdich, Izv. Akad. Nauk SSSR Ser. Fiz. **44**, 1683 (1980).
[4] V. B. Voloshinov et al. Zh. Prikl. Spektrosk. **52**, 284 (1990).
[5] V. M. Epikhin and Yu. K. Kalinnikov, Zh. Tekh. Fiz **59**(2), 160 (1989) [Sov. Phys. Tech. Phys. **34**, 227 (1989)].

The thermodynamics of energy-storing photoprocesess

N. D. Gudkov

Institute of Soil Science and Photosynthesis, Russian Academy of Sciences, 142292 Pushchino, Moscow District, Russia

A solution is found for the most general formulation of the problem of the thermodynamic limitations on the rate of increase of the free energy of a closed photochemical system interacting with a heat bath and a radiation field. A reaction of the form $A_1^* + A_2 \rightarrow A_3 + A_4$ is used to compare the thermodynamically maximum energy yield of the process of storing radiant energy with the ''kinetic'' limit of the same quantity. © *1997 American Institute of Physics.*
[S1063-7842(97)01206-3]

INTRODUCTION

Thermodynamics imposes definite limits on the efficiency of converting radiant energy into ''chemical'' energy: one can show (see, e.g., Ref. 1) that the net flux of free energy of a material through the surface of a reacting volume, involving the introduction of the initial reagents in the volume and the removal of the products of an endoenergetic photochemical reaction (see Fig. 1), i.e., the gain (per unit time) of free chemical energy can never exceed the difference $\Pi_r - T\Sigma_r$, where Π_r and Σ_r are the net fluxes of energy and entropy, respectively, of the radiation obtained by the ''reactor'' through its surface; T is the temperature at which the reaction occurs, and which (by hypothesis) is equal to or only slightly different from the temperature of the surrounding medium.

A mature photosynthesizing leaf (sending the photosynthesis products to the other members of the plant) is a classical (and the most important) example of a ''device'' that utilizes the process. Photosynthesis, however, offers an example of another sort. If the entire plant is placed in a ''climate'' chamber that is supplied with enough reserve required for its lifetime, then we have a system (closed), which as before will function, obviously, as a device for converting light energy into chemical energy. The difference of such a converter (the ''accumulator'' type) from the previously described type (which may be called a ''generator''), is the non-steady-state nature of its operation: under illumination the state of the system changes in time and is accompanied by the increase in the free energy.

The question of thermodynamic limitations on the ''storage'' of radiant energy by this kind of system was first studied in Ref. 2 and independently in a more recent publication.[3] Both investigators solved the problem under the assumption of a steady-state radiation field in the volume of the converter, and they found that the limiting (the thermodynamically attainable) rate of increase dF/dt of the free energy of the latter is determined by the same combinations that limit the power of a photchemical reactor

$$dF/dt\Big|_{\substack{\text{steady-state} \\ \text{radiation field}}} \leq \Pi_r - T\Sigma_r = N_m. \tag{1}$$

In this paper I undertake to solve this problem in a more general formulation, assuming that the states of the ''field'' and ''material'' components are not stationary.[1]

We shall denote by E and S the internal energy and entropy of the material component of the system, which exchanges energy with the radiation field and the surrounding medium and in which the photo-endoenergetic chemical reaction takes place. The energy and the entropy of the radiation field in the interior of the system will be denoted by E_r and S_r, respectively. We also introduce the fluxes Π_r^{in} and Σ_r^{in} of energy and entropy of the radiation incident on the surface of the system, and correspondingly Π_r^{out} and Σ_r^{out} for the radiation leaving the same surface. Finally, q denotes the heat flux that the system exchanges with the surrounding medium (at a temperature T) by direct heat conduction. The balance equation for the total energy and entropy of the system is then written in the form

$$\frac{d}{dt}(E + E_r) = \Pi_r^{\text{in}} - \Pi_r^{\text{out}} - q,$$

$$\frac{d}{dt}(S + S_r) = \Sigma_r^{\text{in}} - \Sigma_r^{\text{out}} - \frac{q}{T} + \frac{d_i S_t}{dt},$$

where $S_t = S + S_r$, and as usual, $d_i S_t/dt$ denotes the rate of generation (''production'') of entropy in the system.

Eliminating the flux q from the equations, we obtain

$$\frac{dF}{dt} = \Pi_r - T\Sigma_r - \frac{dE_r}{dt} + T\frac{dS_r}{dt} + T\frac{d_i S_t}{dt}. \tag{2}$$

Here $F(= E - TS)$ is the free energy (of the material part) of the system, $\Pi_r(= \Pi_r^{\text{in}} - \Pi_r^{\text{out}})$ and $\Sigma_r(= \Sigma_r^{\text{in}} - \Sigma_r^{\text{out}})$ are the net energy and entropy fluxes into the system through its surface. In accordance with the second law of thermodynamics, the last term on the right-hand side of Eq. (2) is intrinsically nonnegative, and consequently the following inequality must necessarily hold

$$\frac{dF}{dt} \leq \Pi_r - T\Sigma_r - \frac{d}{dt}(E_r - TS_r) = N_m, \tag{3}$$

which also expresses in the most general form the limitations imposed by thermodynamics on the energy of an arbitrary photo-endoenergetic process.[2]

2. All the quantities entering into expression (3) for the limiting power N_m are functions of the state of the radiation field in the interior and on the surface of the converter, i.e.,
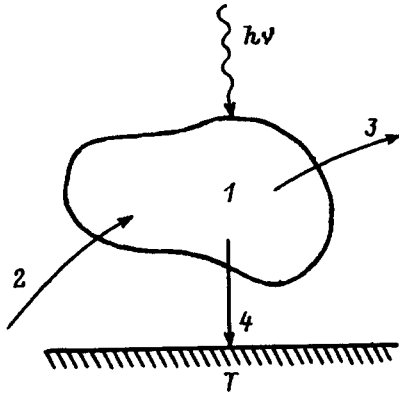
FIG. 1. Diagram of a photochemical converter of radiant energy into free chemical energy. *1* — reactor, whose state does not change with time and which is in contact with the surrounding medium (at a temperature T) and exchanges energy with the radiation field; *2* — initial reagents, *3* — products of the photochemical reaction; *4* — heat flux.

they are determined by the spectral brightness K_ν of the radiation, which must be known for all frequencies ν and directions $\boldsymbol{\omega}$ of propagation of the light ray at all points $\mathbf{r}$ and for every instant of time t. Assuming that $K_\nu(\boldsymbol{\omega},\mathbf{r},t)$ is known we find that

$$N_m = -\int d\sigma \int d\nu \int d\Omega_{\boldsymbol{\omega}} \mathbf{n}\cdot\boldsymbol{\omega}(K_\nu - TL_\nu)$$
$$- \int d\tau \int d\nu \int d\Omega_{\boldsymbol{\omega}} \frac{1}{q_\nu}\left(\frac{\partial K_\nu}{\partial t} - T\frac{\partial L_\nu}{\partial t}\right). \quad (4)$$

Here $d\sigma$ is an element of surface area (with an outer normal $\mathbf{n}$), $d\tau$ is an element of volume of the converter, $d\Omega_{\boldsymbol{\omega}}$ is an element of solid angle in the direction $\boldsymbol{\omega}$, q_ν is the speed of light of frequency ν in the medium, and $L_\nu = L_\nu(K_\nu)$ is the spectral brightness of the radiation.[3)]

In the first term in the sum (4) we use the Gauss-Ostrogradskiĭ theorem to change the integration over the surface to an integration over the volume of the converter. Combining the two volume integrals and recalling that for a constant vector $\boldsymbol{\omega}$ (independent of $\mathbf{r}$), the following is true:

$$\operatorname{div}(K_\nu - TL_\nu)\boldsymbol{\omega} = \boldsymbol{\omega}\cdot\boldsymbol{\nabla}(K_\nu - TL_\nu),$$

and we obtain as a result

$$N_m = -\int d\tau \int d\nu \int d\Omega_{\boldsymbol{\omega}}\Bigg[\boldsymbol{\omega}\cdot\boldsymbol{\nabla}(K_\nu - TL_\nu)$$
$$+ \frac{1}{q_\nu}\left(\frac{\partial K_\nu}{\partial t} - T\frac{\partial L_\nu}{\partial t}\right)\Bigg]$$

or

$$N_m = -\int d\tau \int d\nu \int d\Omega_{\boldsymbol{\omega}}\left(1 - \frac{T}{T_\nu}\right)\left(\boldsymbol{\omega}\cdot\boldsymbol{\nabla}K_\nu + \frac{1}{q_\nu}\frac{\partial K_\nu}{\partial t}\right). \quad (5)$$

Here we have used the relations

$$\boldsymbol{\nabla}L_\nu = \frac{\partial L_\nu}{\partial K_\nu}\boldsymbol{\nabla}K_\nu, \qquad \frac{\partial L_\nu}{\partial t} = \frac{\partial L_\nu}{\partial K_\nu}\frac{\partial K_\nu}{\partial t},$$

and have introduced the temperature T_ν of the light beam of frequency ν, which is defined by[4]

$$\frac{1}{T_\nu} = \frac{\partial L_\nu}{\partial K_\nu}.$$

Finally, instead of the combination of quantities in the second parentheses under the integral sign in Eq. (5) we substitute its value from the radiant energy transport equation

$$\frac{1}{q_\nu}\frac{\partial K_\nu}{\partial t} + \boldsymbol{\omega}\cdot\boldsymbol{\nabla}K_\nu = -\alpha_\nu K_\nu + \varepsilon_\nu, \quad (6)$$

where α_ν and ε_ν are the absorption and transmission coefficients, respectively, for the medium.[5]

The expression for the limiting power N_m then takes the following form

$$N_m = \int d\tau \int d\nu \int d\Omega_{\boldsymbol{\omega}}\left(1 - \frac{T}{T_\nu}\right)(\alpha_\nu K_\nu - \varepsilon_\nu) \geq \frac{dF}{dt}. \quad (7)$$

Formula (7) is the final goal of all of these calculations and of the work as a whole. This inequality solves the problem (of the thermodynamic limitation on the storage of radiant energy) in its most general formulation, at least as applied to converters of the photochemical type. It should be noted that the solution (7) contains, as one might expect, the result (1) of Refs. 2 and 3 as a special case corresponding to the condition[4)] $\partial K_\nu/\partial t = 0$. It should be noted that the extension of the result (1) to the general case in fact reduces to a change in the form of writing expression (1) for N_m (compare relations (1) and (7)) and we find the interesting circumstance (generally not obvious *a priori*) emerging from the properties of the solution: the thermodynamically maximum value of dF/dt is determined, as can be seen from relation (7) by only the ''coordinates'' of the field components of the system (i. e., the values $K_\nu(\boldsymbol{\omega},\mathbf{r})$ at a given instant of time) and does not depend at all on the ''velocities'' (i.e., the quantities that would characterize explicitly a possible non-steady state of the subsystem; compare Eq. (4), and also (1) and (3)).

It should also be mentioned that the initial relation (3) and consequently the inequality (7) follow from the laws of thermodynamics in their ''global'' formulation. Both laws are valid, as is known, even in the local form. In particular, the second law states that not only d_iS_t/dt, but also the volume density of this quantity is intrinsically nonnegative.[6] With respect to our problem, this means that if we introduce the volume density f of the quantity F and thereby replace dF/dt by the integral

$$\frac{dF}{dt} = \int d\tau \frac{df}{dt},$$

then from the inequalities of the sums on the two sides of relation (7) we correctly infer that the inequality applies to the individual terms. As a result we arrive at the following ''local'' formulation of the thermodynamic limitations on the energetics of the photchemical process

$$\frac{df}{dt} \le \int d\nu \int d\Omega_\omega \left(1 - \frac{T}{T_\nu}\right)(\alpha_\nu K_\nu - \varepsilon_\nu). \tag{8}$$

It is to be noted that the coefficients α_ν and ε_ν in these formulas generally are made up of the sum $\alpha_\nu = \alpha_{\nu 0} + \alpha_\nu'$ and $\varepsilon_\nu = \varepsilon_{\nu 0} + \varepsilon_\nu'$, where $\alpha_{\nu 0}$ and $\varepsilon_{\nu 0}$ refer to the "true" absorption and emission and α_ν' and ε_ν' come from the (possible) scattering of the radiation (without change in frequency). By virtue of the so-called "reciprocity theorem"

$$\int d\Omega_\omega (\alpha_\nu' K_\nu - \varepsilon_\nu') = 0,$$

and hence one of the integrals in expression (8) is

$$\int d\nu \int d\Omega_\omega (\alpha_\nu K_\nu - \varepsilon_\nu) = \int d\nu \int d\Omega_\omega (\alpha_{\nu 0} K_\nu - \varepsilon_{\nu 0})$$

and therefore determines the (instantaneous) net radiation power absorbed per unit volume of the reaction mixture.[5] Relation (8) can now be rewritten in the form[6]

$$\frac{df}{dt} \le \varphi_r - \int d\nu \int d\Omega_\omega \frac{T}{T_\nu}(\alpha_\nu K_\nu - \varepsilon_\nu),$$

or equivalently

$$\frac{df/dt}{\varphi_r} = \eta \le 1 - \frac{T}{T_r} = \eta_m, \tag{9}$$

where we have used the notation

$$\frac{1}{T_r} = \int d\nu \int d\Omega_\omega \frac{\alpha_\nu K_\nu - \varepsilon_\nu}{\varphi_r} T_\nu^{-1}, \tag{10}$$

and the quantity η has the meaning of a "local" energy yield (the efficiency) of the process and η_m is consequently its (the yield's) thermodynamic limit.[7]

It should be pointed out that the inequality (9) is entirely general and must hold for any arbitrary photoendoenergetic process. In order to elucidate (even though only partially) the physical content of so general a relation, it would be helpful to "test" it with a specific example of a photochemical reaction used for this purpose to obtain the upper limit of η without recourse to the thermodynamic (entropy) properties of the radiation, but based only on kinetic arguments. The results of one of the possible attempts in this direction is included in the following Appendix for the interested reader.

APPENDIX

We assume that the following processes are taking place in the reaction volume (at constant pressure p and temperature T):

$$A_1 + h\nu \underset{k_r^\downarrow}{\overset{k_r^\uparrow}{\rightleftharpoons}} A_1^* \overset{k_{nr}^\downarrow}{\rightarrow} A_1, \tag{11a}$$

$$A_1^* + A_2 \overset{k_q}{\rightarrow} A_3 + A_4, \tag{11b}$$

$$A_3 + A_4 \overset{k_b}{\rightarrow} A_1 + A_2, \tag{11c}$$

where A_1^* stands for the singlet-excited state of the reagent A_1 and the symbols above the arrows indicate the rate constants of the corresponding reactions.

We shall assume that the reactions are not accompanied by any appreciable changes in the volume of the solution, so that the change in the free energy F is the same as the change in the thermodynamic potential G, and consequently

$$\frac{df}{dt} = \sum_i \mu_i \dot{n}_i + \mu_1^* \dot{n}_1^*. \tag{12}$$

Here μ_i is the chemical potential n_i is the concentration of the ith component of the reaction, μ_1^* and n_1^* refer to the collection of particles A_1^*, and the dot above the variables denotes, as usual, differentiation with respect to the time. We transform the right-hand side of Eq. (12) using the relations that follow from Eqs. (11a)—(11b),

$$\dot{n}_2 = -\dot{n}_3, \quad \dot{n}_3 = \dot{n}_4, \quad \dot{n}_3 = -(\dot{n}_1 + \dot{n}_1^*). \tag{13}$$

We obtain as a result

$$\frac{df}{dt} = \dot{n}_3 \Delta G + \dot{n}_1^* \Delta \mu, \tag{14}$$

where

$$\Delta G = \mu_3 + \mu_4 - \mu_1 - \mu_2$$

is the increase in the isobaric potential in the (endoenergetic) reaction $A_1 + A_2 \rightarrow A_3 + A_4$, and the notation $\Delta \mu = \mu_1^* - \mu_1$ has been introduced.

We assume that the reaction (11b) proceeds spontaneously, and consequently

$$\dot{n}_3 \ge 0, \tag{15}$$

while the change $\widetilde{\Delta}G$ of the thermodynamic potential in this reaction is negative

$$\widetilde{\Delta}G = \mu_3 + \mu_4 - \mu_1^* - \mu_2 = \Delta G - \Delta \mu \le 0. \tag{16}$$

If we take into account relations (15) and (16), we see that equation (14) means that

$$df/dt \le (\dot{n}_3 + \dot{n}_1^*)\Delta \mu$$

or, as follows from Eq. (13),

$$df/dt \le -\dot{n}_1 \Delta \mu. \tag{17}$$

In accordance with Eqs. (11) we then find

$$\dot{n}_1 = (k_r^\downarrow + k_{nr}^\downarrow)n_1^* - k_r^\uparrow n_1 + k_b n_3 n_4, \tag{18}$$

from which we obtain the upper limit on the first of the factors on the right-hand side of relation (17)

$$-\dot{n}_1 \le k_r^\uparrow n_1 - k_r^\downarrow n_1^*. \tag{19}$$

Confining the analysis to the case of a weak solution, we can now estimate the difference $\Delta \mu$ of the chemical potentials of the A_1 molecules in the singlet-excited state and the ground state. In this approximation[8]

$$\Delta \mu = \Delta \mu^0(p, T) + kT \ln(n_1^*/n_1),$$

where $\Delta\mu^0$ is the standard value of $\Delta\mu$, and k is Boltzmann's constant.

When $\dot{n}_1 \leqslant 0$ ((i. e., when the products of the reaction (11b) accumulate, which is assumed)), it follows from the kinetic equation (18) that

$$n_1^*/n_1 \leqslant k_r^\uparrow/k_r^\downarrow$$

and hence

$$\Delta\mu \leqslant \Delta\mu^0 + kT \, \ln(k_r^\uparrow/k_r^\downarrow). \tag{20}$$

The subsequent calculations are considerably simplified if the radiation field is assumed to be isotropic at any point of the medium (we thereby assume that the is initiated by the diffusion of light) and if it is also assumed that the absorption spectrum of the molecule A_1 consists of a single narrow line near some frequency ν. With these assumptions the following relation holds[8] (Refs. 7 and 9).

$$\frac{k_r^\downarrow}{k_r^\uparrow} = \frac{1+n_\nu}{n_\nu} \exp[(\Delta\mu^0 - h\nu)/kT]. \tag{21}$$

Here $\Delta\mu^0$ has the same meaning as in formula (20), and n_ν is the number of quanta of the field oscillators

$$n_\nu = q_\nu^2 K_\nu/2h\nu^3,$$

where h is Planck's constant; we recall that q_ν is the velocity of light and K_ν is the spectral brightness of the radiation. Using Eq. (21), we obtain from relation (20)

$$\Delta\mu \leqslant h\nu\left[1 - \frac{kT}{n\nu}\ln\left(1 + \frac{1}{n_\nu}\right)\right]. \tag{22}$$

From inequalities (17), (19), and (22), we obtain the final result for the ''kinetic'' limit $\widetilde{\eta}_m$ of the energy yield of the photochemical reaction under study

$$\eta = \frac{df/dt}{\varphi_r} \leqslant 1 - \frac{T}{T_\nu} = \widetilde{\eta}_m, \tag{23}$$

where we have used the standard notation φ_r for the net radiation power absorbed per unit volume of the medium

$$\varphi_r = h\nu(k_r^\uparrow n_1 - k_r^\downarrow n_1^*)$$

and have taken into account that the combination $h\nu/k \, \ln(1+n_\nu^{-1})$ determines the temperature T_ν of the radiation of frequency ν.[4]

Let us now compare the result (23) with the thermodynamic limit η_m of the quantity η (formula (9)), which, we recall, we obtained from the most general arguments without inquiring into the detailed mechanism of the processes going on in the system. If the field is isotropic then scattering can be disregarded and we then put in formula (10) $\alpha_\nu = \alpha_{\nu 0}$ and $\varepsilon_\nu = \varepsilon_{\nu 0}$; and in addition the integration over $d\Omega_\omega$ reduces to multiplying the integrand by a factor of $4/\pi$, so that

$$\frac{1}{T_r} = 4\pi \int d\nu \, \frac{\alpha_{\nu 0} K_\nu - \varepsilon_{\nu 0}}{\varphi_r} T_\nu^{-1}.$$

Recalling the (assumed) narrow-band nature of the absorption spectrum of the particles A_1 we then obtain

$$\frac{1}{T_r} = \frac{4\pi\Delta\nu(\alpha_{\nu 0}K_\nu - \varepsilon_{\nu 0})}{\varphi_r} T_\nu^{-1},$$

or since $4\pi\Delta\nu(\alpha_{\nu 0}K_\nu - \varepsilon_{\nu 0}) = \varphi_r$,

$$T_r = T_\nu. \tag{24}$$

Therefore the value of η_m determined by formula (9) coincides exactly with the kinetic limit $\widetilde{\eta}_m$ for the energy yield from the equality (23).

[1] Strictly speaking, the assumption that the radiation field is stationary in the volume of the energy storage system does not enter explicitly in these investigations, but, as will be seen later on (see footnote 2), the result obtained in Refs. 2 and 3 corresponds to the case which is almost surely realized when the material absorbing the light is not consumed during the reaction, but acts as a photosensitizer, as, for instance, in photosynthesis. However, in the general case the state of the radiation field, in the medium varies, along with a change in the optical properties during the photochemical reaction. Of course, the field may vary with time because of changes in the flux of radiant energy to be converted.

[2] Expression (1) for the thermodynamic limit of the ''useful'' power dF/dt of an ''accumulator''type of converter, derived in Ref. 2 and 3 is implied by Eq. (3) if we set $dE_r/dt = dS_r/dt = 0$; i.e., it is true, as already noted, for a stationary radiation field in the volume of the energy-storing system.

[3] The quantity L_ν is derived by exactly the same means as K_ν, that is, by the equality $dS_\nu = -L_\nu \mathbf{n} \cdot \boldsymbol{\omega} d\sigma d\nu d\Omega_\omega dt$, where dS_ν is the entropy of the radiation whose energy includes a frequency interval $(\nu, \nu+d\nu)$, and which in a time dt is transported through an area $d\sigma$ along the normal and within an element of solid angle $d\Omega_\omega$ in the direction of the unit vector ω (Ref. 4).

[4] For the steady-state radiation field Eq. (6) for the transport of radiant energy gives $\alpha_\nu K_\nu - \varepsilon_\nu = -\boldsymbol{\omega} \cdot \nabla K_\nu$ and, correspondingly Eq. (7) becomes

$$N_m = -\int d\tau \int d\nu \int d\Omega_\omega\left(1 - \frac{T}{T_\nu}\right)\boldsymbol{\omega} \cdot \nabla K_\nu,$$

from which, after going over to a surface integral, Eq. (1) follows.

[5] The term ''net'' means in this context that we are accounting for radiant energy that upon absorption goes into any other forms but not into radiant energy again (as through luminescence).

[6] In the integral $\int d\Omega_\omega(T/T_\nu)(\alpha_\nu K_\nu - \varepsilon_\nu)$ the replacements $\alpha_\nu \to \alpha_{\nu 0}$, $\varepsilon_\nu \to \varepsilon_{\nu 0}$ are not permissible in the general case (for more information on the effects due to light scattering see, e.g., Ref. 7).

[7] Dividing both sides of inequality (7) by the quantity $\Phi_r = \int d\tau\varphi_r$, we obtain, clearly, the upper limit similar to the estimate (9) for the efficiency $\hat{\eta} = \Phi_r^{-1}dF/dt$ of the system as a whole.

[8] The condition of thermalization as a necessary condition for the validity of Eq. (21) (the derivative the ''universal relation between the absorption spectrum and the fluorescence spectrum) in condensed media (and, in particular, in solutions) is almost certainly satisfied.[7,9]

[1] N. D. Gudkov, Zh. Fiz. Khim. **67**, 635 (1993).

[2] L. N. Bell, Zh. Éksp. Teor. Fiz. **46**, 1117 (1964) [Sov. Phys. JETP **19**, 756 (1964)].

[3] M. A. Leontovich, Usp. Fiz. Nauk **114**, 555 (1974) [Sov. Phys. Usp. **17**, 963 (1975)].

[4] M. Planck, *Vorlesungen über die Theorie des Wärmestrahlung* 5th edition, (Barth, Leipzig, 1923) [Russ. transl., ONTI, Leningrad, 1935].

[5] V. V. Sobolev, *Transport of Radiant Energy in the Atmospheres of Stars and Planets* [in Russian] (GITTL, Moscow, 1956).

[6] I. Prigogine, *Introduction to the Thermodynamics of Reversible Processes* 3rd edition (Interscience, New York, 1968) [Russian transl. of earlier ed., IL, Moscow, 1961].

[7] L. N. Bell and N. D. Gudkov, *Thermodynamics of Light Energy Conversion* (SPB Academic Publishing bv, The Hague, 1993).

[8] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 2 vols, 3rd ed., Pergamon Press, Oxford (1980) [Russ. original, Nauka, Moscow (1976, 1978)].

[9] R. T. Ross, Photochem. Photobiol. **21**, 401 (1975).

Translated by J. R. Anderson

Phonon focusing in CdSe, ZnS, and ZnO crystals

V. V. Zubritskiĭ

B. I. Stepanov Institute of Physics, Belarus Academy of Sciences, 220072 Minsk, Belarus

The energy focusing patterns of acoustic phonons in hexagonal CdSe, ZnS, and ZnO crystals at room temperature are calculated with allowance for the piezoelectric coupling. Estimates are presented, and it is shown that the piezoelectric stiffening and the values of the elastic and dielectric constants of the crystals, which characterize the processes by which the samples were grown, have a nonmonotonic influence on the focusing. Piezoelectric splitting of the phonon fluxes is observed, and it is found that the slow transverse acoustic (STA) modes are infinitely concentrated. The possibility of modifying the focusing anisotropy of STA phonons in hexagonal II–VI compounds by doping them with lithium is predicted. Experiments on acoustoelectronic generation in CdSe are analyzed, and it is concluded that the variety of high-frequency current oscillation spectra can be controlled by an isolated part of the focusing patterns of different modes. © *1997 American Institute of Physics.* [S1063-7842(97)01306-8]

INTRODUCTION

Phonon focusing has been observed in various crystalline media and continues to be a subject of intense study. Most of the research has been performed at temperatures low enough for superconducting detectors to function. The advent of new methods[1–3] has expanded the possibilities of investigating this effect at higher temperatures and of elucidating its role in various physical phenomena[4–6] and in the operation of specific acoustoelectronic devices: for example, monochromatic generation spectra have been obtained in an investigation of cadmium selenide acoustooptic cells on slabs with the direction lying in the normal angular region in which the concentration of shear waves occurs.[7]

All this calls for a detailed study of the phenomenon. However, the literature does not contain adequate information on the focusing of nonequilibrium phonons for some classes of materials, such as wide-gap piezoelectric semiconductors. For example, the collinear axes and the corresponding concentration factors for II–VI compounds were determined in Ref. 8 without taking the piezoelectric effect into account. In Ref. 9 the influence of the piezoelectric effect on the collinear directions and widths of the peak regions in the cross sections of the group velocities of the FTA modes in CdS and ZnO crystals was demonstrated. The limited information available clearly makes it difficult to interpret experimental results on the corresponding nonequilibrium properties of crystals and, consequently, to perform a purposeful search for new operating regimes for acousto-optoelectronic devices.

Therefore, in the present work we have calculated the focusing patterns of longitudinal acoustic (LA), fast transverse acoustic (FTA), and slow transverse acoustic (STA) phonons for hexagonal CdSe, ZnS, and ZnO crystals in the presence of the piezoelectric effect at room temperature and have analyzed their role in experiments on acoustoelectronic generation in CdSe slabs.

METHOD AND OBJECTS OF INVESTIGATION

The solution of the equation of propagation of the displacement in an anisotropic piezoelectric medium,[10] brought into the form

$$(c_{ijkl}^E n_j n_k + e_{nij} n_n n_j e_{mkl} n_m n_k / \varepsilon_{ij}^s n_i n_j - \rho v^2 \delta_{il}) u_l^0 = 0,$$

where c_{ijkl}^E, e_{nij}, and ε_{ij}^s are the components of the piezoelectric modulus tensor and the dielectric tensor, respectively, ρ is the density of the crystal, the n_i are the components of the unit vectors corresponding to the directions of propagation of the elastic wave, v is its velocity, and **u** is the polarization vector of the medium, was used to find the dependence of the inverse phase velocities on the direction of propagation. Meridian cross sections containing the **z** axis for the refraction surfaces of the LA, FTA, and STA phonons were analyzed using the generating relation[10,11] between the slowness surfaces and the fluxes of energy concentration. The concentration factors A_c for the cross sections were calculated on the basis of the method in Refs. 12 and 13. Both infinite ($A_c \rightarrow \infty$ when $G \rightarrow 0$, where G is the curvature of the slowness surface) and noninfinite ($A_c > 1$) concentration[14] were taken into account in the analysis. The criterion of transverse isotropy was used to check the correctness of the calculations.

Since the focusing can be sensitive to modification of the parameters of the crystals,[8,14] those material constants for which complete sets have been measured on the same type of high-resistivity samples were selected from the set of reference data. Specifically, the CdSe samples were grown from the gas phase and compensated by copper impurity atoms (~ 10 ppm), annealed at 700 °C for an hour and brought to room temperature by sudden cooling to increase its resistivity to $10^9 - 10^{10}$ $\Omega \cdot$cm and to prevent the formation of precipitates.[15] The data for ZnS correspond to crystals of 100% wurtzitic modification with a dark resistivity of $\sim 10^{12}$ $\Omega \cdot$cm and a MgS content equal to 10.3 mol%.[16] The ZnO crystals were annealed in air at 800 °C in a Li$_2$CO$_3$ melt for up to 5 days to increase the dark resistivity to

TABLE I. Material parameters of the crystals used in the calculations.

Crystal	CdSe	ZnS	ZnO	References
Elastic moduli, 10^{10} N/m²:				Refs. 15–17
c_{11}^E	7.41	12.94	20.70	
c_{12}^E	4.52	6.82	11.77	
c_{13}^E	3.93	5.34	10.61	
c_{33}^E	8.36	14.24	20.95	
c_{44}^E	1.317	2.72	4.48	
Piezoelectric moduli, C/m²:				Refs. 15–17
e_{15}	-0.138	-0.118	-0.37	
e_{31}	-0.160	-0.238	-0.62	
e_{33}	0.347	0.265	0.96	
Relative dielectric constants:				Refs. 15–17
ε_{11}^s	9.33	8.25	8.33	
ε_{33}^s	10.20	8.59	8.81	

TABLE II. Directions and concentration factors for longitudinal and transverse fast and slow acoustic phonons in the meridian plane of hexagonal CdSe, ZnS, and ZnO at room temperature.

	LA		FTA		STA	
Crystal	Θ_f, deg	A_c	Θ_f, deg[a]	A_c	Θ_f, deg[a]	A_c
CdSe	0.0	1.78	40.7*	6.85×10^2	60.8*	0.62
	180.0		319.3*		299.2*	
	90.0	1.23	49.2*	5.35×10^2	78.7*	1.09
	270.0		310.8*		281.3*	
			90.0	1.10	101.3*	1.09
			270.0		258.7*	
			130.8*	5.35×10^2	119.2*	0.62
			229.2*		240.8*	
			139.3*	6.85×10^2		
			220.7*			
ZnS	0.0	1.80	40.4*	4.11×10^2	55.6*	0.70
	180.0		319.6*		304.4*	
	90.0	1.38	49.9*	8.01×10^2	76.9*	1.11
	270.0		310.1*		283.1*	
			90.0	1.13	103.1*	1.11
			270.0		256.9*	
			130.1*	8.01×10^2	124.4*	0.70
			229.9*		235.6	
			139.6*	4.11×10^2		
			220.4*			
ZnO	0.0	1.66	36.5*	6.49×10^2	0.0	1.00
	180.0		323.5*		180.0	
	68.5	1.12	38.3*	1.21×10^3		
	291.5		321.7*			
	111.5	1.12	90.0	1.01		
	248.5		270.0			
			141.7*	1.21×10^3		
			218.3*			
			143.5*	6.49×10^2		
			216.5*			

[a]The directions corresponding to mathematically infinite concentration are marked by an asterisk.

10^9-10^{10} $\Omega\cdot$cm.[17] The errors for the determination of the elastic, piezoelectric, and dielectric constants (Table I) of CdSe, ZnS, and ZnO are equal to 0.2–4%, 0.3–4%, and 1–4%, respectively.

RESULTS AND DISCUSSION

Let us examine the influence of the piezoelectric effect on the focusing of the LA mode in CdSe, where this influence is less pronounced, in reference to the available data[8,9] for comparison. The 0.1–1.9% difference between the elastic moduli of the CdSe used in the present work from the values taken in Ref. 8 leads to an approximately 2.3% decrease in A_c. Piezoelectric coupling decreases $A_\perp$ by an additional 3.5%, but it has a reverse influence along the **z** axis, i.e., A_z increases by 11.5%. As a result, as is seen from Table II, in CdSe the increase in the energy flux of the LA phonons along the **z** axis (ΔA_z) in comparison with the isotropic case amounts to 78%, which is 2.15 times smaller than the value of ΔA_z in Ref. 8. In the orthogonal direction the analogous value is ~1.4 times smaller than the value in Ref. 8.

These variations of ΔA, which are caused by modification of the elastic properties of the samples and the inclusion of the piezoelectric coupling in the treatment, can serve as an estimate of the errors introduced into the analyzed intensities of the phonon beams and point out the overestimated values of A_z that were obtained in Ref. 8 without consideration of the piezoelectric effect. This is confirmed by the identical agreement of the results of the analytical calculations and the numerical results when the parameters in Ref. 8 are employed as the starting data. A similar situation is observed for the other crystals; therefore, the data from the numerical calculations are analyzed below.

The intensities of the focused LA phonon beams in ZnS as a function of the values of the constants are modified differently. The 1.1–5.1% change in the elastic moduli alone leads to a 2.0% decrease in $A_\perp$ but a 2.8% increase in A_z. The piezoelectric interaction enhances this tendency by 3.1% and 5.5%, respectively, and the resultant intensities are close to those obtained for CdSe (Table II).

In ZnO, modification of the elastic and piezoelectric constants has a more pronounced influence on the focusing of the LA mode. When the piezoelectric effect is disregarded, the concentration anisotropy in zinc oxide is similar to that obtained for CdSe and ZnS (Fig. 1). The 0.7–5.5% variation of the elastic moduli leads to 6.8% and 7.9% decreases in the focusing intensity along and perpendicular to the **z** axis, respectively. The presence of piezoelectric coupling causes the following changes: the intensity of concentration along the **z** axis increases by 45%, while defocusing is observed in the orthogonal direction ($A_\perp<1$). In addition, the direction of concentration along the twofold axis in the absence of piezoelectric coupling now splits into two (Fig. 1). In other words, focusing cones for the LA mode appear in the bulk in ZnO.

In the samples investigated the most intense focusing is exhibited by the fast transverse phonons. In crystal cuts containing the **c** axis there are eight directions corresponding to mathematically infinite concentration. This means that the energy of the FTA modes propagates within CdSe, ZnS, and ZnO predominantly along the generators of four cones at the angles Θ_f^{FTA} to the **c** axis (Table II). The angular interval between the generators of the inner and outer cones for CdSe and ZnS amounts to 8.5 and 9.5°, respectively, and the di-
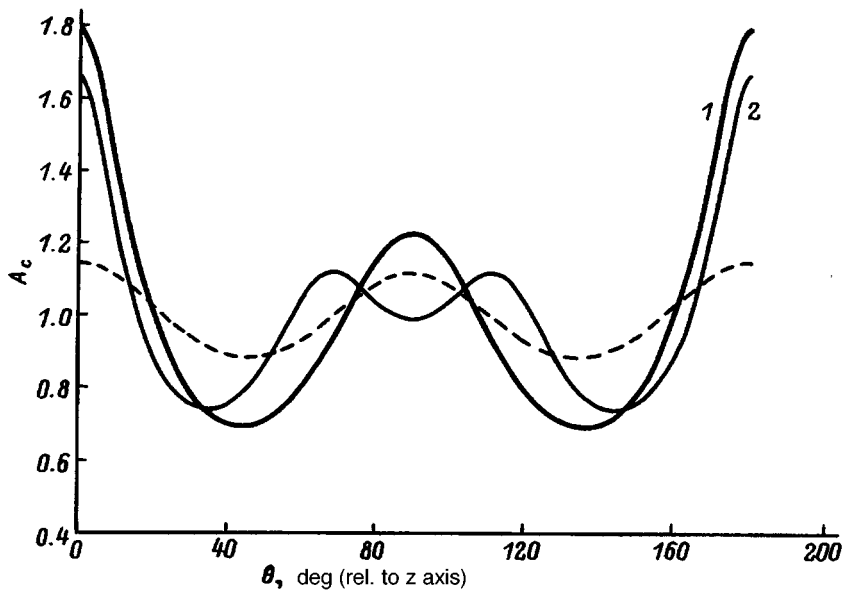
FIG. 1. Angular distribution of the concentration factor for longitudinal acoustic phonons in $\{hk0\}$ CdSe (*1*) and ZnO (*2*) (the dashed curve corresponds to the absence of the piezoelectric effect).

rections of maximum focusing are practically the same in these crystals.

A different situation is observed in ZnO. Here the generators of the inner and outer cones practically coincide, i.e., the angular interval between them equals 1.8° (Table II). The focusing directions are "pressed" about 4 and 11° closer to the **c** axis than in CdSe and ZnS. As is seen from Fig. 2, the role of the piezoelectric effect reduces in this case to a splitting and angular displacement of the distribution to a large increase in intensity, and to the appearance of a new focusing beam in the direction orthogonal to the **c** axis.

The analysis performed for the STA mode shows that in CdSe and 100% wurtzite there are eight directions satisfying the condition of mathematically infinite concentration. The corresponding focusing cones in both CdSe and ZnS "embrace" the focusing cone of the FTA phonons (Table II). However, in contrast to the FTA mode, the focusing direc-

tions of the STA phonons in CdSe and ZnS do not coincide with one another.

The literature known to us does not contain any reports from other investigators regarding the possibility of the infinite concentration of STA phonons in hexagonal semiconductors. The occurrence of such focusing has been noted only for CdS (Ref. 18) in connection with the problem of the incomplete electrical breakdown of crystals. A common feature here is the appreciable difference between the numerical infinite concentration amplitudes for the STA modes and the FTA phonons, which is caused mainly by features of the respective acoustical surfaces: the focusing-producing portions of the refraction surface for the fast transverse phonons are far more extensive than the analogous regions for the STA modes. Although the question of the underestimated numerical intensities is a subject for a separate treatment, the
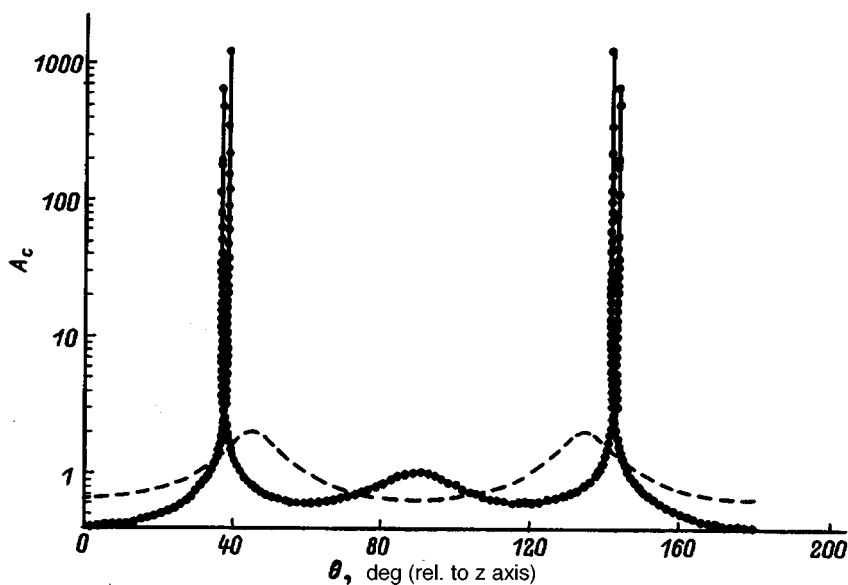
FIG. 2. Anisotropy of the concentration factor of ZnO for fast transverse phonons without (dashed curve) and with consideration of the piezoelectric coupling in a plane containing the **c** axis.
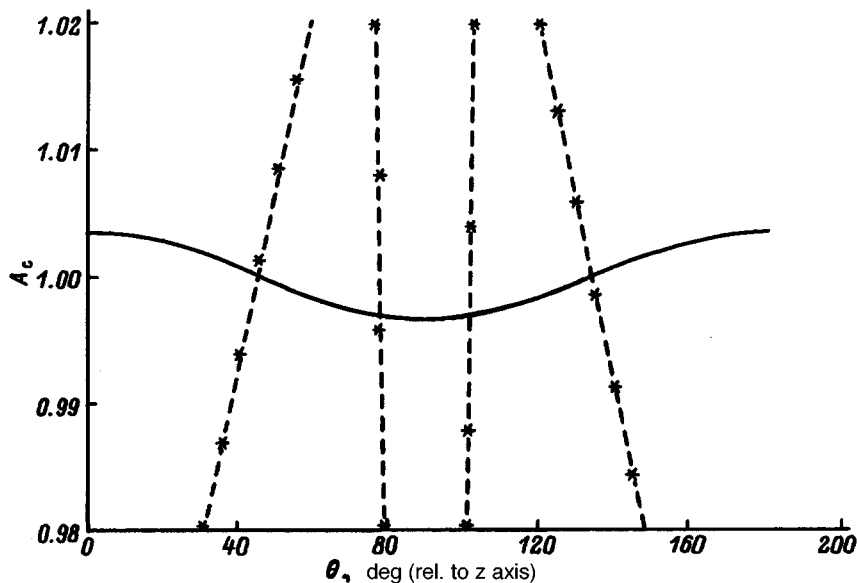
FIG. 3. Distribution of the concentration factor for slow transverse phonons in the meridian plane of ZnO. The dashed curve is the dependence obtained in Ref. 8.

data obtained confirm the occurrence of the focusing just cited in $A^{II}B^{VI}$ compounds.

Similar focusing was not discovered for the slow transverse phonons in the zinc oxide crystals investigated. The STA mode in ZnO is far less anisotropic than the LA modes in the crystals investigated. It is seen from Fig. 3 that the distribution of its concentration factor practically corresponds to an isotropic medium.

As we know,[10] the STA modes are piezoelectrically passive. However, simulation reveals a strong dependence of the anisotropy of their energy flux on the variation of the elastic properties of the samples considered above. As it turns out, the STA mode in ZnO that has not been specially doped has infinite focusing, while it is not exhibited by thermally treated crystals. Thus, the observed (Fig. 3) fundamental variation of the focusing anisotropy is caused by the variation of the elastic properties due to the strong doping of the samples.

The property discovered should also be characteristic of other wide-gap $A^{II}B^{VI}$ compounds and can be verified, for example, by comparing the focusing anisotropy of the STA phonons in CdSe, CdS, or ZnS before and after thermal treatment similar to that described above for ZnO. After doping with lithium, isotropy of the focusing of the energy of the STA modes should be observed in these crystals.

APPLICATION TO ACOUSTOELECTRONIC GENERATION EXPERIMENTS

The absence of exact agreement between the phonon distributions, which presumably follows from the fact that the materials investigated belong to a single class of crystals, and the observed ''elasticity'' of the phonon focusing can be utilized to obtain samples with assigned anisotropy of the distribution of the phonon fluxes. For example, the focusing of the STA modes can be suppressed by doping the crystals with lithium, and/or the angular gap between the inner and outer focusing cones of the FTA phonons can be regulated.

When the devices have finite dimensions, the importance of the latter, like the possibility of uniquely selecting the crystallographic orientation of cuts, is clearly seen, if the results obtained are applied to experimental data[7] from an investigation of the high-frequency current oscillations accompanying the generation of ultrasound in plates of the 15 and 49° CdSe cuts.

Since the structure of the near-contact region was not reported in Ref. 7, we shall assume that the plates used are homogeneous across their entire thickness (0.2 mm), and we shall take into account the area of the electrodes (1–10 mm^2), as well as the mean free path of the focused phonons (l_{ph}), which is of the order of 2 mm in the bulk of semiconductors of different perfection.[19,20] Then, it is seen from the data in Table II and the geometry of the experiment[7] that in the case of a detector of minimal diameter, besides the two intense FTA phonon beams at $\Theta_f^{FTA} = 40.7$ and $49.2°$, beams of LA and STA modes, the geometric lengths of whose paths ($L \approx 0.3$ mm) correspond to l_{ph}, also participated in the acoustoelectronic interaction in plates of the 49° cut. The nonlinearity of the frequency dependence and the switching of the generation to a regime with a different oscillation spectrum as the drift parameter increased were caused mainly by the modulating influence of the beams along $\Theta_f^{FTA} = 40.7°$ and the faster (by a factor of 1.4) beam along $\Theta_f^{LA} = 0$.

Similarly, in devices on plates of the 15° cut, where it was assumed that focusing has no influence, the following three groups of beams had the strongest influence on acoustoelectronic generation: $\Theta_f^{FTA} = 310.9$ and $319.3°$ (I), $\Theta_f^{LA} = 0$ (II), and $\Theta_f^{FTA} = 40.7$ and $49.2°$ (III). The maximum path length for the beams of group I was 1.6–1.9 times greater than the values of L_{FTA} for the equivalent beams of group III. When the illumination was removed, the decrease in the concentration of free charge carriers was sufficient for eliminating the modulating contribution of the beams of group I. This could promote the appearance of gaps in the spectra (see Figs. 4b and 4c in Ref. 7), as well as the exci-

tation of individual spectral lines as the conductivity of the samples varies.

A similar influence of oblique beams of focused phonons could be observed when the analogous spectra were recorded in photoconductive CdS plates oriented perpendicularly to the **c** axis (see Fig. 1a in Ref. 21 and the data in Ref. 18).

The effect indicated above of focused beams of nonequilibrium phonons of different directivity on the stability of the spectra of the high-frequency current oscillations observed in other investigations can be confirmed without invoking tedious methods by selecting the ''required'' portion of the phonon distribution with, for example, more miniaturized electrodes of appropriate geometry,[7] The contribution of the STA mode and the role of the focusing cones of the FTA phonons can also be determined by comparing the generation spectra in CdSe and ZnO plates.

CONCLUSIONS

In contrast to the available literature data, in the present work phonon focusing was assessed for hexagonal CdSe, ZnS, and ZnO, and the diversity of the energy focusing patterns of different phonon modes in similar crystals was demonstrated. The observed nonmonotonic nature of the influence of the modification of the material parameters of the crystals on the focusing points out the need for an individual examination of the focusing of each acoustic mode in a concrete sample, if the measured set of elastic, piezoelectric, and dielectric constants is outside of the range investigated in the present work. The data obtained ensure the required level of consideration of the anisotropy of the transport of the nonequilibrium phonons when various physical properties of the samples investigated are studied and permit the purposeful development of acousto-optoelectronic devices on their basis for various wavelength ranges.

[1] A. A. Kolomenskiĭ and A. A. Maznev, JETP Lett. **53**, 423 (1991).
[2] M. R. Hauser, R. L. Weaver, and J. P. Wolfe, Phys. Rev. Lett. **68**, 2604 (1992).
[3] A. A. Kolomenskii and A. A. Maznev, Phys. Rev. B. **48**, 14 502 (1993).
[4] V. K. Komar', V. A. Kornienko, V. P. Migal' et al., Pis'ma Zh. Tekh. Fiz. **20** (10), 71 (1994) [Tech. Phys. Lett. **20**, 797 (1994)].
[5] V. V. Zubritskiĭ, Zh. Tekh. Fiz. **63** (5), 61 (1993) [Tech. Phys. **38**, 399 (1993)].
[6] J. A. Shields and J. P. Wolfe, Phys. Rev. B. **50**, 8297 (1994).
[7] V. N. Shorin, in *Exact Measurements in Acousto-optics and Optoelectronics. A Collection of Scientific Papers* [in Russian], Moscow, VNIIFTRI (1985), pp. 45–52.
[8] A. K. McCurdy, Phys. Rev. B. **9**, 466 (1974).
[9] A. K. McCurdy, in *Phonon Scattering in Condensed Matter. Proceedings of the 5th International Conference*, A. C. Anderson and J. P. Wolfe (eds.), Springer-Verlag, Berlin and other cities (1986), pp. 296–298.
[10] E. Dieulesaint and D. Royer, *Elastic Waves in Solids*, Wiley, New York (1981) [Russ. transl., Nauka, Moscow (1982)].
[11] F. I. Fedorov, *Theory of Elastic Waves in Crystals*, Plenum Press, New York (1968).
[12] M. Lax and V. Narayanamurti, Phys. Rev. B. **22**, 4876 (1980).
[13] G. Every, Phys. Rev. B. **24**, 3456 (1981).
[14] V. V. Zubritskiĭ, Pis'ma Zh. Tekh. Fiz. **18** (16), 72 (1992) [Sov. Tech. Phys. Lett. **18**, 539 (1992)].
[15] D. Berlincourt, H. Jaffe, and L. R. Shiozawa, Phys. Rev. **129**, 1009 (1963).
[16] I. A. Dan'kov, I. B. Kobyakov, and S. Yu. Davydov, Fiz. Tverd. Tela (Leningrad) **24**, 3613 (1982) [Sov. Phys. Solid State **24**, 2058 (1982)].
[17] E. F. Tokarev, I. B. Kobyakov, I. P. Kuz'mina et al., Fiz. Tverd. Tela (Leningrad) **17**, 980 (1975) [Sov. Phys. Solid State **17**, 629 (1975)].
[18] V. V. Zubritskiĭ, Fiz. Tverd. Tela (St. Petersburg) **38**, 56 (1996) [Phys. Solid State **38**, 31 (1996)].
[19] M. Schreiber, M. Fieseler, A. Masur et al., in *Proceedings of the 18th International Conference on Physics of Semiconductors*, O. Engström (ed.), World Scientific, Singapore (1987), pp. 1373–1376.
[20] E. Held, W. Klein, and R. P. Huebener, Z. Phys. B **75**, 17 (1989).
[21] H. M. Janus and N. I. Meyer, Solid State Commun. **8**, 417 (1970).

Translated by P. Shelnitz

Splitting of an electromagnetic pulse on resonant reflection from a plasma film

M. I. Bakunov and N. S. Gurbatov

N. I. Lobachevski Nizhniĭ Novgorod State University, 603600 Nizhniĭ Novgorod, Russia

It is found theoretically that the temporal profile of a quasimonochromatic electromagnetic pulse is strongly distorted on reflection from a thin (on the wavelength scale) film of a plasma-like (semiconductor, metallic) medium under plasma-resonance conditions. It is shown that an incident Gaussian pulse splits with time (completely or partially) into two reflected pulses, whose amplitudes can be controlled by varying the relationships among the parameters of the incident pulse and the film. © *1997 American Institute of Physics.* [S1063-7842(97)01406-2]

INTRODUCTION

Transverse magnetic (TM) waves can interact effectively with films of a plasma-like (semiconductor, metallic) medium that are thin on the wavelength scale. As was shown for the first time in Ref. 1, thin plasma films are capable of strongly reflecting TM waves provided that the frequency of the wave is close to the plasma frequency of the film (plasma resonance) and that collisions are sufficiently infrequent. The possibility of the strong absorption of TM waves by plasma films was discovered in Ref. 2 also under plasma-resonance conditions. The resonant shielding and resonant absorption of TM waves by thin plasma films has been studied fairly thoroughly for monochromatic waves (see, for example, Refs. 3–5). Some features of the absorption of electromagnetic pulses in plasma films were examined in Ref. 6.

This report describes a new resonance effect, viz., strong distortion of the temporal profile of a quasimonochromatic electromagnetic pulse on reflection from a plasma film located on a perfectly conducting substrate. It is caused by the strong dispersion of the phase of the reflection coefficient in a narrow range of frequencies near the plasma frequency of the film. It is shown that the incident Gaussian pulse splits with time (completely or partially) into two reflected pulses, whose amplitudes can be controlled by varying the relationship between the parameters of the incident pulse and the film.

The reshaping of an electromagnetic pulse under plasma-resonance conditions can be utilized to reshape pulses of submillimeter radiation on reflection from semiconductor films, as well as ultrashort (femtosecond) optical pulses on reflection from metallic films. We note that the controlled reshaping of femtosecond laser pulses has been actively discussed in the literature (see, for example, Refs. 7–9). As will be shown below, the effect considered in the present work differs significantly from the conceptually similar method of transforming pulses under conditions of excitation of surface waves[8] and has some advantages in respect to practical implementation. The preliminary results of this work were reported in Ref. 10.

STATEMENT OF THE PROBLEM. BASIC FORMULAS

Suppose that a Gaussian quasimonochromatic pulse of TM polarization with a magnetic field

$$B_z^i(\xi) = A_i(\xi/\tau)\exp(i\omega_0\xi),$$

$$A_i(\xi/\tau) = B_0 \exp(-\xi^2/2\tau^2), \tag{1}$$

where $\xi = t - (x\cos\theta + y\sin\theta)/c$ and $\omega_0\tau \gg 2\pi$, is incident at an angle θ from a vacuum ($x<0$) onto a homogeneous plasma film ($0<x<d$) located on a perfecting conducting substrate ($x=d$). The film is assumed to be thin on the scale of the wavelength of the incident radiation

$$\frac{\omega_0 d}{c} \ll 1 \tag{2}$$

and when collisions are neglected, it is characterized by the dielectric constant $\varepsilon = \varepsilon_L(1 - \omega_p^2/\omega^2)$, where ω_p is the plasma frequency of the free carriers and ε_L is the dielectric constant of the lattice (in the case of a gas plasma $\varepsilon_L=1$).

The carrier frequency ω_0 of the pulse described by (1) is assumed to be close to the plasma frequency of the film (plasma resonance)

$$\frac{|\omega_0 - \omega_p|}{\omega_p} \ll 1. \tag{3}$$

Let us examine the reflected pulse at the point $x=y=0$ [the results for other points differ only by displacement in time: $t \rightarrow t + (x\cos\theta - y\sin\theta)/c$]. The reflected pulse is specified by the Fourier integral of the product of the spectrum of the incident pulse

$$F(\omega) = \frac{B_0\tau}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}(\omega-\omega_0)^2\tau^2\right] \tag{4}$$

and the reflection coefficient $R(\omega)$ of a monochromatic wave. When conditions (2) and (3) are taken into account, the latter can be described to good accuracy in the form[5,11]

$$R(\omega) = \frac{\omega - \omega_p + i\alpha\omega_p}{\omega - \omega_p - i\alpha\omega_p}, \tag{5}$$

where

$$\alpha = \frac{\omega_p d \sin^2\theta}{2c\varepsilon_L \cos\theta} \tag{6}$$

is a small parameter characterizing the thickness of the plasma film ($\alpha \ll 1$).
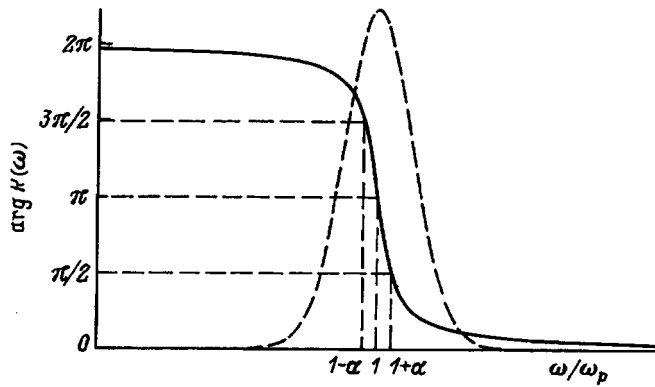
FIG. 1. Dependence of the phase of the reflection coefficient on the frequency for $\alpha = 0.05$ (solid curve) and qualitative form of the spectrum of the incident pulse for $\omega_0 = \omega_p$ (dashed curve).

FIG. 2. Temporal profile of the reflected [$|A_r(t/\tau)|$, solid lines] and incident [$A_i(t/\tau)$, dashed lines] pulses for $\Delta = 0$. γ: a — 0.1, b — 0.22, c — 0.6.

The absolute value of the reflection coefficient (5) equals unity $[|R(\omega)| = 1]$, and its phase

$$\arg R(\omega) = \arctan \frac{2\alpha\omega_p(\omega - \omega_p)}{(\omega - \omega_p)^2 - \alpha^2\omega_p^2} \qquad (7)$$

varies as a function of the frequency by 2π in a narrow range of frequencies with a width amounting to several $\alpha\omega_p$ in the vicinity of the plasma frequency of the film ω_p (Fig. 1).

Since we intend to perform a numerical analysis of the temporal profile $|A_r(t/\tau)|$ of the reflected pulse

$$B_z^r(t) = A_r(t/\tau)\exp(i\omega_0 t), \qquad (8)$$

it would be convenient to go over to the dimensionless variable $\eta = (\omega - \omega_0)\tau$ in the Fourier integral. For the complex amplitude $A_r(t/\tau)$ we obtain

$$A_r(t/\tau) = \frac{B_0}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\eta + \Delta + i\gamma}{\eta + \Delta - i\gamma} \exp\left(-\frac{\eta^2}{2} + i\eta\frac{t}{\tau}\right) d\eta, \qquad (9)$$

where $\Delta = (\omega_0 - \omega_p)\tau$ is a parameter characterizing the detuning from resonance, and the parameter $\gamma = \alpha\omega_p\tau$ characterizes the ratio between the width of the frequency range of the resonant phase change of the reflection coefficient ($\sim \alpha\omega_p$) and the width of the spectrum of the input pulse ($\sim 1/\tau$).

RESULTS

The results of the numerical calculations of the shape of the reflected pulse on the basis of Eq. (9) are presented in Figs. 2 and 3. In the case of exact agreement between ω_0 and ω_p, i.e., when $\Delta = 0$, the reflected pulse splits into two parts, which are completely separated by a zero on the temporal profile (Fig. 2). Qualitatively, this splitting can be attributed to the fact that the central components of the spectrum of the incident pulse, which fall on the steep portions of the function $\arg R(\omega)$ (Fig. 1), experience a time delay,[12] while the peripheral components of the spectrum, corresponding to the gently sloping portions of $\arg R(\omega)$, are reflected without a delay.
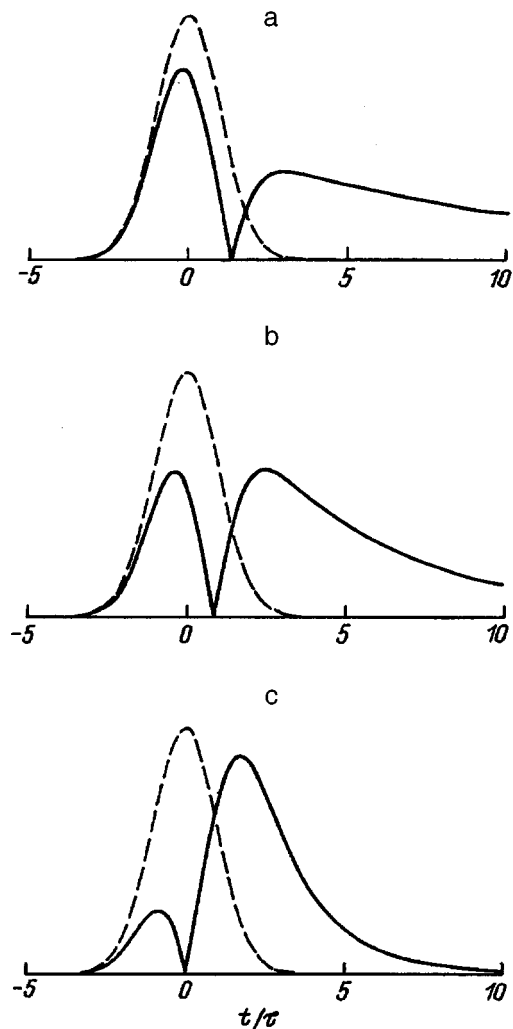
Clearly, as γ increases, i.e., as the width of the resonance region increases in comparison to the width of the spectrum of the incident pulse, the delayed part of the reflected pulse grows, and the leading part shrinks (Fig. 2).

The presence of a zero of the envelope is a characteristic feature of the case $\Delta = 0$. Figure 4 presents the numerically calculated dependence of the time when the envelope drops to zero t_*/τ on γ (solid curve). This dependence can be explained using a vector diagram (Fig. 5), in which the vectors correspond to narrow portions of the spectrum of the reflected pulse and revolve with time at an angular velocity $\omega - \omega_0$. Because of the symmetry of the diagram relative to the $0 - \pi$ line, the vertical components of the vectors cancel one another at each moment in time. The horizontal components cancel out only at a certain time, which depends on the value of γ (compare Figs. 2, 4, and 5).

When $\Delta \neq 0$ the vector diagram is asymmetric, and mutual cancellation of the vectors does not occur. Therefore, the leading and delayed parts of the reflected pulse are connected by a juncture of finite magnitude, and when Δ is sufficiently large, they merge (Fig. 3).

In the limit $\gamma \gg 1$ it is not difficult to obtain a simple

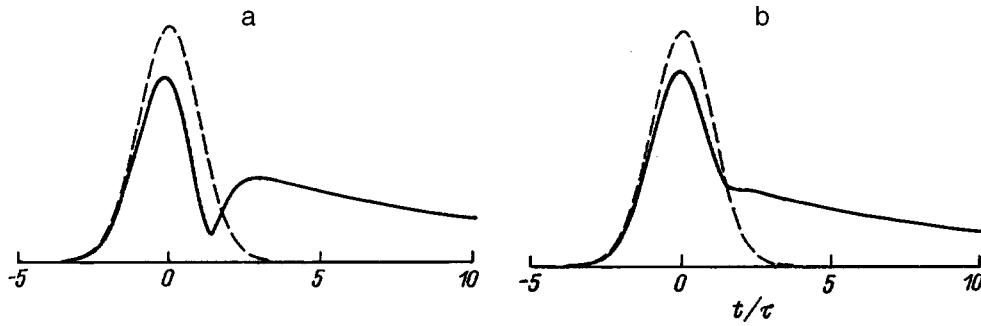FIG. 3. Temporal profile of the reflected $[|A_r(t/\tau)|$, solid lines] and incident $[A_i(t/\tau)$, dashed lines] pulses for $\gamma=0.1$. Δ: a — 0.2, b — 0.7.

analytical expression for the shape of the reflected pulse. In Eq. (9) we go over to the variable $\widetilde{\eta}=\eta/\gamma$. The large parameter γ then appears in the exponent of the exponential function in the integrand, making it possible to employ the method of steepest descent. As a result, under the condition $t/\tau\ll\gamma$, which ensures that the steepest-descent contour $\mathrm{Im}\,\widetilde{\eta}=t/\gamma\tau$ remains far from the pole $\widetilde{\eta}=-\Delta/\gamma-i$ on the plane of the complex values of $\widetilde{\eta}$, we obtain

$$|A_r(t/\tau)|\approx A_i(t/\tau)\sqrt{\frac{(t/\tau+\gamma)^2+\Delta^2}{(t/\tau-\gamma)^2+\Delta^2}}. \tag{10}$$

Equation (10) describes a pulse which differs from the incident pulse by the presence of a small leading part. When $\Delta=0$, it is separated from the main part of the pulse by a zero of the envelope at the point

$$\frac{t_*}{\tau}=-\gamma \tag{11}$$

on the time axis. The straight line described by Eq. (11) corresponds to the asymptote in Fig. 4 (the dashed line). If $\Delta\neq0$, Eq. (10) implies the presence of a juncture between the split parts of the pulse.

One characteristic feature of the temporal structure of the reflected pulse envelope is the presence of a slowly decaying trailing edge or ''tail'' on the delayed part of the pulse (Figs. 2 and 3). It can be shown by the methods of contour integration in the complex η plane that this ''tail'' corresponds to the contribution of the pole of the integrand $\eta=-\Delta+i\gamma$, and that for $t/\tau\gg1$ and $t/\tau\gg\gamma$ it is described by an exponential dependence of the form

$$|A_r(t/\tau)|\approx2\sqrt{2\pi}B_0\gamma\exp\left(\frac{\gamma^2-\Delta^2}{2}\right)\exp\left(-\gamma\frac{t}{\tau}\right). \tag{12}$$

Physically, the presence of the exponential tail described by Eq. (12) is caused by the emission of electromagnetic

FIG. 4. Dependence of the position of the zero on the temporal profile of the reflected pulse on the value of γ for $\Delta=0$.
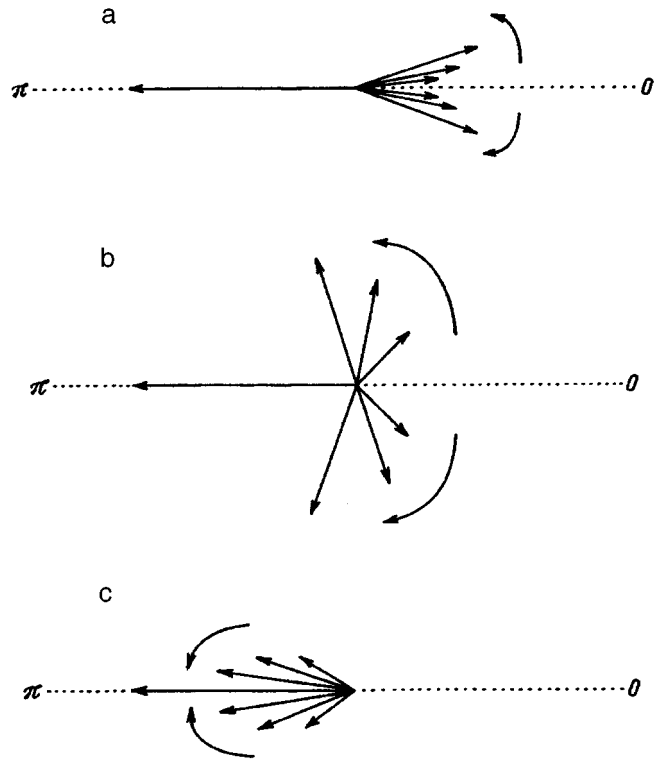

FIG. 5. Vector diagrams for the reflected pulse at $t=0$. a — $\gamma\ll1$, b — $\gamma\sim1$, c — $\gamma\gg1$.

energy stored in the plasma layer as a consequence of the resonant swelling of the component E_x of the electric field.[13]

CONCLUSIONS

In the present work the plasma-resonant splitting of a quasimonochromatic electromagnetic pulse on reflection from a plasma film has been examined in a pure form, i.e., as a dispersion effect. This is ensured by the absence of collisions, which can be neglected under the condition[11]

$$\nu \ll 2\alpha\omega_p, \tag{13}$$

where ν is the effective collision frequency.

When ν increases and inequality (13) is violated, resonant absorption of the electromagnetic wave in the film begins to play a role. As an analysis shows, this is manifested mainly as a decrease in the amplitude of the delayed part of the reflected pulse.

Unlike the method for reshaping laser pulses under conditions of excitation of surface plasmons,[8] the effect considered here is not critical for the angular spread of the incident light beam owing to the lack of a dependence of the resonant frequency on the angle of incidence θ and the weak dependence of α on θ over a broad range of angles of incidence [see Eqs. (5) and (6)]. Moreover, practical implementation of the effect does not require the use of an ATR (attenuated total reflection) prism (cf. Ref. 8).

[1] I. G. Kondrat'ev and M. A. Miller, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **7**, 124 (1964).

[2] R. P. Godwin, Phys. Rev. Lett. **28**, 85 (1972).

[3] A. A. Zharov and I. G. Kondrat'ev, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **20**, 1474 (1977).

[4] A. K. Kotov, Fiz. Plazmy **11**, 629 (1985) [Sov. J. Plasma Phys. **11**, 368 (1985)].

[5] M. I. Bakunov and Yu. M. Sorokin, Opt. Spektrosk. **66**, 692 (1989) [Opt. Spectrosc. (USSR) **66**, 405 (1989)].

[6] A. S. Sakharov, Preprint No. 43 [in Russian], P. N. Lebedev Physics Institute, Russian Academy of Sciences, Moscow (1981).

[7] I. H. Campbell and P. M. Fauchet, Opt. Lett. **13**, 634 (1988).

[8] R. V. Andaloro, H. J. Simon, and R. T. Deck, Appl. Opt. **33**, 6340 (1994).

[9] C. W. Hillegas, J. X. Tull, D. Goswami, *et al.*, Opt. Lett. **19**, 737 (1994).

[10] M. I. Bakunov and N. S. Gurbatov, in *Abstracts of the 2nd International Scientific School-Seminar ''Dynamic and Stochastic Wave Phenomena,''* Nizhny Novgorod, 1994, p. 47.

[11] M. I. Bakunov and S. N. Zhukov, Zh. Tekh. Fiz. **61** (12), 25 (1991) [Sov. Phys. Tech. Phys. **36**, 1340 (1991)].

[12] V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas*, 2nd ed., Pergamon Press, Oxford–New York (1970) [Russ. original, 2nd ed., Nauka, Moscow (1967), Sec. 21].

[13] I. G. Kondrat'ev and M. A. Miller, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **7**, 176 (1964).

Translated by P. Shelnitz

Calculation of the interaction force between a relativistic electron beam and an Ohmic plasma channel

E. K. Kolesnikov and A. S. Manuĭlov

V. I. Smirnov Scientific-Research Institute of Mathematics and Mechanics, 198904 St. Petersburg, Russia

A formula for calculating the interaction force between a relativistic electron beam and a preformed Ohmic plasma channel with an arbitrary offset of the channel axis from the beam axis is obtained in the case of complete charge neutralization. It is shown that this force is repulsive for radial profiles of the conductivity with a peak on the channel axis. © *1997 American Institute of Physics.* [S1063-7842(97)01506-7]

New areas of application of relativistic electron beams (REBs) call for further investigation of the dynamics of the transport of REBs in gas-plasma media.[1–17] A study of the conditions for the stable transport of a beam along Ohmic plasma channels is of special interest among the problems associated with REB transport. In particular, several cases in which plasma channels have a stabilizing influence on REB propagation were considered in Refs. 9–12,14, and 16. The tracking force attracting a beam to a channel was calculated in Ref. 8 in the electrostatic approximation under conditions of a low-conductivity Ohmic channel. In addition, channels in which the bulk of the plasma return current is located outside the beam have been considered in Refs. 10,11,14, and 16. When the REB is laterally offset, this situation leads to weakening of the hose modes.[2,3] The results of a numerical simulation of the beam–plasma interaction during the transport of REBs along Ohmic plasma channels with allowance for the production of conductivity as a result of impact and avalanche ionization of the channel gas were presented in Ref. 9. It was shown that electrostatic tracking (i.e., attraction toward the Ohmic channel) occurs in the head part of the beam, while repulsion of the REB away from the channel is observed in the main part of the beam (the ''body'' of the REB). The latter effect is clearly caused by the increase in conductivity due to impact and avalanche ionization of the channel gas and the corresponding increase in the destabilizing plasma return current near the channel axis. However, it is impossible to assess the magnitude of the interaction force between a beam and an Ohmic channel from the results in Ref. 9.

In the present work we obtain a formula for calculating the beam–channel interaction force under conditions of complete neutralization of the space charge of the REB for arbitrary values of the offset of the channel axis from the beam axis.

Let us consider a paraxial monoenergetic axially symmetric REB with an arbitrary radial profile of the current density $J_b(r)$ propagating in a gas-plasma medium along a preformed Ohmic channel [with a radial conductivity profile $\sigma_{ch}(r)$] parallel to the z axis of a cylindrical coordinate system (r,θ,z). We shall henceforth confine ourselves to the case of a high-conductivity channel, in which the condition of complete neutralization of the REB space charge holds [$4\pi\sigma_{ch}(0)R_b/c \gg 1$, where R_b is the characteristic radius of

the beam and c is the speed of light]. To calculate the interaction force between a beam and a preformed channel we assume that the channel is offset from the beam axis by an arbitrary distance Y_{ch}. Then, using the Biot–Savart–Laplace law, we obtain the force which the plasma channel exerts on one beam electron:

$$F_1 = \frac{ec}{I_b}\int_0^\infty dr\, 2\pi r J_b(r)$$
$$\times \int_r^\infty d\rho\, \rho \int_0^{2\pi} d\theta\, \cos\theta\, \frac{2J_p(|\boldsymbol{\rho}-\mathbf{Y}_{ch}|)}{\rho c^2}, \tag{1}$$

where $J_p(r)$ is the radial profile of the plasma return current density flowing in the channel, I_b is the total beam current, and e is the charge of the electron.

After integration by parts with the use of Ohm's law for J_p, we have

$$F_1 = \frac{2e}{cI_b}\int_0^\infty dr I_b(r)\int_0^{2\pi} d\theta\, \cos\theta E_z(\Psi)\sigma_{ch}(\Psi), \tag{2}$$

where

$$I_b(r) = 2\pi \int_0^r d\rho\,\rho J_b(\rho) \tag{3}$$

is the beam current through a tube of radius r, E_z is the z component of the collective electric field of the plasma–beam system, and

$$\Psi^2 = r^2 + Y_{ch}^2 - 2rY_{ch}\cos\theta. \tag{4}$$

As was shown in Ref. 17, the axial electric field E_z depends weakly on Ψ; therefore, we take E_z out from under the integrals over r and θ. In addition, the force exerted by the channel on a unit length of the beam, F_{dt}, is related to F_1 in the obvious way:

$$F_{dt} = \frac{F_1}{\gamma mc^2}, \tag{5}$$

where γ is the relativistic factor of the beam particles and m is the electron mass.

Then we have

$$F_{dt} \simeq -\frac{2}{\pi}\frac{I_b}{I_A}\frac{f_m}{R_b^2}\int_0^\infty dr\,\tilde{I}_b(r)\int_0^{2\pi} d\theta\,\cos\theta\tilde{\sigma}_{ch}(\Psi), \tag{6}$$

where $I_A = \gamma m c^3/e$ is the limiting Alfvén current,

$$f_m = -\frac{\sigma_{ch}(0)E_z}{(I_b/\pi R_b^2)} \tag{7}$$

is the current neutralization coefficient on the channel axis, $\widetilde{\sigma}_{ch} = \sigma_{ch}/\sigma_{ch}(0)$, and $\widetilde{I}_b(r) = I_b(r)/I_b$.

We use Eq. (6) just derived to calculate the interaction force between a beam and an Ohmic channel in the linear case when $Y_{ch}/R_b \ll 1$, the REB has a Bennett radial beam current density profile, and the conductivity of the Ohmic channel has a "step-like" form

$$\sigma(r) = \begin{cases} \sigma_{ch}^0, & r \leqslant R_{ch}, \\ \sigma_p, & r > R_{ch}. \end{cases} \tag{8}$$

In this case we expand the dependence of σ on Ψ in a Taylor series to terms of first order in the small parameter Y_{ch}. Then we have

$$\sigma(\Psi) = \sigma(r) - Y_{ch}\cos\theta \frac{\partial\sigma}{\partial r}. \tag{9}$$

Substituting Eq. (9) into (6), we ultimately obtain

$$F_{dt} = -\frac{2}{\pi R_b}\frac{I_b}{I_A}\frac{Y}{R_b}(f_m - f_{mp})\frac{R_{ch}^2}{(R_b^2 + R_{ch}^2)}, \tag{10}$$

where $f_m = -\sigma_{ch}^0 E_z/I_b^*$, $f_{mp} = -\sigma_p E_z/I_b^*$, and $I_b^* = I_b/(\pi R_b^2)$.

Taking into account that $E_z \approx \text{const}$ and that $f_m, f_{mp} > 0$, we find that F_{dt} is a tracking force (which attracts toward the channel) if $\sigma_p > \sigma_{ch}^0$ and a detracking force (which repels from the channel) if $\sigma_p < \sigma_{ch}^0$. In the former case ($\sigma_p > \sigma_{ch}^0$) the main return current flows outside the beam (at $r > R_{ch}$); therefore, if the REB is offset from the channel axis it will be repelled toward its original position, in agreement with the results in Refs. 14 and 16. In the latter situation the bulk of the plasma return current flows within the beam itself and the Ohmic channel. When the center of mass of the REB is laterally offset from the channel axis, displacement of the beam from the region of greatest conductivity occurs. It should be noted here that in the present work the choice of coordinate system and displacement vector $\mathbf{Y}_{ch}$ was such that a tracking force would be positive and a detracking force would be negative. In the opposite situation, in which the beam is displaced relative to a stationary plasma channel, the tracking force becomes negative.

We next obtain a formula for calculating the beam–channel interaction force F_{dt} in the nonlinear case, in which $Y_{ch} \geqslant R_b$. For this purpose we assume that the beam and the channel have Bennett radial profiles with different characteristic scales R_b and R_{ch}. Then after integrating over the azimuthal angle θ, from (6) we have

$$F_{dt} = -4\frac{I_b}{I_A}\frac{f_m}{R_b}\int_0^\infty d\rho\, \widetilde{I}_b(\rho)\rho$$

$$\times \frac{\widetilde{Y}_{ch}}{\eta^2\left[\left(1 + \frac{\rho^2 + \widetilde{Y}_{ch}^2}{\eta^2}\right)^2 - \left(\frac{2\rho\widetilde{Y}_{ch}}{\eta^2}\right)^2\right]^{3/2}}, \tag{11}$$
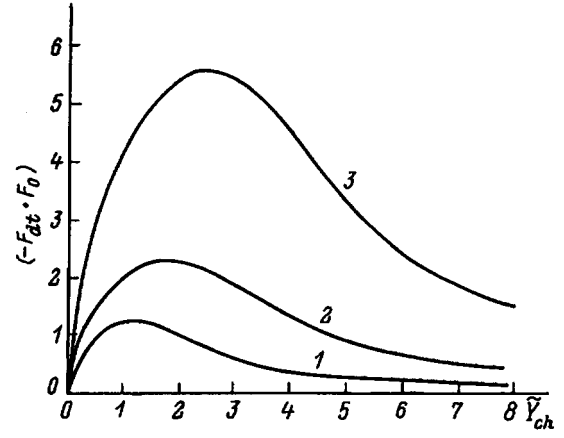

FIG. 1. Dependence of $(-F_{dt} \cdot F_0)$ on the offset of the channel axis $\widetilde{Y}_{ch}$ ($F_0 = 10^{-2}$ [1/cm]) for various values of η. η: 1 — 0.5, 2 — 1, 3 — 2.

where $\eta = (R_{ch}/R_b)$, $\rho = r/R_b$, and $\widetilde{Y}_{ch} = Y_{ch}/R_b$.

Figure 1 presents plots of the dependence of F_{dt} on $\widetilde{Y}_{ch}$ for various values of η ($\eta = 0.5, 1, 2$) and $\xi = 75R_b$ (ξ is the distance from the beam front to the specified transverse section of the REB). The beam parameters were taken as follows: $R_b = 0.5$ cm, $I_b = 10$ kA, and $E = 5$ MeV ($\gamma = 10$). In addition, the beam current obeys the growth law

$$I_b(\xi) = I_b \tanh\left(\frac{\xi}{\xi_r}\right), \tag{12}$$

where $\xi_r = 30R_b$.

As estimates show, under these conditions for $\sigma_{ch}(0) = 6 \times 10^{11}$ s^{-1} the current neutralization coefficient $f_m \simeq 0.5$. It is seen from Fig. 1 that in the case of the high-conductivity Ohmic channel under consideration (in which complete charge neutralization occurs) F_{dt} is a detracking force at the conductivity peak in the center of the channel. We note that this result coincides qualitatively with the data from the numerical simulations in Ref. 9. As the results in Refs. 10,11,14, and 16 and Eq. (10) in the present communication show, an increase in conductivity from the center toward the periphery of the plasma channel is necessary to obtain a tracking force.

Thus we have obtained a formula for finding the beam–channel interaction force in a magnetic regime. We have shown that ejection of the beam from the region occupied by the channel will occur in the case of radial profiles of the Ohmic conductivity with a peak at the center of the channel and a decrease toward the periphery, as is confirmed by the results of the experimental work in Ref. 12.

[1] E. P. Lee, Phys. Fluids **19**, 60 (1976).
[2] E. P. Lee, Phys. Fluids **21**, 1327 (1978).
[3] H. S. Uhm and M. Lampe, Phys. Fluids **23**, 1574 (1980).
[4] W. A. Barletta, E. P. Lee, and S. S. Yu, Nucl. Fusion **21**, 961 (1981).
[5] R. F. Fernsler, R. F. Hubbard, and M. Lampe, J. Appl. Phys. **75**, 3278 (1994).
[6] E. R. Nadezhdin and G. A. Sorokin, Fiz. Plazmy **9**, 989 (1983) [Sov. J. Plasma Phys. **9**, 576 (1983)].
[7] W. M. Sharp and M. Lampe, Phys. Fluids **23**, 2383 (1980).
[8] E. P. Lee, Livermore Lab. Report No. UCID-19674 (1983), p. 10.
[9] B. Hui and M. Lampe, in *5th International Conference on High Power Particle Beams*, San Francisco, 1983, pp. 374–377.

[10] D. P. Murphy, R. E. Pechacek, D. P. Taggart *et al.*, Phys. Fluids. B **4**, 3407 (1992).

[11] D. R. Welch, F. M. Bieniosek, and B. B. Godfrey, Phys. Rev. Lett. **65**, 3128 (1990).

[12] D. P. Murphy, M. Raleigh, R. E. Pechacek *et al.*, Phys. Fluids **30**, 232 (1987).

[13] E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **60**(3), 40 (1990) [Sov. Phys. Tech. Phys. **35**, 298 (1990)].

[14] E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **61**(12), 43 (1991) [Sov. Phys. Tech. Phys. **36**, 1351 (1991)].

[15] E. K. Kolesnikov, A. S. Manuĭlov, and I. V. Abashkina, Zh. Tekh. Fiz. **64** (11), 136 (1994) [Tech. Phys. **39**, 593 (1994)].

[16] R. F. Hubbard, R. F. Fernsler, S. P. Slinker *et al.*, in *5th International Conference on High Power Particle Beams*, San Francisco, 1983, pp. 370–373.

[17] M. Lampe, W. M. Sharp, R. F. Hubbard *et al.*, Phys. Fluids **27**, 2921 (1984).

Translated by P. Shelnitz

Formation of localized silver centers on the surface of titanium dioxide films using a scanning tunneling microscope

V. P. Poroshkov and V. S. Gurin

Scientific-Research Institute of Physicochemical Problems, Belarus State University, 220080 Minsk, Belarus

A scanning tunneling microscope (STM) is used to create changes in the surface relief of titanium dioxide films containing adsorbed silver ions. Structures measuring ~ 10 nm, which presumably consist of silver particles, form on the film surface as a result of the application of short pulses with an amplitude $\geqslant 15$ V to an STM probe operating in the tunneling-current regime. © *1997 American Institute of Physics.* [S1063-7842(97)01606-1]

INTRODUCTION

The control of processes leading to the formation of ultrasmall surface relief elements on the surface of a solid requires the development of special methods for treating surfaces. The employment of a scanning tunneling microscope (STM) makes it possible not only to study the surfaces of solids down to the atomic level, but also to perform local treatment of a surface and to thereby create inhomogeneities in the composition of the surface and structural changes covering distances smaller than 10–100 nm using various materials.[1–7]

A TiO$_2$ film is an example of an electrochemically and photoelectrochemically active semiconductor material, whose properties can be modified by depositing metal particles on its surface.[8,9] In an ordinary electrochemical experiment metal particles are deposited on the surface of a semiconductor in direct contact with an electrolyte solution containing metal ions. Under the conditions of an experiment with an STM, the role of the counterelectrode can be performed by the probe, but the presence of a liquid electrolyte phase is undesirable from the standpoint of forming localized structures. Nanometer-scale metal particles can be formed from a fairly small number of metal ions, which can be contained in a thin adsorbed layer, without the use of an electrolyte containing metal ions. The fact that growth of the particles can then take place without appreciable mass transport through a solution promotes their localization at the sites of primary modification of the film surface.

In the present work TiO$_2$ films with an adsorbed layer of silver ions are proposed as systems for studying the possibility of forming local structures using an STM. The electrochemical properties of such films vary sharply as a result of reduction of the Ag$^+$ ions to metallic particles. Such an STM-stimulated reduction process is investigated as the result of the application of pulses of increased potential ($\leqslant 20$ V) to the probe, which is located over a specific point on the surface, followed by scanning of the particular area to record the resultant surface relief pattern.

The preliminary creation of active sites or other inhomogeneities, whose electrochemical behavior differs from that of the original surface, using an STM in such a manner permits the development of methods for the local selective cathodic deposition of various metals on the surfaces of semiconductor electrodes and for the electrochemical formation of metal films of specified configuration and dimensions.

EXPERIMENTAL

The surface relief was investigated and modified on a computer-controlled system based on an STM fabricated in the "Delta" Scientific-Research Institute (Moscow). The STM probe was made from a platinum wire. The surface relief was recorded at a constant probe potential equal to 0.1 V in the tunneling-current regime (~ 1 nA), and the surface was modified by the action of pulses (of negative polarity)
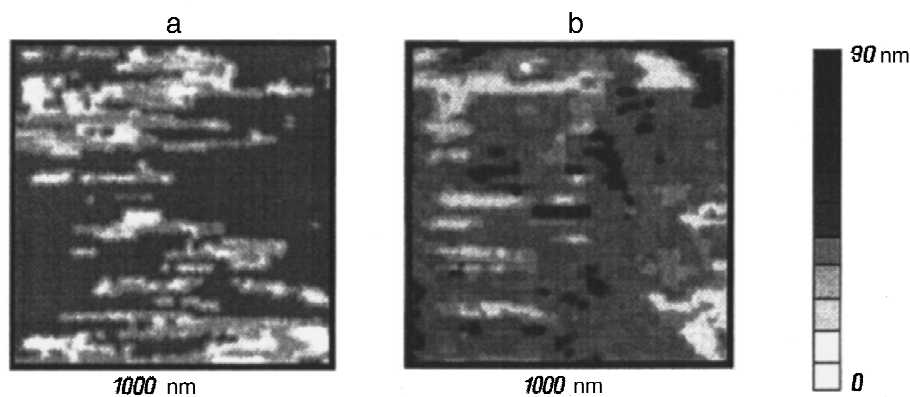

FIG. 1. STM image of the relief of an unmodified TiO$_2$ film (a) and a film modified by electrodeposited silver particles (10^{14} atoms/cm^2) (b). The height of the surface relief is expressed by different shades of gray according to the scale presented.
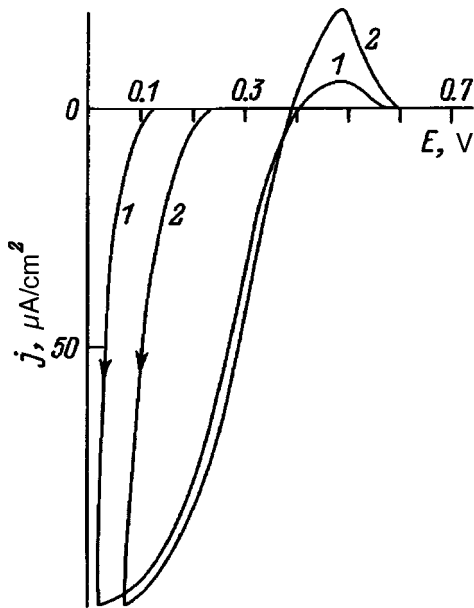
FIG. 2. Voltammograms of unmodified Ti–TiO$_2$ electrodes (1) and electrodes modified by silver particles in an amount equal to 10^{14} atoms/cm^2 (2). The concentration of AgNO$_3$ in the electrolyte was 0.005 M. The potential scale is given relative to a saturated silver chloride electrode.

with amplitudes up to 50 V from an external generator. The pulse duration was $\leqslant 1$ ms, which is less than the time constant of the feedback circuit of the STM.

The TiO$_2$ films were obtained by a hydrolytic method. A film of metallic titanium or gold with a thickness of 100–200 nm was deposited on a polished wafer of single-crystal silicon. This film then served as a current-conducting substrate for the TiO$_2$ film subsequently applied to its surface as a result of the hydrolysis of a 0.001% solution of polybutyltitanate in *tert*-butanol. The polybutyltitanate solution was applied to the horizontally oriented surface of the titanium- or gold-coated silicon wafer and then slowly (over the course of 5–10 h) dried in *tert*-butanol vapor. Next, the sample was heated to 450–500 °C in vacuum for 10 min, during which a polycrystalline film of TiO$_2$ of the anatase modification with a thickness equal to 5–10 nm formed. The surface resistivity of such samples amounted to 10–100 kΩ/cm. The character of the voltammograms obtained in an electrochemical experiment for Ti–TiO$_2$ electrodes fabricated in this manner scarcely differed from those for TiO$_2$ films on a metallic titanium substrate, either in a background electrolyte or during the selective cathodic deposition of silver in a solution of AgNO$_3$.

An adsorbed layer of Ag$^+$ ions was formed by immersing samples with TiO$_2$ films in an aqueous 10^{-2}–10^{-3} M solution of AgNO$_3$ for 100 s. The Ag$^+$ ions adsorbed under the conditions form film structures containing 10^{-7}–10^{-9} g/cm^2 silver,[10] in which particles measuring 1–5 nm form upon irradiation.
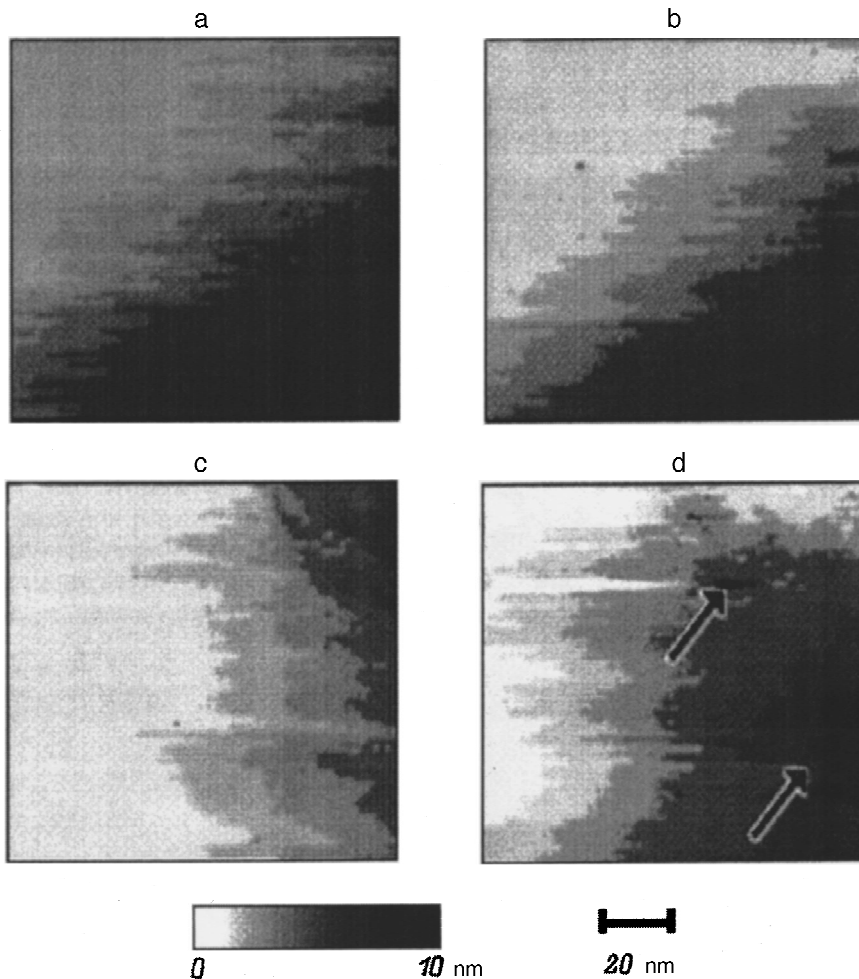

FIG. 3. STM images of the surface relief of a TiO$_2$ film, modified in a 10^{-3} M aqueous solution of AgNO$_3$ for 60 s, before (a, c) and after (b, d) pulsed treatment of the surface in an STM. The pulse amplitude was 12 (b) and 15 V (d). The pulse duration was 100 μs (b, d). The arrows point to embossed structures formed after the pulsed treatment, which are attributable to the silver particles. The height of the surface relief is expressed by different shades of gray according to the scale presented.
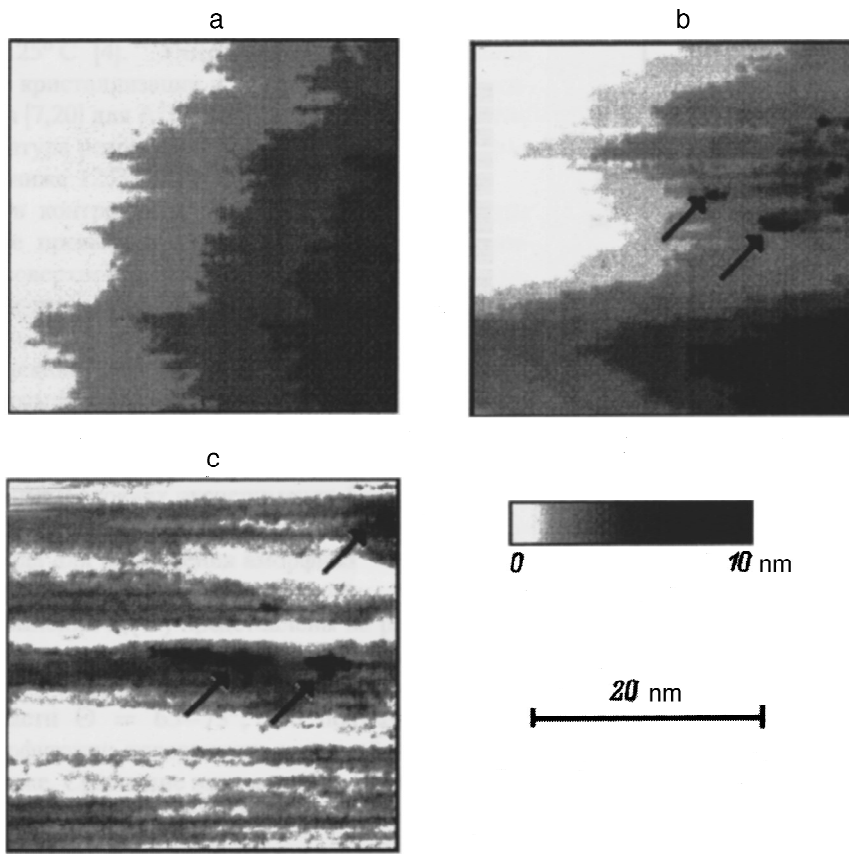
a   b

c

0   10 nm

20 nm

FIG. 4. STM images of the surface relief of a TiO$_2$ film, modified in a 10^{-3} M aqueous solution of AgNO$_3$ for 60 s, before (a) and after (b) pulsed treatment in an STM. The pulse amplitude was 18 V (b); the pulse duration was 1 ms (b). (c) — STM image of the surface of a sample similar to (b) after the construction of a sectional plane that eliminates the deviation of the surface of the TiO$_2$ film from a horizontal orientation relative to the working plane of the microscope. The arrows point to embossed structures formed after the pulsed treatment, which are attributable to the silver particles. The height of the surface relief is expressed by different shades of gray according to the scale presented.

EXPERIMENTAL RESULTS

The STM image of the surface relief of a Ti–TiO$_2$ film electrode not containing silver is presented in Fig. 1a. It is characterized by a relatively even surface with small inhomogeneities associated with the polycrystallinity of the film. When silver particles are electrodeposited outside the STM under galvanostatic conditions, the formation of inhomogeneities of the surface relief with horizontal dimensions of 10–50 nm and vertical dimensions of 10–20 nm is observed on the surface of the TiO$_2$ film (Fig. 1b). These inhomogeneities are appreciably greater than those of the original film, and their dimensions correlate with the results of electron-microscope investigations of similar samples.[12,13] The electrochemical deposition of silver on the surface of such a modified Ti–TiO$_2$ electrode outside the STM takes place with a smaller overvoltage than on an unmodified electrode (Fig. 2). When the potential of the Ti–TiO$_2$ electrode equals ~ 0.25 (relative to a saturated silver/silver chloride electrode), the deposition rate is at least 10 times higher on the electrode modified with particles than on the unmodified electrode.

A TiO$_2$ film with adsorbed Ag$^+$ ions was employed to create localized centers of inhomogeneity on the film surface using an STM. No stable changes in the relief resulting from the action of pulses with amplitudes up to 15–20 V on films not containing silver ions were noted, although higher potentials created considerable nonlocal changes over the entire area scanned. Figures 3a, 3c, and 4a present the STM images of the surface of a TiO$_2$ film modified by Ag$^+$ ions. The relief pattern, which demonstrates the uniform tilting of the

surface of the sample without appreciable inhomogeneities, does not exhibit significant differences from the image of the surface of an unmodified Ti–TiO$_2$ film electrode (compare these figures with Fig. 1a). Thus, the adsorbed ions are not displayed on the image at this STM resolution. After pulsed treatment of the surface of a modified TiO$_2$ film in the STM with a pulse amplitude of ~ 15 V and a duration of 100 μs, areas of inhomogeneity with horizontal dimensions $\leqslant 10$ nm and vertical dimensions equal to 2–5 nm appear on the relief (Fig. 3d). Increases in the pulse amplitude and duration do not lead to the formation of larger objects or an increase in their number (Figs. 4b and 4c), while pulses with an amplitude less than 12 V do not cause any alteration of the relief (Fig. 3b).

The surface relief inhomogeneities formed after the pulses can be attributed to reduction of the Ag$^+$ ions and the formation of silver particles, as is confirmed by the formation of silver particles when a TiO$_2$ film with adsorbed Ag$^+$ ions is UV-irradiated outside the STM.[12] The action of a pulse can lead to the reduction of adsorbed Ag$^+$ ions to atoms followed by growth of the particles to dimensions corresponding to those observed. This is possible for the high mobility of highly disperse silver.[14,15] Larger particles do not form when the pulse amplitude and duration are increased, probably because the number of Ag$^+$ ions adsorbed on the area where the probe acts is small (their total concentration is $\leqslant 10^{-7}$ g/cm^2).

It should be noted that a semiconductor electrode for which the selective cathodic deposition of metals has been established was employed for local modification of the sur-

face in the STM. Although STM-stimulated electrochemical processes are also possible for some other semiconductors (GaAs, Ge, graphite),[16-18] for the TiO_2–Ag system considered here one can carry out a subsequent selective electrodeposition of the metal on the centers created using the STM.

CONCLUSIONS

Thus, the experiments performed have demonstrated the fundamental possibility of using an STM to create localized centers, which are presumably formed by silver particles, as a result of the STM-simulated reduction of silver ions in an adsorbed layer on the surface of TiO_2 films. Such silver particles can serve as metal deposition centers (even outside the STM) for forming metallized surface structures and thin metal films in elements of nanometer-scale electronic devices.

[1] C. F. Quate, *Scanning Tunneling Microscopy and Related Methods*, Kluwer Acad. Publ., Dordrecht (1990), pp. 281–297.
[2] G. M. Shedd and P. Russell, Nanotechnology **1**, 67 (1990).
[3] R. Wiesendanger, Appl. Surf. Sci. **54**, 271 (1992).
[4] G. G. Vladimirov and A. L. Gryazev, Vestnik SPb. Gos. Univ., Ser. 4 (4), 24 (1993).
[5] H. Rohrer, Microelectron. Eng. **27**, 3 (1995).
[6] V. S. Gurin, Zh. Nauchn. Prikl. Fotogr. **38** (3), 67 (1993).
[7] V. S. Gurin, E. A. Tyavlovskaya, A. I. Sharendo, and O. G. Kulinkovich, Zh. Tekh. Fiz. **64**(2), 165 (1994) [Tech. Phys. **39**, 205 (1994)].
[8] A. L. Linsebigler, G. Lu, and J. T. Yates Jr., Chem. Rev. **95**, 735 (1995).
[9] A. I. Kulak, *Electrochemistry of Semiconductor Heterostructures* [in Russian], Izd. Universiteta, Minsk (1986).
[10] Yu. V. Nechepurenko, G. A. Ragoĭsha, V. G. Sokolov, and G. A. Branitskiĭ, Zh. Nauchn. Prikl. Fotogr. Kinematogr. **28**(3), 181 (1983).
[11] V. P. Poroshkov, V. S. Gurin, and N. I. Kuntsevich, Elektrokhimiya **29**, 1275 (1993).
[12] V. P. Poroshkov, V. S. Gurin, N. I. Kuntsevich, and V. V. Sviridov, Vestsi Akad. Navuk Belarusi: Ser. Khim. (1), 21 (1993).
[13] V. S. Gurin, V. P. Poroshkov, K. N. Kasparov, E. A. Tyavlovskaya, and N. I. Kuntsevich, Zh. Fiz. Khim. **67**, 2442 (1993).
[14] K. Heinemann and H. Poppa, Thin Solid Films **33**, 237 (1976).
[15] G. E. Rhead, Int. Mater. Rev. **34** (6), 261 (1989).
[16] T. Thundat, L. A. Nagahara, and S. M. Lindsay, J. Vac. Sci. Technol. A **8**, 539 (1990).
[17] R. Richter and H. L. Harthagel, Int. J. Electron. **69**, 631 (1990).
[18] W. Li, J. A. Virtanen, and R. M. Penner, Appl. Phys. Lett. **60**, 1181 (1992).

Translated by P. Shelnitz

Modeling of surfaces of constant force above a lattice of close-packed atoms in the repulsive mode

E. V. Blagov, G. L. Klimchitskaya, A. A. Lobashev, and V. M. Mostepanenko

Northwestern Correspondence Polytechnical Institute, 191065 St. Petersburg, Russia

Surfaces of constant force (force contours) are calculated for the scanning of an AFM tip over a lattice of close-packed atoms in the repulsive mode. It is shown that discontinuities are observed on the force contours in the regions between the atoms of the surface lattice for sufficiently small initial scanning heights of a tip with a single atom at its end. A cluster model of the tip end, which ensures continuity of the scanning at arbitrary initial heights, is constructed. The dependence of the AFM images on the orientation of the cluster on the tip end relative to the crystallographic axes of the surface is investigated for both an unperturbed lattice of close-packed atoms and a lattice containing point defects. The diagnostic possibilities of the findings are discussed. © *1997 American Institute of Physics.* [S1063-7842(97)01706-6]

INTRODUCTION

Atomic force microscopy (AFM), which is based on the scanning of a tip along a line of constant force exerted on it by a solid surface, can be utilized to investigate the surface relief of any substance.[1–5] As we know, the maximum horizontal resolution of AFM (which is of the order of several angstroms) can be achieved in the repulsive mode (which is sometimes called the contact mode), in which the tip is brought close to the surface of the body under investigation, to within a distance not exceeding the distances between the atoms. The interaction between the tip and the surface is then determined mainly by the repulsive exchange forces between the closest atoms on the tip end and the surface beneath the tip [the contribution of the attractive van der Waals forces in the contact mode does not exceed 10% (Ref. 6)].

The scanning of an AFM tip gives a force contour, which describes the shape of the surface under investigation. Such force contours have been obtained with atomic resolution in the repulsive mode for leucosapphire,[6] boron nitride,[7] molybdenum disulfate,[7] graphite,[7–10] and others. Interpretation of the results requires the performance of theoretical calculations of force contours (see, for example, Refs. 11–14). The results of such calculations, however, depend on the tip models used.[15,16] This dependence is especially significant for the repulsive mode. In fact, in the attractive mode the distance between the tip and the surface under investigation is large in comparison with the interatomic distances; therefore, the tip can be modeled in the form of a continuous paraboloid or cone.[11,16,17] Various models of tips both with a single atom and with a definite atomic cluster at the end have been used to describe the scanning process in the repulsive mode,[12,13,15,18,19] with a large portion of the results pertaining to investigations of graphite surfaces.

This paper discusses the calculation of force contours when a lattice of close-packed atoms is scanned by an AFM tip. It is shown that discontinuities of the force contours appear in the regions between the surface atoms for sufficiently small initial scanning heights of an AFM tip with one atom at its end. The appearance of these discontinuities on the force contours (which were first discussed briefly in Ref. 20)

causes the tip to dip into the surface being investigated during the scanning. Such events had previously been detected experimentally (see, for example, Ref. 13), but they were not interpreted as a consequence of discontinuities on the force contours. In the present work it was established that continuous scanning at fairly small initial heights (i.e., fairly large repulsive forces) is possible only when a tip with a definite atomic cluster at its end is used, and the minimum dimensions of that cluster were found for a lattice of close-packed atoms.

1. A MONATOMIC TIP: DISCONTINUITIES ON THE FORCE CONTOURS

Let us consider the scanning of an AFM tip over the surface of the close-packed lattice shown in Fig. 1 (the dark circles are atoms in the surface layer of the sample, and the light circles are the projections of atoms in the second layer of the lattice onto the surface layer). The interaction of the atoms in the sample and the tip is usually described[2,3,7–9,11,15,19] using the pairwise Lennard-Jones potential

$$V_{ij} = \frac{\alpha}{r_{ij}^{12}} - \frac{\beta}{r_{ij}^{6}}, \tag{1}$$

where i labels the atoms in the sample and j labels the atoms in the tip.

As was noted in the Introduction, when the scanning is carried out in the repulsive mode, the contribution of the attractive forces is small;[6] therefore, the second term on the right-hand side of Eq. (1) can be neglected.

The potential energy $U(x,y,z)$ of the interaction between the tip and the sample is obtained by summing the potentials (1) over the possible values of i and j (or only over i in the case of a monatomic tip) and depends on the coordinates of the tip end x, y, and z. It is convenient to use a coordinate system in which the xy plane coincides with the surface of the sample and the z axis is directed toward the tip.

If we neglect the displacement of the surface and tip atoms during the scanning [in any case, this is certainly valid
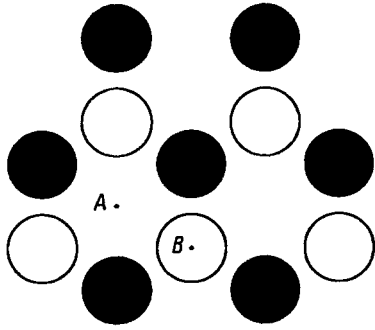
FIG. 1. Crystal cell on the surface of a lattice of close-packed atoms with projections of the atoms of the lower-lying layer.
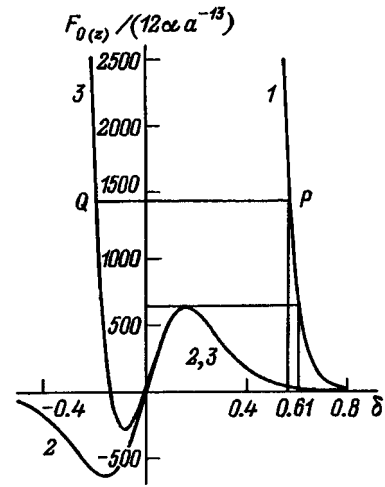

FIG. 2. Dependence of the normalized initial force $F_0/(12\alpha a^{-13})$ between a monatomic tip located above a surface atom on the relative height of the tip above the surface $\delta = d/a$ (1) and dependence of the normalized vertical component of the force $F_z/(12\alpha a^{-13})$ on the relative distance $\delta = z/a$ between a monatomic tip and the surface for tip positions A (B) above the surface (Fig. 1) (2, 3). Values $\delta < 0$ correspond to positions of the tip below the surface layer.

for a scanning force $F_0 \leq 10^{-9}$ N (Refs. 12 and 13)], the force contour can be found by solving the equation

$$-\frac{\partial U(x,y,z)}{\partial z} = F_0 \qquad (2)$$

for z.

To obtain results which do not depend on the choice of the value of the semiempirical parameter α in Eq. (1), it is convenient to go over to dimensionless (relative) expressions for the forces and coordinates. The equilibrium distance a between the atoms in the crystal lattice is naturally used as the unit of length, and the force which acts on a monatomic (i.e., having one atom at its end) tip located at an initial height d above a surface atom is employed as the unit of force. It is easy to see that in the case of a monatomic tip Eq. (2) for finding the force contour takes the form

$$\left(\frac{a}{d}\right)^{13} = \sum_i \frac{z - z_i}{[(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2]^7}, \qquad (3)$$

where the values of the coordinates are indicated in relative units.

The solution of Eq. (3) was sought by numerical methods, the summation being carried out over all the atoms of the surface under investigation that are sufficiently close to the tip to make some significant contribution to the result. It was established that no solution of Eq. (3) exists for $d < 0.61a$ at certain values of x and y, indicating the presence of a discontinuity on the force contour. This can easily be seen, if we compare the dependence of the repulsive force between the tip and surface on the height above the surface at three characteristic points: in a position directly above a surface atom and in positions A and B above the centers of interatomic triangles (Fig. 1). Plots of this dependence are presented in Fig. 2 (curves 1–3, respectively). If we compare curves 1 and 2 in Fig. 2, it becomes clear that a continuous force contour exists only when the condition $d \geq 0.61a$ is satisfied. In fact, when $d < 0.61a$, the value of the initial force corresponding to curve 1 (for example, the value of the force at point P) exceeds the maximum force value corresponding to curve 2, so that Eq. (3) cannot be satisfied. Therefore, the force contour has a discontinuity in the region above point A.

To determine the absolute value of the force at which the surface corresponding to it becomes discontinuous, knowl-

edge of the exchange-coupling constant α is required. Unfortunately, it is presently known with very poor accuracy. Nevertheless, using the value $\alpha \approx 3 \times 10^{19}$ H/nm[13] and $d = 0.6a \approx 0.61$ nm,[12] we obtain $F = 12\alpha/d^{13} \approx 8 \times 10^{-9}$ N for the force.

We note that position A in Fig. 1 corresponds to a point on the close-packed surface under which there is no second-layer atom. Let us now consider a position of the AFM tip above a second-layer atom of the sample (position B in Fig. 1). It is easily seen that Eq. (3) has a solution in the vicinity of point B, although the tip drops below the surface layer of atoms in the sample. This situation can be illustrated by comparing curves 1 and 3 in Fig. 2. When curve 3 was obtained, it was assumed that the shortest distance between a surface atom and an atom in the lower-lying layer is $0.9a$ when relaxation is taken into account. This distance corresponds to a distance of $0.78a$ between the layers. As follows from a comparison of curves 1 and 3 in Fig. 2, each point P on curve 1 corresponds to a point Q on curve 3 with the same value of the force. Thus, the force contour above position B is continuous for any value of d, if the interaction of the tip with the second-layer atoms of the sample is taken into account. Nevertheless, here the tip drops below the surface of the sample under investigation at some values of d. This can also be interpreted as dipping of the tip into the surface of the sample. It can easily be seen that consideration of the third layer of atoms in the sample and higher-lying layers of atoms in the tip does not significantly alter the picture just described (in performing the calculations it was assumed for simplicity that the crystal structure of the tip material is the same as that of the sample).

A qualitative proof of the existence of discontinuities on the force contours when the initial scanning heights are sufficiently small can be obtained by considering the scanning of a monatomic tip over one surface atom, as was done in
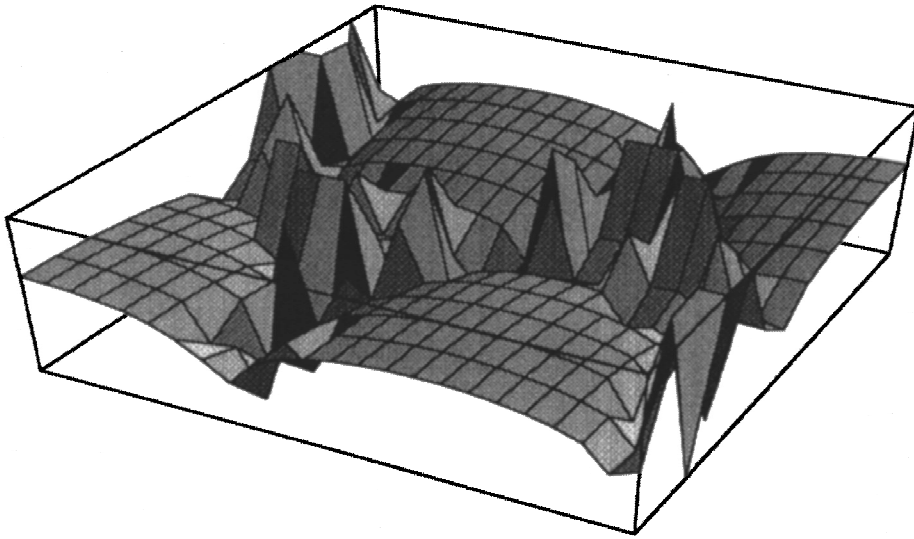
FIG. 3. Three-dimensional image of a force contour.

Refs. 20 and 21. In fact, a simple calculation shows that at small values of d the surface atom ceases to support the monatomic tip during scanning before the neighboring surface atom begins to support it.

An attempt to solve Eq. (3) by numerical methods at the discontinuities on the force contours leads to a divergent iteration process. Termination of such a process at any fixed number of iteration cycles yields a highly unordered set of peaks. To illustrate this, Fig. 3 presents a three-dimensional image of the force contour obtained by solving Eq. (3) for $d = 0.5a$. The portions of the force contour in the regions above atoms of the surface under investigation and the sharp peaks in the regions between atoms, where the force contour has a discontinuity, are clearly seen in Fig. 3. We note that the initial scanning height selected $d = 0.5a$ exceeds the sum of the radii assigned to the tip and surface atoms in the hard-sphere model in Ref. 22. Therefore, the use of the potential defined by Eq. (1) to describe the repulsive forces is perfectly justified.

2. POLYATOMIC TIP MODELS

Along with the simplest tip model (one atom at the end), some more complicated models have recently been discussed in the literature,[11–16,18,20,21,23] being of interest for numerous reasons. For example, it has been found that the model of a monatomic tip does not permit a faithful description of several features of the experimentally obtained force contour for graphite.[12,15,23] A model of a tip containing a cluster of several atoms at its end has also been used to describe the deformation of the surface and the tip during scanning[14,18,24] and to ensure the continuity of the scanning of a tip over the surface of a close-packed lattice, including a surface with point defects.[20,21] There is also considerable interest in such models in connection with the practice of using AFM data to determine the real structure of the tip end.[16,25]

The wide variety of problems to be solved is responsible for the diversity of the tip models used. The simplest generalization of a monatomic tip is a model consisting of a small number (2–4) of atoms arranged in a single plane.[14,15] The use of such models, however, does not permit a faithful de-

scription of the experimental results obtained from investigations of graphite.[14] In addition, direct calculations confirm that the corresponding force contours are not continuous, as in the case of a monatomic tip (Sec. 1).

The models in which the macroscopic shape of the tip (paraboloid or cone) and the structure of the crystal lattice are taken into account are more realistic. Two types of lattices are usually considered, viz., the tetragonal diamond structure[14,18,23] and the close-packed structure,[20,21,24] and an atomic cluster consisting of several atoms is found on the tip end. On the basis of general arguments regarding the interaction of atoms[22,26] it is difficult to expect that all the atoms in the cluster would be located in a single plane perpendicular to the axis of the tip. It would be more natural to assume that one of the atoms in the cluster is lower than the others (specifically at the tip end), while the other atoms are arranged on the surface of the tip in a plane perpendicular to its axis. If the tip is a paraboloid of revolution, its surface (in the coordinate system described in Sec. 1) can be assigned by the equation

$$z_t = \frac{x_t^2 + y_t^2}{2R} + d, \qquad (4)$$

where R is the radius of curvature of the tip.

The coordinates of the cluster atoms on the tip end should satisfy Eq. (4); however, due to the surface relaxation of the tip atoms and their repulsion from the surface under investigation, the mean distance between the atoms should be smaller than in the unperturbed lattice.

To obtain the theoretical force contour for such a polyatomic tip, the following equation instead of (3) must be solved for z:

$$\left(\frac{a}{d}\right)^{13}$$

$$= \sum_j \sum_i \frac{z + z_j' - z_i}{[(x + x_j' - x_i)^2 + (y + y_j' - y_i)^2 + (z + z_j' - z_i)^2]^7}, \qquad (5)$$

where x_j', y_j', and z_j' are the coordinates (in relative units) of the cluster atoms at the tip end in a coordinate system whose origin is at the very end.

The calculations previously performed[20,21] showed that the smallest configuration which provides for scanning stability (i.e., the absence of discontinuities on the force contour at any possible values of the initial height d) is a cluster of seven atoms. Six of them form a regular hexagon with a side $b=0.85a$, and the seventh is located at a distance $h=b^2/(2R)$ beneath its center. When $R=(5/3)a$, we have $h=0.22a$. It is clear that the force contour obtained by solving Eq. (5) depends on the orientation of the cluster on the tip end relative to the crystallographic axes of the surface,

$$z = z(x,y,\varphi), \tag{6}$$

where φ is the angle between the axes of the atomic hexagons of the cluster on the tip end and of those on the close-packed surface.

The calculation procedure depends highly on how the scanning is carried out. If the orientation of the cluster relative to the surface varies randomly during the scanning, averaging must be performed, and the averaged force contour

$$z(x,y) = \frac{3}{\pi} \int_0^{\pi/3} z(x,y,\varphi)\, d\varphi \tag{7}$$

must be compared with experimental data.

If, on the other hand, the orientation of the cluster on the tip end remains strictly fixed during the scanning, all the calculations must be performed for that fixed orientation. The surfaces (6) for different φ, of course, can differ markedly.

The averaged surfaces (7) were examined quite thoroughly in the preceding studies.[20,21] Below we shall examine the dependence on φ of the force contours (6) obtained by solving Eq. (5) for scanning over both an unperturbed surface of close-packed atoms and a surface containing point defects in the form of vacancies and divacancies.

3. FORCE CONTOURS ABOVE A CLOSE-PACKED LATTICE: DEPENDENCE ON THE TIP ORIENTATION

This section presents the force contours (6) obtained by numerical solution of Eq. (5), which describes the scanning of an AFM tip with a cluster of seven atoms, as discussed in the preceding section, on its end. The two-dimensional images presented are the cross sections of the corresponding surfaces formed by ten horizontal planes constructed with a spacing $\Delta z = (z_{max} - z_{min})/10$. A lighter tone corresponds to a higher position of the tip above the surface, and a darker tone corresponds to a lower position. In all the cases considered $z_{max} = d = 0.5a$, so that white corresponds to the height interval $(z_{max}, z_{max} - \Delta z)$. The values of z_{min} and Δz are indicated separately for each concrete case.

Figure 4 presents the results of the calculation of $z(x,y,\varphi)$ for an unperturbed surface of close-packed atoms with different orientations of the cluster on the tip end relative to the crystallographic axes of the surface. Figure 4a corresponds to the case of coinciding axes of the atomic hexagons of the cluster and the surface of the sample

($\varphi=0$). In this case $z_{min}=0.11a$, and $\Delta z=0.039a$. Figure 4b corresponds to the case $\varphi=30°$ ($z_{min}=0.28a$, $\Delta z=0.022a$). As is seen from a comparison of Fig. 4a and 4b, the results depend weakly on the orientation of the tip in the regions directly above the atoms of the surface under investigation. This situation is perfectly natural, since in the regions directly above the atoms the force contour can also be faithfully described using the model of a monatomic tip. An appreciable dependence on the orientation is observed when a cluster tip scans an interatomic space. For example, as φ is varied by 30° the relief depth varies by 1.7 fold.

Let us now examine the force contour near several types of point defects in a close-packed lattice. As an example, Fig. 5 presents the results of the calculations for a vacancy and two tip orientations: a — $\varphi=0$ ($z_{min}=0.07a$, $\Delta z=0.043a$), b — $\varphi=15°$ ($z_{min}=0.2a$, $\Delta z=0.03a$). In this case, too, a comparison of the figures reveals considerable variation of the relief depth and even variation of the character of the force contour above a vacancy as the tip orientation is varied.

Finally, Fig. 6 presents the results of calculations of force contours for a defect in the form of a divacancy for different tip orientations (the angle φ is measured from the axis of the divacancy in this case). Figures 6a and 6b, respectively, present three- and two-dimensional images of the force contour for $\varphi=0$ ($z_{min}=0.09a$, $\Delta z=0.041a$). Figure 6c presents a two-dimensional image of the force contour for $\varphi=15°$ ($z_{min}=0.19a$, $\Delta z=0.041a$). It is easy to see that as the tip orientation is varied in this case, not only does the relief depth vary, but also the symmetry of the force contour about the axis of the divacancy is broken.

We note that the averaged force contours (7) are always symmetric about all the symmetry axes of the object under investigation.[20,21] Thus a comparison with experiment of the calculated force contours for a fixed orientation (like those obtained in this section) and those obtained as a result of averaging over different orientations[20,21] creates possibilities for diagnosis of the atomic structure of the end of the AFM tip and the conditions of its scanning of the surface under investigation.

4. DISCUSSION OF RESULTS

As was shown in Sec. 1 of this paper, the force contours above a lattice of close-packed atoms are, generally speaking, not continuous and have discontinuities in the regions between the lattice atoms. This effect occurs only when the initial scanning heights of a monatomic tip are sufficiently small ($d<0.61a$, where a is the value of the equilibrium distance between the lattice atoms) and corresponds to a thrusting of the AFM tip into the surface in the region of a discontinuity. The existence of discontinuities of the force contours above the surface of a solid is not in itself surprising. They make possible such phenomena as, for example, the absorption of low-energy atoms by the surface.

The conclusion that there are discontinuities on force contours was drawn in the present work without allowance for relaxation, i.e., the displacements of the surface and tip atoms during scanning. It can easily be seen, however, that
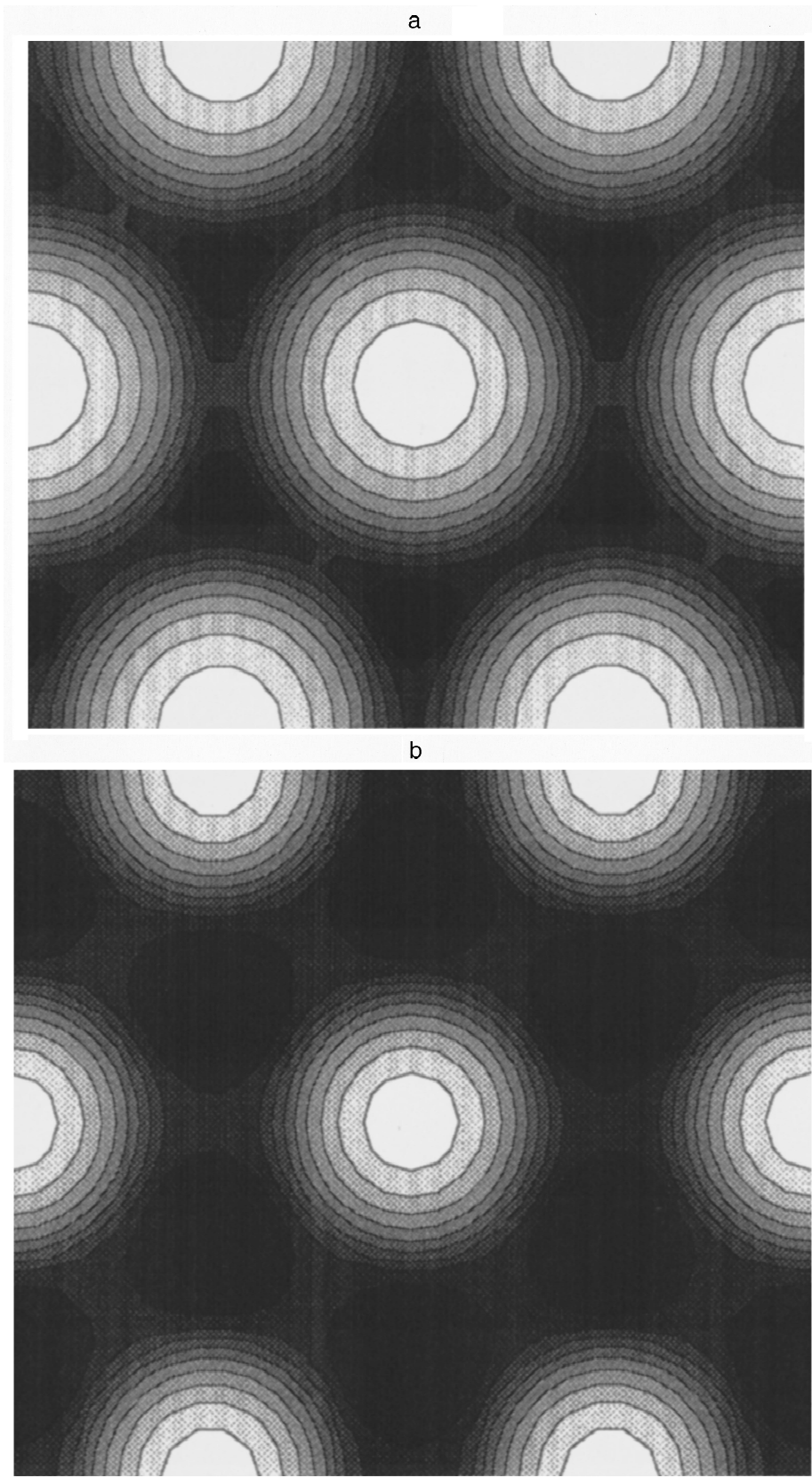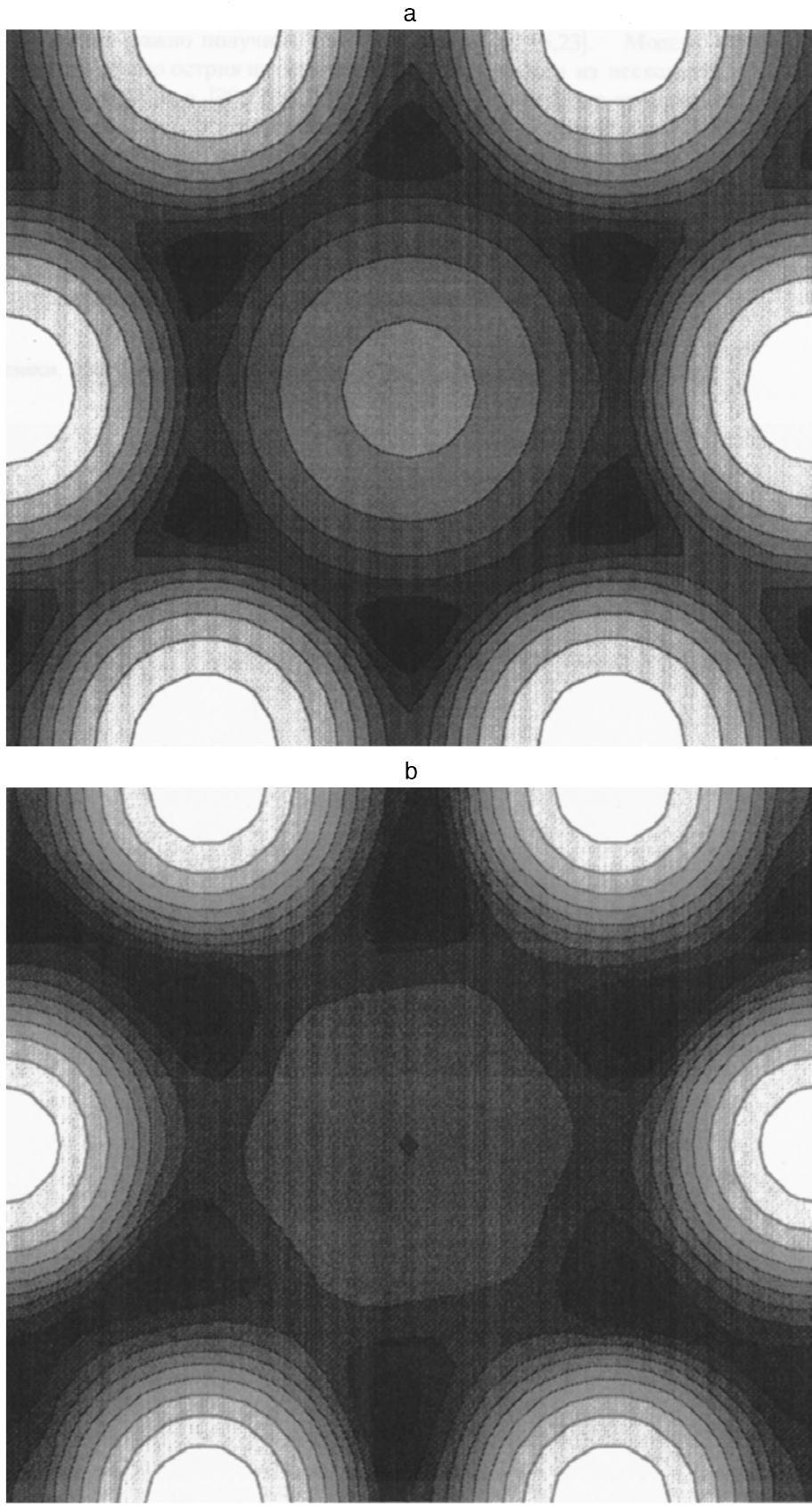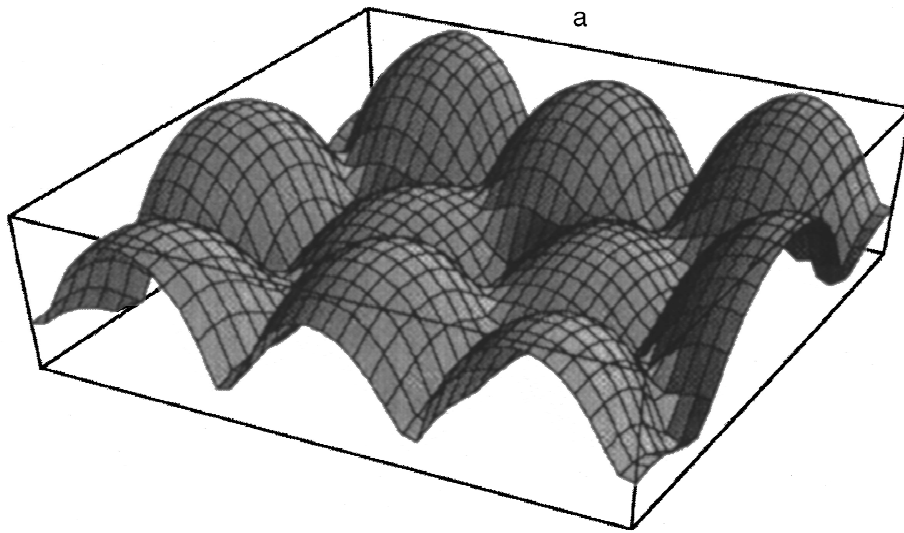
FIG. 4. Two-dimensional image of the force contour obtained by scanning over the surface of a lattice of close-packed atoms for various tip orientations.

taking the relaxation into account would not alter the conclusion that discontinuities exist. In fact, the vertical relaxation (i.e., the depression of surface atoms into the sample by the tip located above them) can be estimated in a first approximation at each fixed surface point and does not affect the description of the force interaction between the tip and sample atoms. Thus, taking it into account cannot eliminate discontinuities on the force contour if they appear in the absence of vertical relaxation. Taking into account the horizontal relaxation can, clearly, lead only to an increase in the

a

b

FIG. 5. Two-dimensional image of the force contour obtained by scanning near a surface vacancy in a lattice of close-packed atoms for various tip orientations.

initial heights of the tip above the sample at which the discontinuities on the force contour appear.

In Sec. 3 of this paper we proposed a cluster model of the tip end, which ensures continuous scanning at any initial heights above the surface of a lattice of close-packed atoms. The corresponding cluster contains seven atom, six of which form a regular hexagon, while the seventh is located at a definite distance beneath its center. The continuity of the
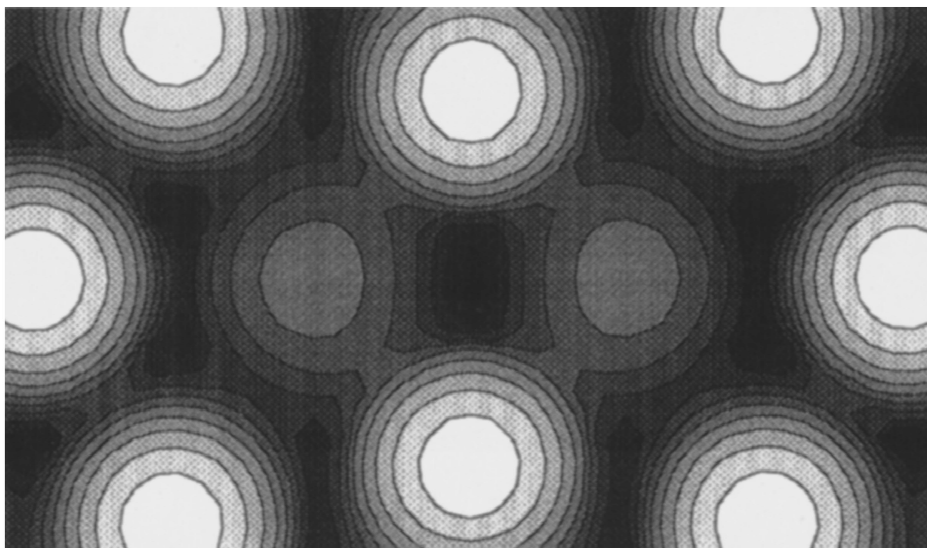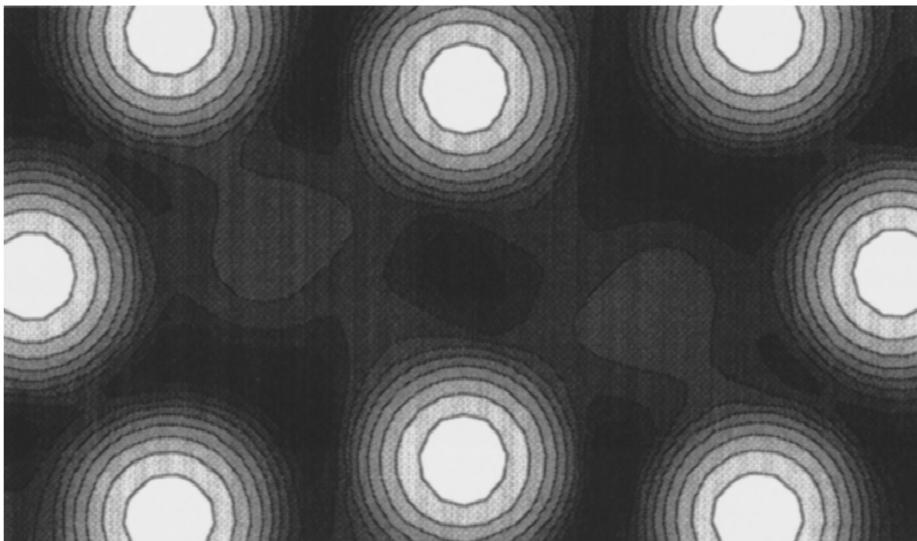
FIG. 6. Image of the force contour obtained by scanning near a surface divacancy in a lattice of close-packed atoms for various tip orientations.

scanning of such a tip is observed not only over a regular lattice, but also over a lattice with point defects in the form of vacancies and divacancies. However, the images of the crystal surfaces investigated can be less than perfectly symmetric and depend on the orientation of the cluster on the tip end relative to the crystallographic axes of the lattice surface.

The circumstances noted in principle permit diagnosis of the structure of a tip on the basis of a set of experimental AFM images of a single sample with a close-packed lattice. If the tip dips into the surface during scanning with an initial height $d < 0.61a$, it can be concluded that a single atom is found on the tip end, whose interaction makes a decisive contribution to the force interaction with the surface. If the scanning at any height is found to be continuous, the tip end has a cluster structure, the minimum parameters of which were determined in this work. In the latter case the AFM images of the surface under investigation can be independent of the initial orientation of the tip relative to the surface. This means that the orientation of the cluster on the tip end relative to the crystallographic axis of the surface varies randomly during the scanning. If the AFM images of a surface depend on the initial tip orientation, it means that the orientation remains constant during the scanning. Then the concrete value of the orientation φ between the axes of the atomic hexagons of the cluster on the tip end and the surface can be established from the nature of the AFM images of the surface under investigation.

[1] G. Binning, C. F. Quate, and C. Gerber, Phys. Rev. Lett. **56**, 930 (1986).

[2] D. Sarid, *Scanning Force Microscopy with Applications to Electric, Magnetic, and Atomic Forces*, New York (1991).

[3] O. Marti and A. Amrein, *STM and SFM in Biology*, San Diego (1993).

[4] R. Wiesendanger and H. J. Guentherodt (eds.), *Scanning Tunneling Microscopy: Theory of STM and Related Scanning Probe Methods*, Berlin (1993).

[5] C. F. Quate, Surf. Sci. **299/300**, 980 (1994).

[6] Yu. N. Moiseev, V. M. Mostepanenko, V. I. Panov, and I. Yu. Sokolov, Pis'ma Zh. Tekh. Fiz. **15**(20), 5 (1989) [Sov. Tech. Phys. Lett. **15**, 789 (1989)].

[7] T. R. Albrecht and C. F. Quate, J. Appl. Phys. **62**, 2599 (1987).

[8] O. Marti, B. Drake, and P. R. Hansma, Appl. Phys. Lett. **51**, 484 (1987).

[9] T. R. Albrecht and C. F. Quate, J. Vac. Sci. Technol. A **6**, 271 (1988).

[10] Yu. N. Moiseev, V. I. Panov, and S. V. Savinov, Pis'ma Zh. Tekh. Fiz. **17**(10), 24 (1991) [Sov. Tech. Phys. Lett. **17**, 360 (1991)].

[11] Yu. N. Moiseev, V. M. Mostepanenko, V. I. Panov, and I. Yu. Sokolov, Phys. Lett. A **132**, 354 (1983).

[12] S. A. C. Gould, K. Burke, and P. K. Hansma, Phys. Rev. B **40**, 5363 (1989).

[13] F. F. Abraham and I. P. Batra, Surf. Sci. **209**, L125 (1989).

[14] U. Landman, W. D. Luedtke, and A. Nitzan, Surf. Sci. **210**, L177 (1989).

[15] F. F. Abraham, I. P. Batra, and S. Ciraci, Phys. Rev. Lett. **60**, 1314 (1988).

[16] C. Odin, J. P. Aimé, Z. El. Kaakour, and T. Bouhacina, Surf. Sci. **317**, 321 (1994).

[17] M. Bordag, G. Klimchitskaya, and V. Mostepanenko, Surf. Sci. **328**, 129 (1995).

[18] U. Landman, W. D. Luedtke, and M. W. Ribarsky, J. Vac. Sci. Technol. A **7**, 2829 (1989).

[19] E. V. Blagov, G. L. Klimchitskaya, V. M. Mostepanenko *et al.*, Pis'ma Zh. Tekh. Fiz. **19**(8), 73 (1993) [Tech. Phys. Lett. **19**, 254 (1993)].

[20] E. V. Blagov, G. L. Klimchitskaya, A. A. Lobashev, and V. M. Mostepanenko, Pis'ma Zh. Tekh. Fiz. **21** (3), 73 (1995) [Tech. Phys. Lett. **21**, 121 (1995)].

[21] E. Blagov, G. Klimchitskaya, A. Lobashov, and V. Mostepanenko, Preprint No. FL-030495, Friedmann Laboratory Publishing, St. Petersburg (1995). p. 29.

[22] J. M. Israelachvili, *Intermolecular and Surface Forces*, Academic Press, New York–London (1985).

[23] H. Tang, C. Joachim, and J. Devillers, Surf. Sci. **291**, 439 (1993).

[24] O. Marti, J. Colchero, and J. Mlynek, in *Nanosources and Manipulations of Atoms under High Fields and Temperatures. NATO ASI Series. Series. E: Applied. Science*, Vol. 235 (1993), pp. 253–270.

[25] L. Montelius and J. O. Togenfeldt, Appl. Phys. Lett. **62**, 2628 (1993).

[26] I. M. Torrens, *Interatomic Potentials*, Academic Press, New York–London (1972).

Translated by P. Shelnitz

Influence of the instrumental functions of electrostatic and magnetic analyzers on the processing of experimental data

V. A. Kurnaev and V. A. Urusov

Moscow State Engineering Physics Institute, 115409 Moscow, Russia

Equations relating the output signal of a dispersion analyzer and the energy distribution function of the charged particles entering it are obtained and solved on the basis of an analysis of the motion of charged particles in such analyzers. The influence of corrections on the reconstruction of the energy distribution in comparison with the standard procedure is considered. © *1997 American Institute of Physics.* [S1063-7842(97)01806-0]

INTRODUCTION

Many modern methods for investigating solids and plasmas are based on analysis of the energy spectra of charged particles. Electrostatic and magnetic analyzers are employed in such investigations. The fact that the output signal of an analyzer conveys the shape of the energy spectrum of the particles with distortions raises the problem of reconstructing the true spectrum of the particles from the output signal of the analyzer.

The problem of reconstructing the true distribution for electrostatic analyzers was reduced in several papers (see, for example, Ref. 1) to solving the convolution integral equation

$$I(W) = C \int_0^{+\infty} A(W-E) f(E) dE, \quad (1)$$

where $I(W)$ is the output signal of the analyzer, $f(E)$ is the energy distribution function of the particles, $A(W-E)$ is the instrumental function of the analyzer, W is the tuning energy of the analyzer, and C is a constant.

When the energy distribution (or the momentum distribution for a magnetic analyzer) is reconstructed, the output signal of the analyzer $I(W)$ [or $I(p)$] is divided by W (or p, respectively).[2] The problem of reconstructing the true energy distribution was reduced in Ref. 3 on the basis of an approximation of the experimental data for a concrete analyzer to the solution of an integral equation of the form

$$I(W) = C \int_0^{+\infty} A(W/E) f(E) dE. \quad (2)$$

A general solution of this equation was obtained in integral form, and it was also shown that the approximate solution of Eq. (2) for a broad spectrum is obtained by dividing the output signal $I(W)$ of the analyzer by W. In Ref. 4 an equation similar to Eq. (2) was also used on the basis of an approximation of experimental data, and a solution was obtained in the form of a series. However, the question of when the equations of the former and latter types should be used to treat the experimental data remained open in the general case. The purpose of the present work is to attempt to reconstruct the true distribution from the analyzer output signal in the general case.

EQUATIONS OF TRAJECTORIES

To obtain the equations of the trajectories of the particles in an analyzer we use the approach described in Ref. 5, expressing the particle velocity $\boldsymbol{\nu}$ in terms of the radius vector of the particle $\mathbf{R}$ and the coordinate S coinciding with its trajectory

$$\boldsymbol{\nu} = \frac{d\mathbf{R}}{dt} = \frac{d\mathbf{R}}{dS}\frac{dS}{dt} = \frac{d\mathbf{R}}{dS}\,\nu, \quad (3)$$

where ν is the absolute velocity of the particle.

Then, taking into account the equation of motion of a nonrelativistic charged particle in an electrostatic field

$$m\frac{d\boldsymbol{\nu}}{dt} = -qe\nabla U, \quad (4)$$

where U is the electric field potential, and expressing the kinetic energy in terms of the total energy E_0, after some relatively simple transformations we obtain

$$2\frac{d^2\mathbf{R}}{dS^2}\left(1 - \frac{qeU}{E_0}\right) - \frac{d\mathbf{R}}{dS}\frac{d(qeU/E_0)}{dS} = -\nabla\left(\frac{qeU}{E_0}\right). \quad (5)$$

Similarly, in the case of a relativistic charged particle in a constant magnetic field, if we substitute Eq. (3) into the equation of motion

$$\frac{d}{dt}\left(\frac{m\boldsymbol{\nu}}{\sqrt{1-\nu^2/c^2}}\right) = \frac{qe}{c}(\boldsymbol{\nu}\times\mathbf{H}), \quad (6)$$

take into account that the kinetic energy of the particles does not vary in the magnetic field ($d\nu/dS = 0$), and introduce the vector $\mathbf{h} = \mathbf{H}/H$, for the path equation we obtain

$$\frac{d^2\mathbf{R}}{dS^2} = \frac{qeH}{pc}\left(\frac{d\mathbf{R}}{dS}\times\mathbf{h}\right), \quad (7)$$

where $p = m\nu/\sqrt{1-\nu^2/c^2}$ is the magnitude of the momentum.

INSTRUMENTAL FUNCTION OF AN ANALYZER AND TREATMENT OF MEASUREMENT DATA

Let us consider the motion of charged particles in an electrostatic analyzer. Let a particle enter the analyzer at a point with the coordinates (η, ξ) (the system of coordinates

coincides with the plane of the entrance electrode) in the direction assigned by the angles α and β with an energy $E=E_0/e$, where e is an elementary charge. The distribution of the field in the analyzer is assigned by the potentials U_i, where $i=1,\ldots,n$, on the electrodes relative to the entrance electrode, which is at zero potential. At the nth or exit electrode of the analyzer the particle is at a point with the coordinates (η_1,ξ_1) (the coordinate system lies in the plane of the exit electrode).

The coordinates (η_1,ξ_1) are found by solving the equation of motion (3). Since the trajectory of the particle [Eq. (5)] remains unchanged as its energy and potential vary over the entire space by the factor L, the coordinates (η_1,ξ_1) remain unchanged when E and U_i, where $i=1,\ldots,n$, change simultaneously by the factor L. This condition holds only if the energy and the potentials appear in the functions of the coordinates (η_1,ξ_1) in the form of a ratio. Similarly, it can be shown that the charge and the energy also appear in the form of a ratio.

To find the relation between the output signal of the analyzer and the energy distribution function of the particles, we use the method described in Ref. 6. If the particles at the entrance to the analyzer have a distribution function with respect to the coordinates of the cross section of the beam formed by the surface of the entrance diaphragm, the angles, and the energy $f(\eta,\xi,\alpha,\beta,E)$, the number of particles which have an energy in the range from E to $E+dE$ and emerge in the direction assigned by the angles α and β into a solid-angle element $d\Omega$ from an element dS_0 of the entrance diaphragm area per unit time equals

$$d^3I=I_0f(\eta,\xi,\alpha,\beta,E)d\Omega dS_0dE, \tag{8}$$

where

$$I_0=\int\int\int\frac{d^3I}{dEdS_0d\Omega}d\Omega dS_0dE$$

is the number of particles passing through the hole in the entrance diaphragm per unit time.

To find the number of particles I passing through the hole in the exit diaphragm per unit time, Eq. (8) must be integrated over all the trajectories passing through the hole in the exit electrode. For this purpose we express the angles α and β in terms of the coordinates (η_1,ξ_1) of the coordinate system of the exit diaphragm

$$\alpha=\alpha\left(\eta,\xi,\eta_1,\xi_1,\frac{qU_1}{E},\ldots,\frac{qU_n}{E}\right),$$

$$\beta=\beta\left(\eta,\xi,\eta_1,\xi_1,\frac{qU_1}{E},\ldots,\frac{qU_n}{E}\right), \tag{9}$$

and substituting these expressions into (8), we integrate over the energy and the areas of the entrance and exit apertures

$$I=I_0\int_0^{+\infty}\int_{S_0}\int_{S_1}f\left(\eta,\xi,\eta_1,\xi_1,\frac{qU_1}{E},\ldots,\frac{qU_n}{E},E\right)$$

$$\times J(\alpha,\beta,\eta_1,\xi_1)\sin\alpha\,dS_1dS_0dE, \tag{10}$$

where

$$J\left(\alpha,\beta,\eta_1,\xi_1,\eta,\xi,\frac{qU_1}{E},\ldots,\frac{qU_n}{E}\right)=\frac{\partial(\alpha,\beta)}{\partial(\eta_1,\xi_1)}$$

is the Jacobian of the transformation .

We assume that the energy distribution in the beam being analyzed does not depend on the distribution with respect to the cross section and the angles

$$f(\eta,\xi,\alpha,\beta,E)=f_1(\eta,\xi,\alpha,\beta)f_2(E). \tag{11}$$

Using such a function, we can represent expression (10) in the form

$$I=I_0\int_0^{+\infty}A\left(\frac{qU_1}{E},\ldots,\frac{qU_n}{E}\right)f_2(E)dE, \tag{12}$$

where

$$A\left(\frac{qU_1}{E},\ldots,\frac{qU_n}{E}\right)$$

$$=\int_{S_0}\int_{S_1}f_1\left(\eta,\xi,\eta_1,\xi_1,\frac{qU_1}{E},\ldots,\frac{qU_n}{E}\right)$$

$$\times J(\alpha,\beta,\eta_1,\xi_1)\sin\alpha\,dS_1dS_0.$$

The function $A(qU_1/E_0,\ldots,qU_n/E_0)$ is the instrumental function of the analyzer, since it expresses the dependence of the output signal of the analyzer on the electrode potentials for a monoenergetic beam of particles. It should be noted that the instrumental function of the analyzer will be a function of the ratios of the electrode potentials to the particle energy, even if the condition of a one-to-one correspondence between the angles α and β at the entrance to the analyzer and the exit coordinates (η_1,ξ_1) is not satisfied, since the instrumental function is the integral over all the trajectories passing through the hole in the exit diaphragm, and each trajectory in the analyzer is a function of the ratios of the electrode potentials to the particle energy, for example, when the beam is focused on a point.

It is not difficult to show that convolution equation (1) is incompatible with Eq. (12) and is, thus, inapplicable to analyzers operating in the spectrometer regime. In fact, if it is assumed that the instrumental function simultaneously satisfies (1) and (12), the condition

$$W-E=F(qU_1/E,\ldots,qU_n/E) \tag{13}$$

must be satisfied, i.e., the difference between the tuning energy of the analyzer and the energy of the particles must be a function of the ratios between the electrode potentials and the energy. On the other hand, the tuning energy W of an electrostatic analyzer should not depend on the energy E with which the particle enters the analyzer, in contradiction with condition (13).

In order that the instrumental function of the analyzer would have the form $A(W/E)$, when the energy of the particles changes by a factor of L, the electric fields must vary by the same factor over the entire trajectory of the particles. This condition is strictly satisfied only if the similarity condition for the electric field holds in the entire space of the analyzer. Therefore, in this case the potentials on the electrodes should be linearly related:
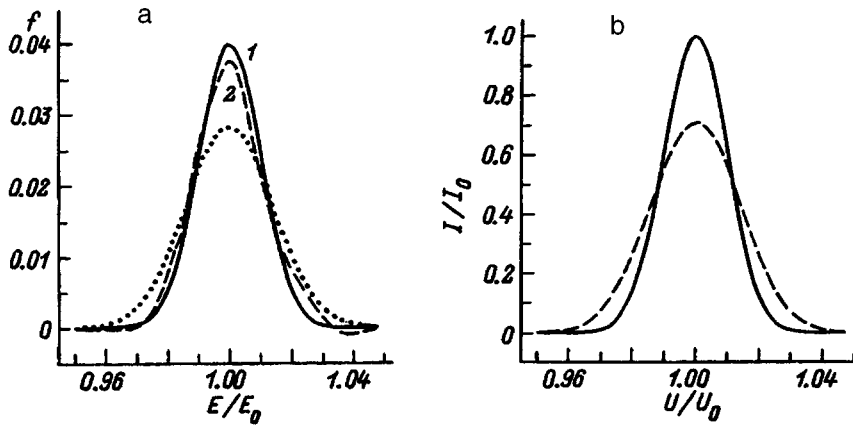
FIG. 1. Results of the reconstruction of an energy distribution. a: solid curve — true distribution; dotted curve — distribution obtained by dividing the output signal by the energy; dashed curve — distribution [from Eq. (20) to within a correction associated with the second derivative] based on the known output signal (dashed curve in Fig. 1b) and the instrumental function (solid curve in Fig. 1b) (the full width of the true energy distribution at half-maximum is equal to the full width of the instrumental function at half-maximum).

$$\frac{U_2}{U_1}=\lambda_2;\ \ldots;\ \frac{U_n}{U_1}=\lambda_n, \tag{14}$$

where $\lambda_2,\ \ldots,\lambda_n$ are constants.

Then the tuning energy and the potential on one of the electrodes, in terms of which the current of the particles at the entrance to the analyzer is measured, will be related by the expression

$$W=kU_1, \tag{15}$$

where $k=$ const is the analyzer constant.

In this case the relation between the current at the analyzer exit and the energy distribution function of the particles is described by the equation

$$I(U_1)=I_0\int_0^{+\infty}A\left(\frac{qU_1}{E},\lambda_2,\ \ldots,\lambda_n\right)f(E)dE. \tag{16}$$

The solution of Eq. (16) for the energy distribution function of the particles can be found in an integral form using the Mellin transform[3]

$$f(k\breve{U})=\frac{1}{I_0}\frac{1}{2\pi i}\int_{x_0-i\cdot\infty}^{x_0+i\cdot\infty}\frac{I_{x-1}}{A_{x-1}}(k\breve{U})^{-x}dx, \tag{17}$$

where

$$I_{x-1}=\int_0^{+\infty}I(U_1)U_1^{x-2}dU_1,$$

$$A_{x-1}=\int_0^{+\infty}A\left(\frac{qU_1}{E}\right)\left(\frac{U_1}{E}\right)^{x-2}d(U_1/E). \tag{18}$$

The expression obtained is not convenient for practical use; therefore, we find the solution of the equation in the form of a series. For this purpose, we assume that $I(U_1)$ is infinitely differentiable and can be expanded into a Taylor series in the vicinity of the point $\breve{U}$:

$$I(U_1)=\sum_{n=0}^{+\infty}\frac{I^{(n)}(\breve{U})(U_1-\breve{U})^n}{n!}. \tag{19}$$

Substituting expression (19) into integral (18) and assuming, for simplicity, that $\breve{U}>0$ and $U_1>0$, for Eq. (17) we obtain

$$f(k\breve{U})=\frac{1}{I_0}\sum_{n=0}^{+\infty}I^{(n)}(\breve{U})\breve{U}^{n-1}\frac{1}{2\pi in!}$$

$$\times\int_0^{+\infty}\int_{x_0-i\cdot\infty}^{x_0+i\cdot\infty}\frac{(U_1/\breve{U}-1)^n}{A_{x-1}k^2}\left(\frac{U_1}{kU}\right)^{x-2}dxd(U_1/\breve{U})$$

or

$$f(k\breve{U})=\frac{1}{I_0}\sum_{n=0}^{+\infty}B_nI^{(n)}(\breve{U})\breve{U}^{n-1}, \tag{20}$$

where $I^{(n)}(\breve{U})$ is the nth derivative of the current at the analyzer exit with respect to $\breve{U}$, and the B_n are constants.

We express the coefficients B_n in terms of the moments of the instrumental function. For this purpose, after expanding $I^{(n)}(\breve{U})$ into a Taylor series and substituting expression (20) into Eq. (16), we obtain the equation

$$I(U_1)=\sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty}\frac{B_nC_{nm}}{m!k^{n+m-1}}I^{(n+m)}(U_1)U_1^{n+m}, \tag{21}$$

where

$$C_{nm}=\int_0^{+\infty}z^{n-1}(z-k)^mA(q/z)dz. \tag{22}$$

Since Eq. (21) holds for any function $I(U_1)$, the coefficient in front of $I^{(n+m)}(U_1)$ for $n=0$ and $m=0$ equals unity, and the sum of the coefficients in front of the remaining derivatives equals zero. Then

$$B_0=\frac{1}{kC_{00}},\quad B_n=-\frac{1}{C_{n0}}\sum_{i=0}^{n-1}B_i\frac{C_{i(n-i)}}{(n-i)!}. \tag{23}$$

The analyzer constant can be expressed so as to satisfy the condition $B_1=0$; in that case it equals

$$k=\frac{C_{10}}{C_{00}}. \tag{24}$$

Therefore, the correction associated with the first derivative $I'(U_1)$ can be eliminated by adjusting the analyzer constant. This allows us to assume, in contrast to the results obtained in Ref. 3, that the first derivative of the current has little influence on the shape of the energy distribution. Figure
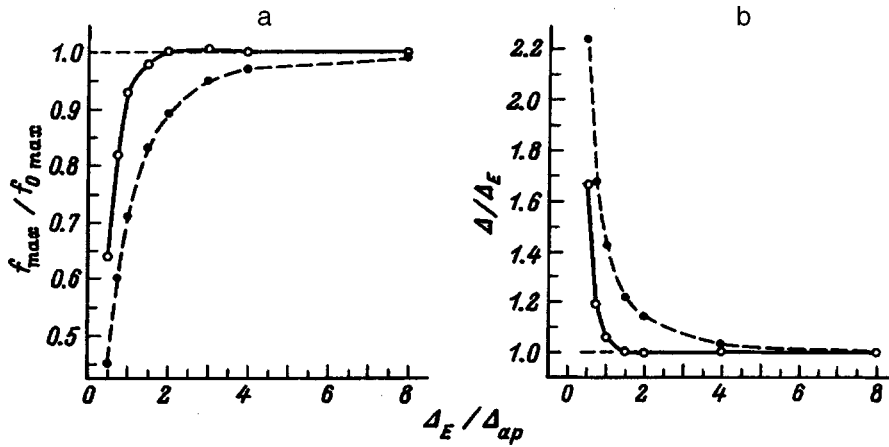
FIG. 2. Dependence of f_{max}/f_{0max} and Δ/Δ_E on Δ_E/Δ_{instr}, where f_{max} is the height of the reconstructed spectrum, f_{0max} is the height of the true distribution, Δ is the full width of the reconstructed spectrum at half-maximum, Δ_E width of the true distribution, and Δ_{instr} is the width of the instrumental function: solid curve – with consideration of the correction associated with the second derivative; dashed curve – for spectra obtained by simple division by the energy.

1 shows examples of the processing of the spectrum in a first approximation and with the correction associated with the second derivative of the current. The normal Gaussian distribution was employed as a trial function for the true distribution function. It can be concluded on the basis of the results obtained (Fig. 2) that the correction associated with the second derivative makes a significant contribution to the distribution function in the case in which the width of the true distribution function is of the order of the width of the instrumental function. In this case the energy distribution obtained with the correction is considerably closer to the true distribution than are the distributions obtained without it. In addition, it is not difficult to show that the distribution function thus obtained satisfies the normalization condition.

Using the normalization condition for $f(E)$, it is not difficult to show that $I(U_1)$ tends to zero when $U_1 \to 0$ and $U_1 \to \infty$. Taking this into account, we find that

$$\int_0^{+\infty} I^{(n)}(U_1)U_1^{n-1}dU_1 = 0, \tag{25}$$

where $n \geq 1$.

Expression (25) makes it possible to compare the intensities of two currents with quasimonoenergetic energy distribution functions of the particles having a width of the order of the width of the instrumental function without refining the form of the distribution function. It is noteworthy that a simple comparison of the current maxima at the analyzer exit gives an incorrect result:

$$\frac{I_{10}}{I_{20}} = \frac{\int_0^{+\infty} I_1(U)/U \, dU}{\int_0^{+\infty} I_2(U)/U \, dU} \neq \frac{I_{1max}}{I_{2max}}, \tag{26}$$

where I_{10} and I_{20} are the particle currents at the entrance to the analyzer, and I_{1max} and I_{2max} are the maximum values of the particle currents $I_1(U)$ and $I_2(U)$ at the analyzer exit.

Let us now consider a magnetic analyzer of charged particles operating in the spectrometer regime. Assuming that the magnetic field in the analyzer is created by magneto-optical elements and that the strengths of fields created by these elements are linearly related, for a particle which has the coordinates (η, ξ) at the entrance diaphragm of the analyzer and emerges in the direction assigned by the angles α

and β with the momentum p, from the equations of motion we obtain the coordinates of the particle at the exit diaphragm of the analyzer

$$\eta_1 = \eta_1(\eta, \xi, \alpha, \beta, qH/p), \quad \xi_1 = \xi_1(\eta, \xi, \alpha, \beta, qH/p), \tag{27}$$

where H is the magnetic field strength at an arbitrarily selected fixed point.

The distribution function of the particles with respect to the magnitude of the momentum, which does not depend on the distribution with respect to the angles and the cross section, and the number of particles passing through the hole in the entrance diaphragm per unit time are related by an expression similar to the expression for electrostatic analyzers

$$I(H) = I_0 \int_0^{+\infty} A(qH/p)f_2(p)dp. \tag{28}$$

The momentum corresponding to tuning of the analyzer and the magnetic field strength H are related by the expression

$$p_1 = kH. \tag{29}$$

For a distribution function whose value varies weakly across the width of the instrumental function, the approximate solution of Eq. (28) will have the form

$$f_2(kH) \approx \frac{I(H)}{CI_0H}, \tag{30}$$

where $C = \int_0^{+\infty} A(q/z)dz$ is a constant and $z = p/H$.

CONCLUSIONS

Let us briefly review the main results of this work.

1. It has been shown in this work that for all electrostatic analyzers operating in the spectrometer mode, the energy distribution function of the charged particles at the entrance to the analyzer and the number of particles passing through the aperture in the exit electrode per unit time depend on the ratio between the electrode potentials and the particle energy and are related by Eq. (12). It should be noted that the expressions (12) and (16) obtained above for describing the relationship between the current at the analyzer exit and the

energy distribution function of the particles remain valid when the stray electrostatic fields caused by the actual geometry of the electrodes are taken into account.

2. A detailed analysis shows that the convolution equation (1) previously proposed[1] for describing the relationship between the particle current at the analyzer exit and the energy distribution function of the particles is not applicable to an analyzer operating in the spectrometer mode.

3. Under the condition of a linear relationship between the potentials on the analyzer electrodes, the particle current at the analyzer exit and the energy distribution function of the particles are related by Eq. (2), in which the instrumental function is a function of the ratio of the tuning energy W to the particle energy E. For a magnetic analyzer operating in the spectrometer mode, the momentum distribution function of the particles and the particle current at the analyzer exit are related by an analogous equation (28) provided the fields created by the magnetooptical elements are linearly related.

4. Solution (20) in the form of a series in derivatives of the current at the analyzer exit, which was obtained for an arbitrary continuous energy distribution function of the particles, and recurrence relations (23) for the coefficients in the series permit the reconstruction of energy spectra with a width of the order of the width of the instrumental function.

5. Since the distribution function of the beam with respect to the coordinates of the cross section formed by the entrance diaphragm and the angles appears in expression (12) for the instrumental function, the monoenergetic particle beam used to calibrate an analyzer must have a distribution function with respect to the cross section and angles that is close to the spectra which are to be measured by the analyzer. For example, if the analyzer is intended for investigating particles reflected or emitted from a surface and the surface area ''visible'' to the analyzer is smaller than the emission area, and the angular distribution function of the particles varies weakly within the angular aperture of the analyzer, it is best to employ a broad monoenergetic particle beam with an angular distribution that is quasi-isotropic within the angular aperture of the analyzer to calibrate the latter.

6. Preliminary retardation of the particles is often used to improve the resolving power of a dispersion analyzer. As a rule, analyzers with preliminary retardation operate in two regimes:[1] in one regime the retarding potential remains constant, and the spectrum is scanned by varying the potential on the deflecting electrodes. One deficiency of this regime is that it is unsuitable for treating the results of measurements of broad spectra. In the other regime the potential difference on the deflecting electrodes remains constant, and the scanning is performed by varying the retarding potential. One deficiency of this regime is that the angular distribution function of the particles varies after the retardation system, making the ensuing treatment of the spectra difficult. The deficiencies just enumerated can be avoided by using an analyzer operating in a regime in which the retarding potential and the potentials on the deflecting electrodes are linearly related by Eq. (14). The use of a retarding potential permits improvement of the resolving power of the analyzer. On the other hand, the particle current at the analyzer exit and the energy distribution function are related by Eq. (16), and thus the treatment of the spectrum reduces, in a first approximation, to division of the signal by the energy.

[1] V. P. Afanas'ev and S. Ya. Yavor, *Electrostatic Analyzer for Charged Particle Beams* [in Russian], Nauka, Moscow (1978).
[2] K. Siegbahn (ed.), *Beta and Gamma-Ray Spectroscopy*, Interscience, New York (1955) [Russ. transl., Fizmatgiz, Moscow (1959), pp. 63–106].
[3] G. I. Zhabrev and S. K. Zhdanov, Pis'ma Zh. Tekh. Fiz. **49**, 2450 (1979) [Sov. Tech. Phys. Lett. **49**, 1377 (1979)].
[4] N. N. Koborov, V. A. Kurnaev, and V. A. Urusov, *Interactions of Ions and Plasmas with Solid Surfaces* [in Russian] (1986), pp. 22–31.
[5] A. A. Sysoev and M. S. Chupakhin, *Introduction to Mass Spectrometry* [in Russian], Atomizdat, Moscow (1977).
[6] V. P. Afanas'ev and S. Ya. Yavor, Pis'ma Zh. Tekh. Fiz. **1**, 227 (1975) [Sov. Tech. Phys. Lett. **1**, 109 (1975)].

Translated by P. Shelnitz

Influence of potential fluctuations on the instrumental functions of electrostatic analyzers

V. A. Kurnaev and V. A. Urusov

Moscow State Engineering-Physics Institute, 115409 Moscow, Russia

An equation which relates the output signal of an electrostatic dispersion analyzer and the energy distribution function of the charged particles entering it is derived with the fluctuations of the potentials on the defecting electrodes taken into account. Solutions of this equation are obtained. The influence of noise on the instrumental functions of analyzers is considered. © *1997 American Institute of Physics.* [S1063-7842(97)01906-5]

INTRODUCTION

It was shown in the preceding paper[1] that the energy distribution function of the charged particles at the entrance to an electrostatic analyzer and the output signal of the analyzer are related by the expression

$$I(U_1, \ldots, U_n) = \int_0^{+\infty} f(E)A(U_1/E, \ldots, U_n/E)dE, \quad (1)$$

where $A(U_1/E, \ldots, U_n/E)$ is the instrumental function of the analyzer; $U_1, \ldots, U_n$ are the potentials on the analyzer electrodes; and $f(E)$ is the energy distribution function of the charged particles.

Relation (1) also holds when the stray fields arising from the actual geometry of the analyzer are taken into account. However, the form of the instrumental function of a charged-particle analyzer is determined not only by its geometry, but also by the fluctuations of the fields within the analyzer, i.e., by the noise. If fluctuations whose characteristic period is much greater than the time of flight of a particle in the analyzer are considered, it can be assumed that the tuning energy of the analyzer fluctuates.

In this paper we shall consider the influence of noise on the instrumental function of an electrostatic dispersion analyzer.

INFLUENCE OF FLUCTUATIONS ON THE INSTRUMENTAL FUNCTION OF AN ANALYZER WITH ONE DEFLECTING ELECTRODE

Let us assume for an electrostatic analyzer, in which only one electrode is under a potential, that the voltage $\check{U}$ on the analyzer electrode fluctuates about the mean potential U with a distribution function $\varphi(\check{U}, U)$, which satisfies the following conditions

$$\int_{-\infty}^{+\infty} \varphi(\check{U}, U)d\check{U} = 1,$$

$$U = \langle \check{U} \rangle = \int_{-\infty}^{+\infty} \check{U} \cdot \varphi(\check{U}, U)dU,$$

$$\sigma^2(\check{U}) = \int_{-\infty}^{+\infty} (\check{U} - U)^2 \varphi(\check{U}, U)dU. \quad (2)$$

If the measurement time of the particle current at the analyzer exit is much greater than the characteristic period of the fluctuations, then, using Eq. (1), we can represent the mean value of the current at the analyzer exit in the form

$$I(U) = \int_{-\infty}^{+\infty} \varphi(\check{U}, U)\widetilde{I}(\check{U})d\check{U}$$

$$= I_0 \int_{-\infty}^{+\infty} \varphi(\check{U}, U)\int_0^{+\infty} f(E)A\left(\frac{q\check{U}}{E}\right)dEd\check{U} \quad (3)$$

or

$$I(U) = I_0 \int_0^{+\infty} f(E)B(U, E)dE, \quad (4)$$

where

$$B(U,E) = \int_{-\infty}^{+\infty} \varphi(\check{U}, U)A(q\check{U}/E)d\check{U} \quad (5)$$

is the instrumental function of the analyzer with consideration of the noise.

Let us find an approximate solution of this equation. Expanding $\widetilde{I}(\check{U})$ into a Taylor series about the point U, assuming that $\widetilde{I}(U) \approx I(U)$ in a first approximation, and taking into account the terms of the series containing derivatives no higher than the second, from (3) we obtain the expression

$$\widetilde{I}(U) \approx I(U) - \frac{\sigma^2}{2}\frac{d^2 I}{dU^2}. \quad (6)$$

A solution of the equation

$$\widetilde{I}(U) = I_0 \int_0^{+\infty} f(E)A\left(\frac{qU}{E}\right)dE \quad (7)$$

can be obtained in the form of a series:[1]

$$f(kU) = \sum_{n=0}^{+\infty} B_n U^{n-1}\frac{d\widetilde{I}^n(U)}{dU^n}, \quad (8)$$

where

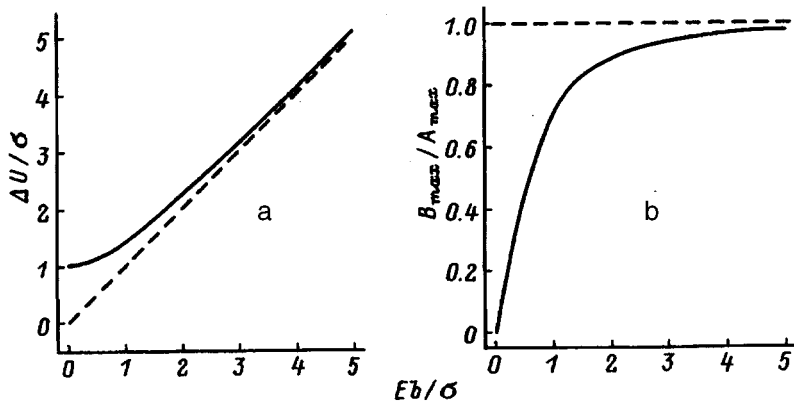$$B_0 = \frac{1}{kC_{00}}, \quad B_n = -\frac{1}{C_{n0}}\sum_{i=0}^{n} B_i \frac{C_{i(n-i)}}{(n-i)!}.$$

FIG. 1. Curves for an electrostatic dispersion analyzer: a—The ratio between the standard deviation of the instrumental function ΔU (with allowance for the influence of the fluctuations of the potential) and the standard deviation σ of the noise and, b—the ratio of the height of the instrumental function $B_{\max}$ (with consideration of the influence of the potential fluctuations) to the height of the instrumental function $A_{\max}$ (without consideration of the noise) as functions of the ratio of the standard deviation of the instrumental function (without consideration of the noise) bE to the standard deviation σ of the fluctuations of the potentials on the analyzer electrodes.

Ultimately, to within the terms with a second derivative, for the "true distribution" as a function of the output signal $I(U)$ we obtain

$$f(kU) \approx \frac{i(U)}{C_{10}I_0U} - \frac{Ud^2I/dU^2}{2C_{10}I_0}\left(1 - \frac{C_{10}^2}{C_{00}C_{20}} + \frac{\sigma^2}{U^2}\right), \quad (9)$$

where the $C_{nm} = \int_0^{+\infty} z^{n-1}(z-k)^m A(q/z)dz$ are constants.

It is expedient to select the analyzer constant such that in the range of energies where the influence of the noise is negligibly small, the analyzer constant would be determined from the expression[1]

$$k = C_{10}/C_{00}. \quad (10)$$

Let us now consider the influence of the noise on the form of the instrumental function of the analyzer. For an arbitrary noise distribution function $\varphi(\check{U},U)$, an approximate analytical expression for the instrumental function $B(U,E)$ of the analyzer can be obtained in two limiting cases: in the range of energies where the width of the instrumental function is determined mainly by the noise and in the range of energies where the influence of the noise on the instrumental function is weak.

In the former case, applying the theorem of the mean[2] to the integral in Eq. (5) and assuming that $A(q/z) \equiv 0$ for $z \leq 0$, we obtain

$$B(U,E) = C_{00}\hat{U}\varphi(\hat{U},U). \quad (11)$$

Assuming that $\check{U}\varphi(\check{U},U)$ varies weakly across the width of the instrumental function $A(q\check{U}/E)$, we have

$$\hat{U}\varphi(\hat{U},U) \approx \frac{E}{k}\varphi\left(\frac{E}{k},U\right). \quad (12)$$

Substituting this expression into (11), we ultimately obtain

$$B(U,E) \approx (C_{00}/k)E\varphi\left(\frac{E}{k},U\right), \quad (13)$$

i.e., the form of the instrumental function is determined mainly by the noise distribution function, and the value of the instrumental function at the maximum depends on the energy of the particles.

For example, for a noise distribution function of the form $\varphi(\check{U}-U)$ it follows from expression (13) that the

transmission of the analyzer will decrease as the energy of the particles decreases. We use the term transmission to refer to a coefficient of an analyzer that is equal to the ratio of the maximum output current of the analyzer to the current of a monoenergetic beam at its entrance.

In the second case, since the noise distribution function $\varphi(\check{U},U)$ is considerably narrower than the instrumental function $A(q\check{U}/E)$, applying the theorem of the mean and assuming that $A(q\hat{U}/E) \approx A(qU/E)$, we obtain

$$B(U,E) \approx A\left(\frac{qU}{E}\right), \quad (14)$$

i.e., the width of the instrumental function will increase linearly with increasing energy, and the transmission of the analyzer will remain unchanged.

Let us trace how the form of the instrumental function of the analyzer varies in the presence of noise as a function of the energy in the case of a noise distribution function of the form $\varphi(\check{U},U) = \varphi(\check{U}-U)$. For this purpose we treat the instrumental function $B(U,E)$ as a distribution function with respect to U, which can be characterized by the mean $\langle U \rangle$ and the variance $\langle \Delta U^2 \rangle$,

$$\langle U \rangle = \frac{\int_{-\infty}^{+\infty} UB(U,E)dU}{\int_{-\infty}^{+\infty} B(U,E)dE}, \quad (15)$$

$$\langle \Delta U^2 \rangle = \frac{\int_{\infty}^{+\infty}(U-\langle U \rangle)^2 B(U,E)dU}{\int_{-\infty}^{+\infty} B(U,E)dU}. \quad (16)$$

After some relatively simple mathematical manipulations, we find that the mean value of the potential is proportional to the energy:

$$\langle U \rangle = \frac{C_1}{C}E. \quad (17)$$

Here $C = \int_{-\infty}^{+\infty} A(qx)dx$, and $C_1 = \int_{-\infty}^{+\infty} xA(qx)dx$, where $x = \check{U}/E$.

We calculate the variance of the instrumental function by substituting expression (17) into Eq. (16) and performing the integration:

$$\langle \Delta U^2 \rangle = \sigma^2 + b^2 E^2. \quad (18)$$

Here $b^2 = (C_2 C - C_1^2)/C^2$, where $C_2 = \int_{-\infty}^{+\infty} x^2 A(qx)dx$.

Figure 1 shows qualitatively the behavior of the width of the instrumental function and its value at the maximum as a function of E. The expressions (13), (14), and (18) obtained allow us to conclude that in the range of energies where $E \ll \sigma/b$ the width of the instrumental function scarcely depends on the energy, and its value at the maximum increases linearly with the energy. Conversely, in the range of energies where $E \gg \sigma/b$ the width of the instrumental function increases linearly with increasing energy, and the value at the maximum remains constant. The quantity $E^2 b^2$ corresponds to the variance of the instrumental function with respect to U in the absence of potential fluctuations and can be determined experimentally.

It should be noted that a solution of Eq. (3) can be found in general form for a noise distribution function of the form under consideration $\varphi(\check{U} - U)$. Applying the Fourier transformation to Eq. (3), we obtain a solution in the form of a series[3]

$$\int_0^{+\infty} f(E) A\left(\frac{qU}{E}\right) dE = \sum_{s=0}^{+\infty} D_s \frac{d^s I(U)}{dU^s} = \widetilde{I}(U), \qquad (19)$$

where

$$D_0 \int_{-\infty}^{+\infty} \varphi(z) dz = 1, \quad \sum_{s=0}^{m} D_s \int_{-\infty}^{+\infty} z^{m-s} \varphi(z) dz \frac{1}{(m-s)!} = 0. \qquad (20)$$

A series solution of Eq. (19) was found in Ref. 1. In the present case it has the form

$$f(kU) = \sum_{n=0}^{+\infty} B_n U^{n-1} \frac{d\widetilde{I}^n(U)}{dU^n}, \qquad (21)$$

where

$$B_0 = \frac{1}{kC_{00}}, \quad B_n = -\frac{1}{C_{n0}} \sum_{i=0}^{n} B_i \frac{C_{i(n-i)}}{(n-i)!}. \qquad (22)$$

Ultimately, for the general solution we obtain

$$f(kU) = \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} D_m B_n U^{n-1} \frac{d^{m+n} I(U)}{dU^{m+n}}, \qquad (23)$$

where the analyzer constant k is specified by expression (7).

INFLUENCE OF NOISE ON THE INSTRUMENTAL FUNCTION OF AN ANALYZER WITH TWO DEFLECTING ELECTRODES

All the arguments advanced above referred to the case in which a potential was supplied to one electrode. Let us now consider the case in which potentials are supplied to two electrodes. The expression for the mean value of the current at the analyzer exit under the condition that the measurement time is much greater than the characteristic period of the fluctuations has the form

$$I(U_1, U_2) = I_0 \int \int_{-\infty}^{+\infty} \varphi(\check{U}_1, \check{U}_2, U_1, U_2)$$

$$\times \int_0^{+\infty} A\left(\frac{\check{U}_1}{E}, \frac{\check{U}_2}{E}\right) f(E) dE d\check{U}_1 d\check{U}_2$$

$$= \int \int_{-\infty}^{+\infty} \varphi(\check{U}_1, \check{U}_2, U_1, U_2) \widetilde{I}(\check{U}_1, \check{U}_2) d\check{U}_1 d\check{U}_2, \qquad (24)$$

where $\varphi(\check{U}_1, \check{U}_2, U_1, U_2)$ is the normalized distribution function of the potential fluctuations, which satisfies the conditions

$$\int \int_{-\infty}^{+\infty} \varphi(\check{U}_1, \check{U}_2, U_1, U_2) d\check{U}_1 d\check{U}_2 = 1,$$

$$\int \int_{-\infty}^{+\infty} \check{U}_1 \varphi(\check{U}_1, \check{U}_2, U_1, U_2) d\check{U}_1 d\check{U}_2 = U_1,$$

$$\int \int_{-\infty}^{+\infty} \check{U}_2 \varphi(\check{U}_1, \check{U}_2, U_1, U_2) d\check{U}_1 d\check{U}_2 = U_2,$$

$$\int \int_{-\infty}^{+\infty} (\check{U}_1 - U_1)^2 \varphi(\check{U}_1, \check{U}_2, U_1, U_2) d\check{U}_1 d\check{U}_2 = \sigma_1^2,$$

$$\int \int_{-\infty}^{+\infty} (\check{U}_2 - U_2)^2 \varphi(\check{U}_1, \check{U}_2, U_1, U_2) d\check{U}_1 d\check{U}_2 = \sigma_2^2,$$

$$\int \int_{-\infty}^{+\infty} (\check{U}_1 - U_1)(\check{U}_2 - U_2) \varphi(\check{U}_1, \check{U}_2, U_1, U_2) d\check{U}_1 d\check{U}_2$$

$$= \mathrm{cov}(\check{U}_1, \check{U}_2). \qquad (25)$$

Assuming that the function $\widetilde{I}(\check{U}_1, \check{U}_2)$ varies weakly across the width of the noise distribution function, expanding $\widetilde{I}(\check{U}_1, \check{U}_2)$ in a Taylor series about the point (U_1, U_2), assuming that $\widetilde{I}(U_1, U_2) = I(U_1, U_2)$ in a first approximation, and keeping the terms with derivatives no higher than the second order, we obtain from Eq. (24)

$$\widetilde{I}(U_1, U_2) \approx I(U_1, U_2) - \frac{1}{2}\left(\sigma_1^2 \frac{\partial^2 I}{\partial U_1^2} \right.$$

$$\left. + 2\mathrm{cov}(\check{U}_1, \check{U}_2) \frac{\partial^2 I}{\partial U_1 \partial U_2} + \sigma_2^2 \frac{\partial^2 I}{\partial U_2^2} \right). \qquad (26)$$

If the mean values of the potentials are linearly related, i.e., if $U_2 = \lambda U_1$, and if we take into account that the expression for the energy distribution function in terms of $\widetilde{I}(U_1, U_2)$ has the form[1]

$$f(kU_1) = \frac{1}{I_0} \sum_{n=0}^{\infty} B_n U_1^{n-1} \frac{d\widetilde{I}^n(U_1, \lambda U_1)}{dU_1^n}, \qquad (27)$$

the solution of Eq. (24) to terms with a second derivative will be as follows:

$$f(kU_1) \approx \frac{1}{I_0}\left(\frac{I(U_1, \lambda U_1)}{C_{10}U_1} \right.$$

$$\left. - \frac{U_1}{2}\frac{C_{20}C_{00}-C_{10}^2}{C_{00}C_{10}C_{20}}\frac{d^2I(U_1,\lambda U_1)}{dU_1^2} \right)$$

$$- \frac{U_1}{2I_0}\left(\sigma_1^2 \frac{\partial^2 I(U_1,U_2)}{\partial U_1^2} \right.$$

$$+ 2\mathrm{cov}(\check{U}_1,\check{U}_2)\frac{\partial^2 I(U_1,U_2)}{\partial U_1 \partial U_2}$$

$$\left. + \sigma_2^2 \frac{\partial^2 I(U_1,U_2)}{\partial U_2^2} \right)\Big|_{U_2=\lambda U_1}. \tag{28}$$

It follows from expression (28) that a more exact reconstruction of the energy spectrum with the noise corrections under the condition that the mean values of the potentials are linearly related requires knowledge of the values of the output signal not only at points where the potentials are linearly related, but also in a certain vicinity of the latter in the general case. If the additional conditions on the potential fluctuations

$$\sigma_2^2 = \lambda^2 \sigma_1^2, \quad \mathrm{cov}(\check{U}_1,\check{U}_2) = \lambda \sigma_1^2 \tag{29}$$

are satisfied, expression (28) transforms into an expression in total derivatives, and to reconstruct such a spectrum to within the corrections associated with a second derivative, it is sufficient to know the signal at the points where the mean values of the potentials are linearly related.

We note that the results obtained for spectrometers with two electrodes can easily be generalized to the case of an analyzer with n electrodes.

CONCLUSIONS

In conclusion, let us briefly describe the results obtained.

1. Equations (3) and (24) make it possible to describe the relationship between the energy distribution function of the charged particles and the output signal of an electrostatic analyzer operating in the spectrometer mode with the influence of fluctuations of the potentials on its electrodes taken into account.

2. The approximate solutions (9) and (28) of these equations for an arbitrary distribution function of the potential fluctuations permit taking into account the corrections associated with noise to the reconstructed energy distribution.

3. For analyzers with several electrodes under different potentials, it is not enough to know the values of the output signal at the points where the mean values of the potentials are linearly related. More exact reconstruction of the true energy distribution requires knowledge of the values of the output signal in the vicinity of these points and the dispersion of the noise.

4. An examination of the behavior of the instrumental function as a function of various parameters in the case of an analyzer with one deflecting electrode shows that at energies for which the width of the instrumental function of the analyzer is determined mainly by the potential fluctuations, the transmission of the analyzer will fall off as the energy decreases.

[1] V. A. Kurnaev and V. A. Urusov, Zh. Tekh. Fiz. **67**, 86 (1997) [Tech. Phys. **42**, 663 (1997)].

[2] G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers*, McGraw-Hill, New York (1961) [Russ. transl., Nauka, Moscow (1977)].

[3] V. V. Raznikov and M. O. Raznikova, *Analytical Mass Spectrometry* [in Russian], Nauka, Moscow (1992).

Translated by P. Shelnitz

Influence of heat treatment and exposure to laser radiation on a vanadium-silicon composite

A. M. Chaplanov and A. N. Shibko

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

The influence of exposure to weak laser radiation during heat treatment on a vanadium–silicon composite is investigated. It is found that the changes in the phase composition of the contact cause changes in its electrophysical parameters. Application of a combined treatment permits the formation of a rectifying vanadium–silicon contact with definite prespecified electrophysical parameters. © *1997 American Institute of Physics.* [S1063-7842(97)02006-0]

Refractory metals and their silicides have recently come to be actively employed in microelectronics because of their thermodynamic features, their promising electrophysical properties, and their stability at high temperatures. One of the ways to modify the electrophysical parameters of a metal–semiconductor system is heat treatment, which alters the electronic states at the interface. When metal–semiconductor systems are subjected to thermal annealing in a vacuum, oxidation processes generally take place, making it difficult to use them in the fabrication of semiconductor devices.[1] One of the techniques which make it possible to influence the oxidation of a surface layer is to irradiate it during annealing by a photon beam of definite energy.[2,3]

In the present work we investigated the phase transformations and changes in the electrophysical parameters of the V–Si system during heat treatment and simultaneous exposure to a laser photon beam with an energy $h\nu = 1.96$ eV.

A vanadium film of thickness ~ 100 nm was obtained by thermal deposition on n-type silicon(111) in a vacuum with a residual pressure equal to 3×10^{-4} Pa (Fig. 1a). The substrate temperature during deposition was 373 K. The silicon plate was chemically treated according to the method described in Ref. 4 before thermal deposition. The films obtained were polycrystalline and highly disperse, with a mean grain diameter of 15–20 nm.

The samples obtained were placed in a vacuum system and subjected to thermal treatment at a residual pressure of 1.5×10^{-4} Pa and simultaneous exposure to laser radiation with $\lambda = 0.63$ μm. The annealing temperature was $T = 500-700$ °C, and the annealing times were $\tau = 5$, 15, and 30 min. The power of the LGN-215 laser was varied and monitored by an IMO-2 calorimeter during the treatment, and it was equal to 25 or 55 mW.

The samples treated were investigated by electron diffraction analysis on a JEM-120 electron microscope and by x-ray photoelectron spectroscopy (XPS). The Schottky barrier height was determined from the current–voltage characteristics according to the method described in Ref. 5. The area of the vanadium–silicon contact was 0.07 mm^2.

During thermal annealing of the system without exposure to laser radiation, various oxide phases with a large oxygen content, viz., V_2O_5 and V_3O_7, form on the surface, depending on the annealing time and temperature. The vanadium oxides form as a result of the reaction of vanadium with oxygen that had been adsorbed by the film during its deposition or which diffused from the surrounding medium. As follows from Table I, when the annealing temperature is 700 °C and the annealing time is 15 or 30 min, V_2O_5 and VO_2 are present on the surface of the vanadium film (Fig. 1b).

The phase composition of the surface oxides does not undergo any changes as a result of thermal annealing of the system and simultaneous exposure to a laser photon beam with a power equal to 25 mW (the photon energy $h\nu = 1.96$ eV) (Table I). However, the appearance of vanadium silicide V_3Si at an annealing temperature of 700 °C and annealing time of 15 or 30 min (Fig. 1c) must be noted. Under these annealing conditions silicon diffuses into the vanadium, forming the silicide.

When the power of the laser radiation is increased to 55 mW, the phase composition of the surface of the system consists of V_2O_3 and V_2O_5, and at an annealing temperature of 700 °C it consists only of V_2O_3 (Fig. 1d, Table I). There are no vanadium oxides with a higher oxygen content, because the additional irradiation of the annealed system by a photon beam with an energy of 1.96 eV and a definite flux density suppresses the oxidation processes, and the dissolved oxygen becomes ''chemically'' inactive.[6] Analyzing Table I, we can conclude that the oxides V_3O_2 and VO_2, as well as the silicide V_3Si, are not present on the surface of the system when the samples are irradiated at a power of 55 mW. The photons interact with the M–O σ and π bonds. As a result, the screening of the M–M bonds decreases, preventing the formation of the high-resistivity and silicide phases on the surface. Incident laser radiation with a higher power is needed to suppress the oxidation processes, since thermal annealing results in the occurrence of an interaction between oxygen and the vanadium film surface, and a higher photon flux density is needed to neutralize the M–O bonds.

The XPS investigations performed showed that the amount of oxygen in the subsurface layer decreases as a result of heat treatment of the V–Si system at $T = 500$ °C for $\tau = 30$ min and simultaneous exposure to laser radiation with $\lambda = 0.63$ μm and a power of 55 mW. The investigations were performed using the oxygen 1S peak (Fig. 2). Figure 3 presents the variation of the oxygen concentration in the subsurface layer for various treatment conditions. When a heated V–Si composite is exposed to a laser beam, a photon is
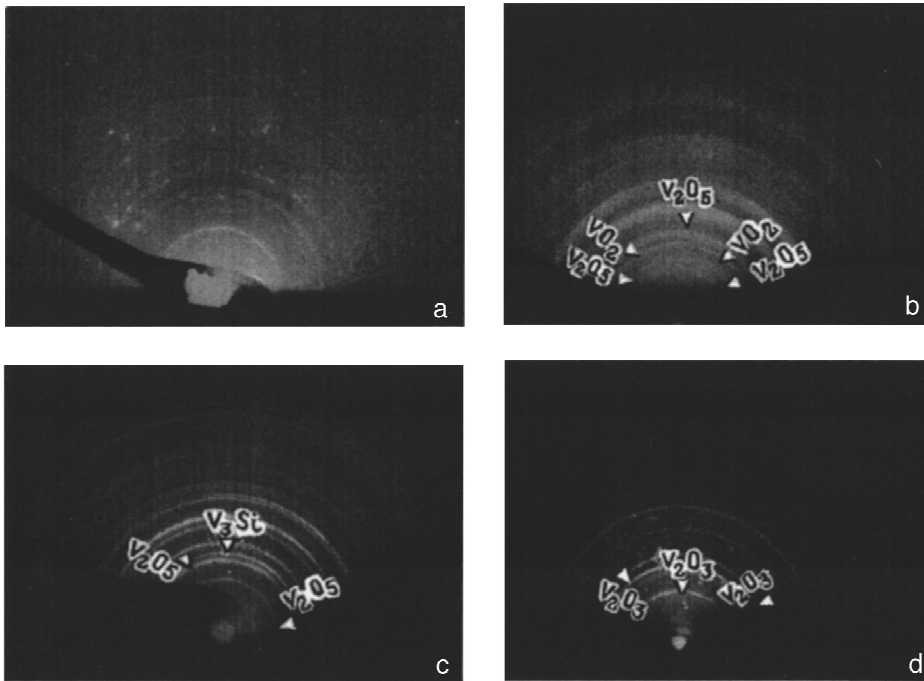
FIG. 1. Electron diffraction patterns of a V–Si composite: a — original sample; b — after heat treatment at $T = 700$ °C for $\tau = 15$ min; b — without exposure to laser radiation; c, d — with exposure to 25 and 55 mW laser radiation, respectively.

absorbed by the surface and oxygen atoms undergo a transition to one of seven states, which are characterized by weak chemical activity. This causes the appearance of different oxides, depending on the power of the laser radiation with $\lambda = 0.63$ μm striking the vanadium–silicon system during the thermal annealing. The results obtained by electron microscopy and electron diffraction analysis correlate with the results obtained by XPS.

The current–voltage characteristic of the V–Si system had an asymmetric form, and the breakdown voltage was 30 V. This is evidence that the native silicon oxide SiO_2 is present between the vanadium film deposited and the silicon, as is confirmed by XPS. Figure 4 presents the current–voltage characteristics of the V–Si system as a function of the treatment conditions. The results of the measurement of the electrophysical parameters after heat treatment and after the combined treatment are listed in Table II. A decrease in

the concentration of the various defects in the oxide film on the metal–semiconductor interface occurs as a result of thermal annealing. The structure of the SiO_2 layer tends to a more equilibrium state, and the number of electron traps decreases. When group IVa metals react with SiO_2, the formation and reduction of SiO_2 by these metals take place.[7] Therefore, vanadium diffuses into the SiO_2 layer and partially reduces it under the treatment conditions employed. This leads to a decrease in the thickness of the oxide. Consequently, when the annealing temperature and time are increased, the breakdown voltage in the V–Si system has a lower value. At the same time, the Schottky barrier height increases (Table II). When the V–Si system is subjected to

TABLE I. Variation of the phase composition of the V–Si system after heat treatment and after the combined treatment.

|				T °C				
	500			600			700	
				τ, min				
5	15	30	5	15	30	5	15	30
Without exposure to laser radiation								
V_3O_7	V_2O_3	V_2O_3	V_3O_7	V_3O_7	V_3O_7	V_2O_3	VO_2	VO_2
V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5
								V_3Si
With exposure to 25 mW laser radiation								
V_5O_9	V_6O_{13}	V_2O_3	V_2O_3	V_2O_3	V_2O_3	V_3O_7	V_3Si	VO_2
V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5
								V_3Si
With exposure to 55 mW laser radiation								
V_2O_3	V_2O_3	V_2O_3	V_2O_3	V_2O_3	V_2O_3	V_2O_3	V_2O_3	V_2O_3
V_6O_{13}	V_2O_5	V_2O_5	V_2O_5	V_2O_5	V_2O_5			
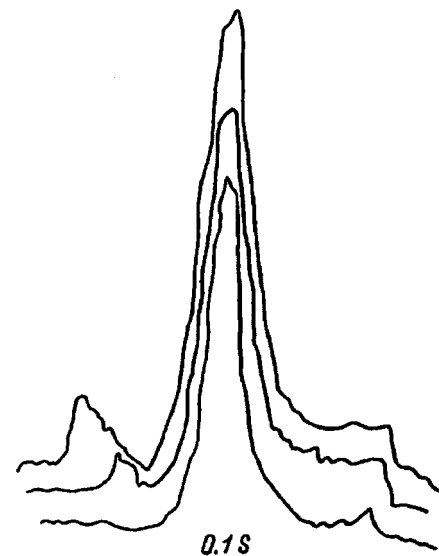
FIG. 2. XPS spectrum of oxygen in a V–Si composite after thermal annealing and exposure to laser radiation with $\lambda = 0.63$ μm.
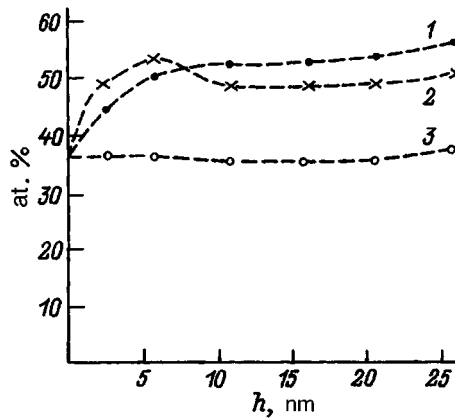
FIG. 3. Variation of the atomic percent oxygen in the subsurface layer of a V–Si contact: *1* — original sample; *2* — after heat treatment at $T = 600\ °C$ for $\tau = 10$ min; *3* — after combined treatment at $T = 600\ °C$ for $\tau = 10$ min and exposure to laser radiation with $\lambda = 0.63\ \mu m$ and a power equal to 55 mW.

TABLE II. Changes in the electrophysical parameters of the V–Si system after thermal treatment.

T [a]	τ	φ_1	φ_2	U_1	U_2	n_1	n_2	a_1	a_2
Original		0.56	0.58	30	20	1.18	1.08	33	36
500	5	0.57	0.59	25	20	1.15	1.04	34	37
	15	0.57	0.59	25	20	1.15	1.04	34	37
	30	0.57	0.60	25	15	1.14	1.04	34	37
600	5	0.58	0.60	30	15	1.10	1.04	35	37
	15	0.58	0.60	15	15	1.09	1.04	36	37
	30	0.58	0.61	15	10	1.08	1.03	36	38
700	5	0.59	0.61	20	15	1.08	1.04	36	37
	15	0.59	0.61	20	15	1.08	1.04	36	37
	30	0.59	0.61	20	15	1.08	1.04	36	37

[a]In Tables II–IV $T(°C)$ is the treatment temperature, τ (min) is the treatment time, φ (eV) is the Schottky barrier height, U (V) is the breakdown voltage, n is the ideality factor, a is a parameter of the current–voltage characteristic, the subscript 1 indicates the values after treatment, and the subscript 2 indicates values after treatment and etching.

thermal annealing, the density of the surface states in the silicon varies in the contact region of the metal–silicon junction and influences the barrier height. The current–voltage characteristic of the contact was characterized using the ideality factor n, which was determined from experimental plots[5] (Table II). It was shown that the ideality factor decreases as the annealing time and temperature increase. This is attributed to the variation of the dielectric parameters of the oxide layer at the interface and the variation of the concentration of electron traps during the heat treatment. Thus, when the annealing temperature and time are increased, the current–voltage characteristic of the V–Si contact tends to the ideal form.

Like the current–voltage characteristic of a contact, the parameter a characterizes the variation of the current with the applied voltage. Its value is determined from the formu-
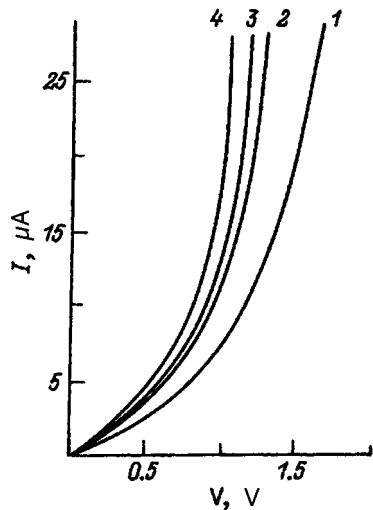
las presented in Ref. 8 and is given in the tables. The parameter a increases with increasing annealing temperature and time due to depletion of the near-contact region of impurities[8] or a decrease in the negative charge in the surface states. The calculations performed show that the inequality $a < e/(kT)$ holds for $n > 1$ so that the rectifying contact formed can be described within the diode theory.[9]

As a result of thermal annealing and simultaneous exposure of a V–Si system to a flux of laser photons with an energy $h\nu = 1.96$ eV, changes occur in its electrophysical parameters. The barrier height and a increase, while the breakdown voltage and the ideality factor decrease. It follows from a comparison of Tables III and Tables IV that the barrier height of the contact is higher after exposure to a 55 mW laser beam than after exposure to a 25 mW beam. The other parameters also differ. This is due primarily to the influence of the photons with $h\nu = 1.96$ eV on the chemical reactions occurring in the V–Si system, which are stimulated by the thermal annealing and the neutralization of oxygen.[6] This is indicated by the decreases in the breakdown voltage and the ideality factor n. In Tables II–IV the parameters of the V–Si system after removal of a surface layer of thickness ~ 10 nm are labeled with the subscript 2. These parameters

FIG. 4. Current–voltage characteristics of a V–Si contact after thermal treatment at $T = 600\ °C$ for $\tau = 10$ min and exposure to laser radiation with $\lambda = 0.63\ \mu m$: *1* — original contact; *2* — after thermal treatment; *3* and *4* — after thermal treatment and exposure to 25 and 55 mW laser radiation, respectively.

TABLE III. Changes in the electrophysical parameters of the V–Si system after heat treatment and exposure to laser 25 mW laser radiation.

T	τ	φ_1	φ_2	U_1	U_2	n_1	n_2	a_1	a_2
500	5	0.58	0.60	20	15	1.10	1.05	34	37
	15	0.58	0.60	15	10	1.06	1.04	37	37
	30	0.58	0.60	15	10	1.06	1.04	37	37
600	5	0.59	0.61	15	10	1.08	1.03	36	38
	15	0.59	0.61	15	15	1.08	1.04	36	37
	30	0.58	0.60	20	15	1.09	1.05	36	37
700	5	0.58	0.60	25	15	1.10	1.05	35	37
	15	0.58	0.60	20	15	1.10	1.05	35	37
	30	0.59	0.61	20	15	1.08	1.04	36	37

TABLE IV. Changes in the electrophysical parameters of the V–Si system after heat treatment and exposure to laser 55 mW laser radiation.

T	τ	φ_1	φ_2	U_1	U_2	n_1	n_2	a_1	a_2
500	5	0.59	0.63	15	5	1.04	1.02	37	38
	15	0.59	0.63	15	5	1.04	1.02	37	38
	30	0.59	0.63	15	5	1.04	1.02	37	38
600	5	0.58	0.62	25	10	1.10	1.02	35	38
	15	0.58	0.61	25	10	1.10	1.03	35	38
	30	0.59	0.62	15	10	1.08	1.02	36	38
700	5	0.59	0.62	20	10	1.08	1.02	36	38
	15	0.58	0.60	20	15	1.08	1.05	36	37
	30	0.58	0.60	25	15	1.10	1.05	35	37

differ form the parameters measured immediately after the treatment. For example, the barrier height of the contact increases. The electrophysical properties of the contact are altered, because in the first case the measurements were performed on the V_xO_y/VSi/SiO$_2$/Si system, and after the oxide layer was removed, the V_xSi_y/SiO$_2$/Si system was measured directly, as is confirmed by XPS. Thus, the 0.01–0.02 eV change in the barrier height is the value of the parameter introduced by V_xO_y. It is noteworthy that no strong ''surges'' in the values of n and a are observed, i.e., a rectifying contact forms as a result of thermal annealing and simultaneous exposure of the V–Si system to laser radiation with different values of the power.

[1] J. M. Poate, K. N. Tu, and J. W. Mayer (eds.), *Thin Films — Interdiffusion and Reactions*, Wiley-Interscience, New York (1978) [Russ. transl., Mir, Moscow (1982)].
[2] D. T. Alimov, V. K. Tyugaĭ, and P. K. Khabibulaev, Dokl. Akad. Nauk UzSSR (4), 23 (1985).
[3] D. T. Alimov, V. K. Tyugaĭ, P. K. Khabibulaev *et al.*, Zh. Fiz. Khim. **61**, 3065 (1987).
[4] *Handbook of Thin-Film Technology*, L. I. Maissel and R. Glang (eds.), McGraw-Hill, New York (1970) [Russ. transl., Sov. Radio, Moscow (1977)].
[5] E. H. Roderick, *Metal–Semiconductor Contacts*, Clarendon Press, Oxford (1980) [Russ. transl., Radio i Svyaz', Moscow (1982)].
[6] A. M. Chaplanov and A. N. Shibko, Neorg. Mater. **29**, 1477 (1993).
[7] S. P. Murarka, *Silicides for VLSI Applications*, Academic Press, New York (1983) [Russ. transl., Mir, Moscow (1986)].
[8] V. M. Strikha, Radiotekh. Elektron. **11**, 2092 (1966).
[9] V. M. Strikha, E. V. Buzaneva, and I. A. Radzievskiĭ, *Schottky Barrier Semiconductor Devices* [in Russian], Sov. Radio, Moscow (1974).

Translated by P. Shelnitz

Possibility of mechanically stimulated transmutation of carbon nuclei in ultradisperse deuterium-containing media

A. G. Lipson, V. A. Kuznetsov, T. S. Ivanova, E. I. Saunin, and S. I. Ushakov

Institute of Physical Chemistry, Russian Academy of Sciences, 117915 Moscow, Russia

The variation of the concentration of ^{14}C in graphite subjected to vibratory milling in the presence of heavy (or light) water is investigated using methods of radiocarbon β-spectral analysis. It is discovered that the $C+D_2O$ system exhibits a statistically significant increase in the β counting rate in comparison to the $C+H_2O$ system milled under similar conditions. A quantitative model of the transmutation of the ^{13}C isotope into the ^{14}C isotope involving an interaction of multiphonon excitations with deuterons in the graphite crystal lattice in the presence of background thermal neutrons is proposed. © *1997 American Institute of Physics.*
[S1063-7842(97)02106-5]

INTRODUCTION

We previously[1–3] showed that the absorption of elastic energy in crystalline particles takes on a quantum character when they achieve sufficiently small dimensions ($r_C \lesssim 10^{-5} - 10^{-4}$ cm) during pulsed mechanical loading. This means that the entire portion of elastic energy supplied is absorbed not uniformly by each atom in a particle but is concentrated in individual microregions of the particle in the form of quanta of an electromagnetic field.[1] This process becomes possible only when the interaction of a flux of elastic energy with a highly disperse crystalline medium leads to the formation of "giant" fluctuations of the elastic energy density. The possibility of the appearance of such fluctuations is associated in the general case with the presence of a broad size distribution of the particles being milled. When the mean diameter of an aggregate is of the order of several microns, the dispersed powder always contains a certain fraction with particle diameters smaller than 0.1 μm.[4] Therefore, it is clear that when an identical portion of elastic energy is supplied to all the particles being milled, in the small fraction of the powder with radii that are 1.5–2.0 orders of magnitude smaller than the mean value, the concentrated elastic energy density can be 5–6 orders above the mean value. In the case of so-called "hard" concentrating, in which the elastic energy density E'_{elas} is commensurate with the energy density of the optical phonons in a crystal (with a Debye frequency ω_D), according to the model in Refs. 1 and 2, a virtual electromagnetic bunch in the form of a coherent multiphonon excitation (CME) having a finite density that is commensurate with the nuclear density is created when elastic energy is concentrated in a particle[2] (in the case of "soft" concentrating the final density of such an electromagnetic bunch is much smaller than the nuclear density). This situation provides grounds for assuming that the case of the "hard" concentrating of elastic energy corresponds to the concentrating of the energy of a CME in one or more nucleons of a hypothetical atomic nucleus located at the center of symmetry of the disperse particle. As was shown in Refs. 5 and 6, the coherent concentrating of such multiphonon excitations in a nucleus in the crystal lattice of a disperse particle can result in the splitting off of a virtual neutron. This process takes place only if $W'_S > E_n$, i.e., if the CME energy per nucleon is greater than the binding energy of a neutron in the nucleus. For its part, the virtual neutron emitted by the neutron-donor nucleus can be recaptured by a neutron-acceptor nucleus located within the field of action of the nuclear forces. If there is no acceptor nucleus near (in terms of nuclear distances) the donor nucleus, the virtual neutron remains in the donor nucleus, since the time of action of a CME is commensurate with a nuclear time ($\tau \sim 10^{-22}$ s). It was postulated in Ref. 7 that the capture of virtual neutrons by acceptor nuclei can be a more intensive process, if there are background thermal neutrons in the lattice. In this case a background neutron catalyzes the transfer of the virtual neutron, i.e., it materializes the latter. Therefore, when there is a background of thermal neutrons, the distance between the neutron donor and acceptor nuclei can be far greater than the nuclear distance and can be of the order of the lattice constant.

Since the typical values of the energy of a CME in "hard" crystals lie in the range 5–10 MeV/nucleon,[3] the formation of a virtual neutron in deuterium-containing particles will be facilitated by the small value of the binding energy of a deuteron ($E_n = 2.22$ MeV). Therefore, the occurrence of nuclear transmutation (alteration of the isotopic composition of the crystal lattice) in such systems is most probable in the presence of a flux of elastic energy and a background of thermal neutrons.

In the present work an experiment on the mechanical milling of crystalline carbon (graphite) in the presence of an addition of D_2O was performed to test these hypotheses. The data obtained on the alteration of the isotopic composition of graphite, as will be shown below, confirm the possibility of the mechanical transmutation of carbon under the action of "giant" fluctuations of the elastic energy density, which lead to the formation of CMEs in the graphite particles.

EXPERIMENTAL METHOD

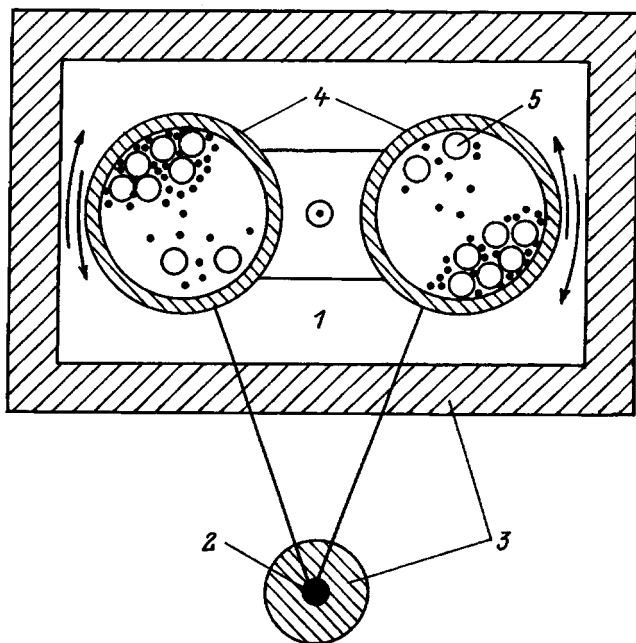The choice of graphite as the carrier of neutron-acceptor nuclei was motivated by the fairly high "stiffness" of this

FIG. 1. Diagram of the experimental setup: *1* — vibratory mill, *2* — neutron source, *3* — moderator, *4* — steel drums, *5* — beads.

substance,[3] which leads to the absorption of a considerable portion of the elastic energy by its particles from the mechanical energy supplied. In addition, carbon consists of two stable isotopes, ^{12}C and ^{13}C, and has only one long-lived β-active nuclide ^{14}C, and the variation of its concentration can be reliably traced using the methods of radiocarbon analysis.

The milling of graphite (C) together with D_2O (H_2O) was carried out in an M-35 vibratory mill with a power density equal to 10 W/g in stainless steel drums using steel beads with a diameter of 6 mm (Fig. 1). To increase the intensity of the mechanical action, the drums were cooled by liquid nitrogen after every 3 min of milling. To prevent the combustion of graphite, the drums were filled with argon to atmospheric pressure. Ultrapure MMPG-8 reactor-grade graphite with an initial specific surface $S_{sp} = 6.3$ m^2/g was used in the experiments. To increase the comminution rate some of the graphite was subjected to vacuum (10^{-4} Torr) annealing at 800 °C. The deuterium carrier used was heavy water of 99.9% purity, which was added to the drum with graphite in an amount equal to 5 wt.%. The control experiments were performed using a similar system, in which D_2O was replaced by H_2O (twice-distilled water), as well as by separately milling graphite without additions of water. The C+D_2O and C+H_2O systems were milled for fixed time intervals: 70 min in the first series of experiments (annealed graphite) and 160 min in the second series.

The kinetics of the comminution of the graphite particles were monitored by determining the specific surface S of the powders from the low-temperature adsorption of nitrogen (by the BET method) after different milling times. The dimensions of the crystallites and the magnitude of the microstrains corresponding to points on the $S(\tau)$ curve (where τ is the milling time) were determined by evaluating the mean

diameters of the coherent x-ray scattering regions (CSRs). These measurements were performed on a DRON-3 diffractometer (Cu K_α radiation), on which the Debye powder diffraction patterns of samples that had been ground to different degrees were obtained.

To prevent the random activation of carbon by neutrons from the time-fluctuating cosmic background, as well as to raise the number of background neutrons in the graphite and to increase the reproducibility of the results for ^{14}C, the background of thermal neutrons in the working zone was held constant and exceeded the mean value of the cosmic background by approximately ninefold. For this purpose, a ^{252}Cf neutron source in a polyethylene protective shell (with a thickness of 2.5 cm) with an intensity equal to 300 n/s in a 4π solid angle was placed at a distance of 90 cm from the working drum of the vibratory mill (Fig. 1). The measurements with a neutron detector[8] showed that the absolute intensity of the background of thermal neutrons in the working zone of the vibratory mill was $i_0 \simeq 5 \times 10^{-2}$ n/s·cm^2 (as opposed to the typical value for the cosmic background of thermal neutrons $i_0 \sim 6 \times 10^{-3}$ n/s·cm^2).

The radiocarbon analysis was performed using a β scintillation spectrometer based on a spectrometric detector of β radiation (a plastic scintillator), which was fastened to the photomultiplier of a VAS-968 detection unit. The signal from the detection unit was fed into a BUS2-47 spectrometric amplifier and then into a ChNO-1024-90 multichannel pulse-height analyzer. To minimize the contribution of the high-energy β radiation in the soft region of the spectrum, a thin Tesla SPF35V04 β detector with a thickness of 700 μm and a diameter of 40 mm, which is partially transparent to β particles with energies greater than 600 keV, was employed. To reduce the background of cosmic rays, the detection unit was placed within a lead protective shell with a thickness of 10 cm. The β activity of the dispersed graphite was measured using Al substrates with a diameter of 50 mm, on which a portion of the powder having a mass equal to 0.1 g and a thickness equal to 30 μm was placed. The substrate was placed at a distance of 2 mm from the detector. The detector was calibrated using standard metrological sources of ^{14}C, ^{99}Tc, and ^{204}Tl (Fig. 2, curves *1* and *3*). The efficiency of the detector with respect to ^{14}C with consideration of its calibration was 35% (73% for ^{204}Tl). To improve the accuracy of the detection of the β radiation from ^{14}C, the counting of this radionuclide was performed only in channels 21 to 60 (Fig. 2, curve *1*). At the same time, to determine any traces of high-energy β nuclides in the samples, the counting rate was continuously monitored in channels 61–250 of the analyzer. The measurements of the β count of graphite were alternated with background measurements (without the weighed portions of graphite). The values obtained were normalized to the mean background levels measured on the respective work day with the corresponding graphite samples.

EXPERIMENTAL RESULTS

The kinetics of the comminution of preliminarily annealed and unannealed graphite together with H_2O and D_2O are presented in Fig. 3. As is seen from the figure, preliminary annealing results in a sharp increase in the com-
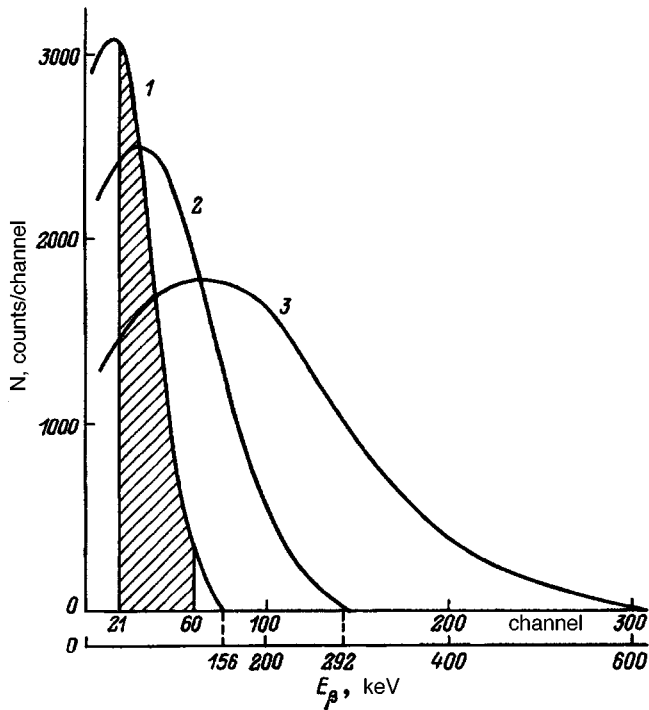
FIG. 2. Spectra of the β radiation of nuclides measured using an SPF35U04 plastic scintillator. The hatched portion of curve *1* corresponds to the recorded β count for ^{14}C.

minution rate in comparison with the unannealed samples. In the case of dry milling the specific surface of a graphite sample reaches $S_{sp} = 340$ m^2/g (Fig. 3, curve *1*). The introduction of water into the system lowers the comminution rate of the annealed sample, so that specific surfaces equal to 95 and 75 m^2/g are achieved after a time $\tau = 70$ min in $C + H_2O$ and $C + D_2O$ samples, respectively (Fig. 3, curves 2
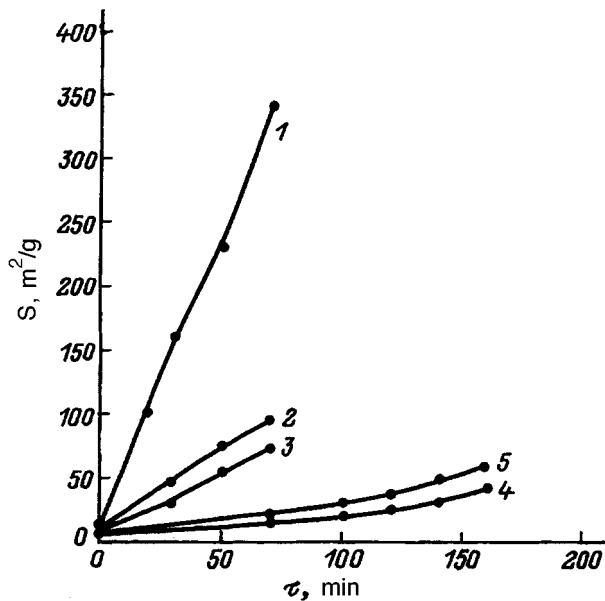
FIG. 3. Kinetics of the vibratory comminution of graphite: *1* — annealed graphite without admixtures, *2* — annealed graphite $+5$:H_2O, *3* — annealed graphite $+5$:D_2O, *4* — unannealed graphite $+5$:H_2O, *5* — unannealed graphite $+5$:D_2O.

and *3*). The values of S for unannealed samples milled with water are equal to only 42 and 59 m^2/g after 160 min of milling (Fig. 3, curves *4* and *5*).

Figures 4a–c present plots of the mean values of the diameter D of the coherent scattering regions, the microstrains ε, and the linewidth θ as functions of S. As is seen from these figures, in the region of values $S \sim 100$ m^2/g the $D(S)$, $\varepsilon(S)$, and $\theta(S)$ curves display a clear-cut discontinuity, which attests to a change in the character of the fracture of the graphite particles. In fact, the discontinuity on the $\theta(S)$ curve, which is associated with considerable broadening of the reflections on the Debye powder diffraction patterns, indicates that the graphite powder is strongly amorphized at $S \gtrsim 100$ m^2/g. A sharp increase in the microstrains ε and a decrease in the diameter D of the coherent scattering regions occur simultaneously. This supports the conclusion that because of the amorphization of the graphite particles no elastic energy is absorbed in them, and plastic flow takes place, leading to the ''vaporization'' of graphite. As follows from Fig. 4c, after a milling time $\tau = 70$ min for graphite in the presence of D_2O (H_2O) the mean width of the reflections taken from the [002], [004], and [110] crystallographic directions (like D and ε), does not differ from the analogous value for the original powder. At the same time (Fig. 5), when the milling time of the $C + H_2O$ (D_2O) systems is increased above 70 min, there is a sharp increase in the linewidth, the linewidth of the samples milled for $\tau = 160$ min being almost three times greater than the value of θ for the analogous samples with $\tau = 70$ min. This indicates that after milling for 70 min, graphite still exhibits crystalline properties like the original (unmilled) powder, while strong amorphization occurs when $\tau = 160$ min. Therefore, the fluctuational quantum model of the fracture of particles[1,2] (the formation of a CME in them) is applicable in the present case only to $C + D_2O$ (H_2O) systems milled for $\tau = 70$ min and cannot be used for powders milled for 160 min. We note that in the case of the amorphization of the graphite particles, the concept of a limiting minimum diameter is meaningless; therefore, when $S > 100$ m^2/g (Fig. 3, curve *1*) graphite can be ground to particle diameters as small as we wish. These data also provide evidence that the region of applicability of the fluctuational quantum model of the fracture of graphite particles corresponds to $S < 100$ m^2/g when $\tau = 70$ min, where they exist predominantly in the crystalline, rather than the amorphous state.

The results of the β counting for different graphite samples are presented in Table I. As we see, for the control samples of the original and separately milled graphite, as well as the $C + H_2O$ system after milling during the times $\tau_1 = 70$ min and $\tau_2 = 160$ min, the β counting rate does not exceed the background values within the measurement error over the entire region of the β spectrum investigated. At the same time, statistically significant upward deviations from the background in the control measurements with the $C + H_2O$ system were obtained for the analogous $C + D_2O$ samples in analyzer channels 21–60, which correspond to the β spectrum of ^{14}C. In the high-energy channels $(61-250)$ the $C + D_2O$ samples milled for $\tau_1 = 70$ min do not display values exceeding the background in the control
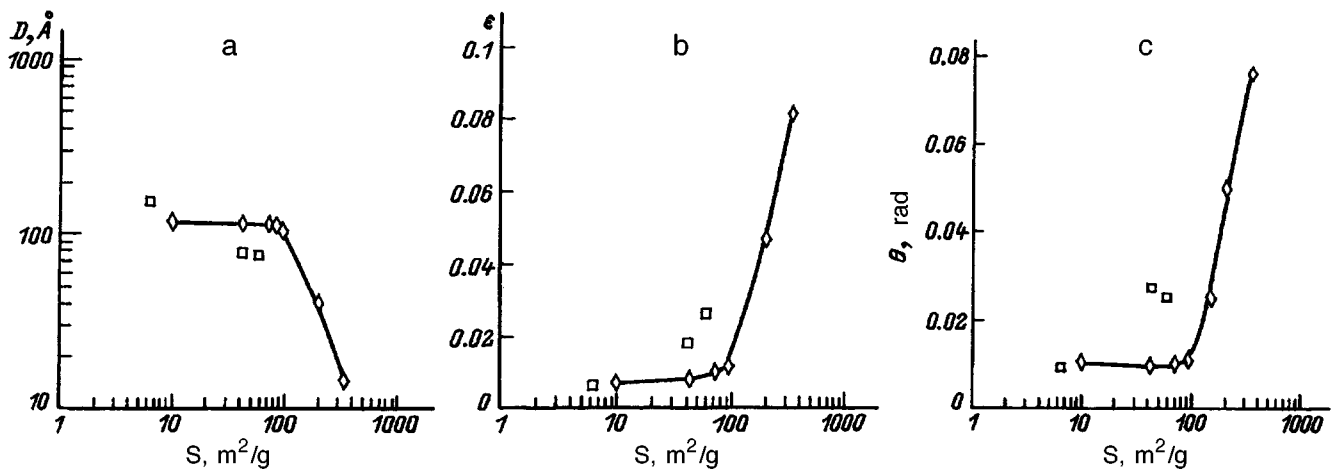
FIG. 4. Variation of the dimensions of the coherent scattering regions (a), the microstrain (b), and the linewidth (c) as a function of the specific surface of the powders for $\tau = 70$ min. The points correspond to the dimensions of the coherent scattering regions for $\tau = 160$ min as a function of S_{sp} for unannealed graphite milled together with H_2O or D_2O.

experiments, while there is a small upward deviation of the β activity in comparison to the $C + H_2O$ system for the samples milled for $\tau_2 = 160$ min. Thus, the $C + D_2O$ systems milled for 70 and 160 min exhibit a statistically significant increase in the β counting rate in the energy range corresponding to the ^{14}C spectrum in comparison to the $C + H_2O$ system milled under similar conditions. The upward deviations from the background in the control experiments equal 8 for $\tau_1 = 70$ min and 17 for $\tau_2 = 160$ min. The mean rate of the generation of ^{14}C with allowance for the calibration is approximately the same in both cases and amounts to $\langle \dot{N}(^{14}C) \rangle = 2 \times 10^6$ ^{14}C atoms/g·s. The increase in the β counting rate in the energy range corresponding to the ^{14}C spectrum for the milled $C + D_2O$ samples in comparison with the analogous $C + H_2O$ samples can have several causes.

1. The detection of β particles emitted by the tritium found in D_2O is eliminated by the fact that the detection region of the ^{14}C β particles lies at energies $E > 43$ keV, while the cutoff energy in the β spectrum of ^3T $E_{\text{cutoff}} \simeq 18$ keV.

2. The influence of the radioactivity of the "attrition products" formed as a result of the wear of the steel beads and the drum.

a) It can be assumed that carbon (including the ^{14}C concentrated in it) contained in the steel passes into the milled product. The influence of this factor, however, is eliminated, because the mass of the attrition products relative to the mass of the disintegrated graphite does not exceed 0.1 wt.%. When the carbon content in the steel used is taken into account, the amount of ^{14}C passing into the milled product does not exceed 10^{-3}% of the amount of ^{14}C found in the original portion of graphite and cannot make a measurable contribution to the β activity of the sample.

b) The influence of the "hard" β isotopes of the radium and thorium series adsorbed by the steel (Fig. 2, curves 2 and 3), which formed as a result of radioactive contamination of the environment. In this case upward deviations from the background level should be observed in the β spectra of the milled products at $E > 156$ keV. This effect was not detected

after short milling times ($\tau = 70$ min) (see Table I). At the same time, when $\tau = 160$ min, some increase in the count in analyzer channels 61–250 in comparison with the background is observed for the $C + D_2O$ and $C + H_2O$ systems. Therefore, the theoretical possibility of "hard" radionuclides passing from the steel into the ground product during prolonged mechanical treatments has not been ruled out. However, their contribution to the β spectrum of radiocarbon is small, since ΔN is small, although the detection efficiency of "hard" β particles is more than two times greater than that of ^{14}C β radiation. In addition, the intensity maximum of these radionuclides is shifted toward higher energies in comparison to the spectrum of ^{14}C; therefore, they cannot be the main cause of the increase in the β count in the energy range corresponding to radiocarbon.
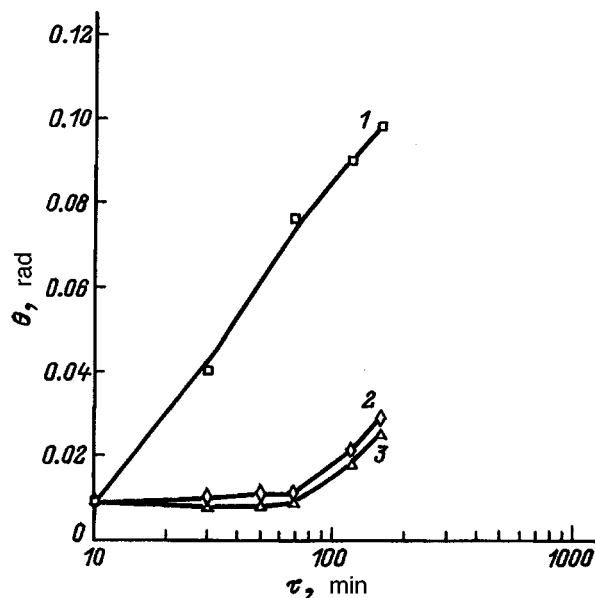

FIG. 5. Variation of the width of the x-ray lines as a function of the milling time. 1 — Unannealed graphite without admixtures, $C + H_2O$, 3 — $C + D_2O$.

TABLE I. Parameters of the β count of graphite samples milled under different conditions.

Sample No.	Sample	S_{sp},[a] m²/g	$N \times 10^{-2}$, counts/s (21–60)	$N \times 10^{-2}$, counts/s (61–250)	$\Delta N \times 10^{-2}$, counts/s (21–60)	$\Delta N \times 10^{-2}$, counts/s (61–250)	$\Delta n(^{14}C)$, Bq/g (21–60)
1	Background	–	7.00±0.14	14.60±0.31	–	–	–
2	C (original)	10	6.97±0.14	14.55±0.31			
					–	0.05±0.47	–
3	C (τ=70 min)	340	6.95±0.14	14.60±0.31			
4	C+H₂O (τ=70 min)	95	7.00±0.14	14.52±0.31			
					0.58±0.21	0.00±0.46	1.63±0.44
5	C+D₂O (τ=70 min)	72	7.58±0.15	14.52±0.31			
6	C+H₂O (τ=160 min)	42	7.05±0.15	15.20±0.33			
					1.20±0.23	1.00±0.50	3.43±0.45
7	C+D₂O (τ=160 min)	59	8.25±0.17	16.21±0.33			

[a]Here S_{sp} is the specific surface of the sample; N is the β counting rate of samples 1–7, respectively, in channels 21–60 and 61–250; ΔN denotes the differences between the counting rates of samples 3 and 2, 5 and 4, and 7 and 6; and $\Delta n(^{14}C)$ is the activity of the radiocarbon generated with consideration of the detector efficiency.

3. The influence of chemical isotope effects. Let us assume that for some reason preferential ''burnup'' of the principal (''light'') isotopes ^{12}C and ^{13}C occurs in the C+D₂O system in comparison with the C+H₂O system. In this case a significant loss of mass from the original graphite charge (8–17%) should be observed in order for the concentrating of ^{14}C to occur. This, however, does not occur, since the mass loss of graphite does not exceed 0.3%. In addition, it is well known that the influence of isotope effects in solid-phase systems increases sharply as the specific surface of the samples increases. Therefore, the samples with a large specific surface should have had a larger concentration of radiocarbon, which was not observed (see Table I). According to the experiment, the concentration of ^{14}C in the C+D₂O samples depends only on the milling time (the dose of elastic energy supplied) and does not depend on S_{sp}.

Thus, the accumulation of ^{14}C in the graphite samples when they are milled in the presence of D₂O and the absence of such an effect in the C+H₂O system cannot be attributed to the influence of radioactive contaminants, attrition products, or chemical isotope effects. Therefore, it should be assumed that the increase in the concentration of radiocarbon during the milling of the C+D₂O system is a consequence of nuclear transformation (transmutation) of the ^{13}C isotope into the ^{14}C isotope when multiphonon excitations interact with deuterons in the graphite crystal lattice in the presence of background thermal neutrons.[5,6]

In the next section we consider the possible causes of the observed transmutation of ^{13}C into ^{14}C and their relationship to the concentrating of elastic energy in the dispersed graphite particles.

DISCUSSION OF THE RESULTS

On the basis of the model presented, the mean number of carbon transmutation events ($^{13}C \rightarrow {}^{14}C$) in a disperse graphite particle per unit time $\dot{N}(^{14}C)$ is specified by the following relation:

$$\dot{N}(^{14}C) = N_S P_d P_{13C} P_n,\tag{1}$$

where N_S is the mean number of CMEs formed during milling per unit time for which $W_S > E_n$ (the binding energy of the neutron in a deuteron), P_d is the probability of the interaction of a CME with a deuteron in the graphite lattice, P_{13C} is the probability of finding a ^{13}C atom near that deuteron, and P_n is the probability of finding a background neutron near a deuteron or the probability of the interaction of such a neutron with a CME.

We must first prove that the CME energy per nucleon (W'_S) will be sufficient for splitting a deuteron, i.e., that (W'_S) > 2.22 MeV. For this purpose we utilize the equation of state of a degenerate neutron gas,[9] which, as was shown in Ref. 2, determines the maximum CME energy density

$$W'_S = \frac{3}{5}(3\pi^2)^{2/5} \frac{\hbar^{6/5}}{m_n^{3/5}} (\alpha L)^{2/5}.\tag{2}$$

Here m_n is the neutron mass, $\hbar$ is Planck's constant, α is the conversion factor for the transformation of elastic energy into the energy of lattice optical phonons of frequency ω_D (the Debye frequency), and L is a parameter having the meaning of the limiting elastic energy which can be stored by the crystals: $L = B_0/V_0\rho$, where B_0 is the equilibrium bulk modulus, V_0 is the unit cell volume, and ρ is the density of the crystal (in particular, for graphite $L = 4.2 \times 10^3 \, \text{erg/g} \cdot \text{cm}^3$).

To estimate W'_S using Eq. (2), we must determine the value of α. The value of α, in turn, which is also the probability of the excitation of phonon modes with the frequency $\omega_{D'}$ in a particle, is given in the hard case as[2]

$$\alpha = \frac{1}{\pi^{5/2}} \frac{m_n c^2}{(\hbar \omega_D)^2} (\hbar m_n)^{1/2} (\beta E_t^{1/2}),\tag{3}$$

where c is the speed of light, E_t is the total energy supplied to the crystalline particle, and β is a coefficient which is equal to the fraction of the elastic energy that is absorbed by

the particle during the mechanical treatment out of the total energy supplied (E_t), so that $\langle E_{elas}\rangle = \beta E_t$ (in a first approximation β can be regarded as the ratio of the energy stored in the ''elastic layer'' of a particle of radius r_{min} to the total energy supplied to it[3,10]).

It is not difficult to see that β can be evaluated from the x-ray structural data for the milled ($\tau = 70$ min) C+D$_2$O (H$_2$O) powders which retain their crystalline properties (Fig. 5). In fact, by definition, in a particle of radius r_{min} we have $\beta = EV_D/EV_{min}$, where V_D is the volume of the layer in which the elastic energy is stored and V_{min} is the volume of a particle of radius r_{min}. Assuming that the particle consists of n unit cells and has a radius nr_0 (where r_0 is the lattice constant) in the simplest case, we have $\beta = 4\pi(nr_0)^2 n\Delta r/(4/3)\pi r_0^3 n^3$, where Δr is the elementary distortion of a bond r_0, whence $\beta \simeq \Delta r/r_0 = \varepsilon$. Thus, in a first approximation β is equal to the predominant lattice microstrain ε appearing in the crystalline systems predominantly when elastic strain energy is stored. For the graphite milling time $\tau = 70$ min we have $\varepsilon \simeq \beta = 0.01$ (Fig. 4b). The substitution of $\beta = 0.01$ and $E_t = 10$ W/g into Eq. (3) with consideration of the value $\omega_D = 10^{14}$ Hz leads to the value $\alpha \simeq 0.25$. The value of the elastic energy conversion factor α can also be obtained from a treatment of the x-ray diffraction patterns of the graphite samples. As was previously shown in Refs. 1–3, the minimum particle radius that can be achieved when a flux of elastic energy is supplied to a dispersed crystalline medium is expressed as

$$r_{min} = \mu\omega_D^2/\alpha L, \tag{4}$$

where μ is the Poisson ratio, which specifies the smallest linear dimension of an asymmetric particle (particularly, in graphite).

When $\mu = 0.25$, $\omega_0 = 10^{14}$ Hz, and $\alpha < 1$, the theoretical minimum radius $(r_{min})_{th} = 58$ Å, which is very close to the radius of a crystallite (a coherent scattering region) $D/2 = 56$ Å for the C+D$_2$O system milled for 70 min (Fig. 4a). In fact, since a coherent scattering region has the diameter of a crystallite D, it is reasonable to assume that D should equal the mean minimum linear dimension of the particles that can be obtained by pulverizing an aggregate having the radius determined from the low-temperature absorption of nitrogen (by the BET method) (Fig. 3). Since in real cases $\alpha < 1$ in (4), setting $r_{min} = r_{BET}$ and $(r_{min})_{th} = r_{CSR} = D/2$, we obtain an expression for α for the BET coherent scattering region

$$\alpha = r_{CSR}/r_{BET}, \tag{5}$$

whence $\alpha \simeq 0.28$, which is in satisfactory agreement with the analogous value obtained from expression (3) for the case of the ''hard'' concentrating of elastic energy.[2] We note that the determination of α for $\tau = 160$ min is meaningless, since Eqs. (3) and (4) are not applicable in this case.

Knowledge of the value of α permits estimation of the critical radius of the graphite particles (r_c). When the radius of the particles $r < r_c$, the absorption of elastic energy in them takes on a quantum character,[1] which permits the use of Eqs. (2) and (3) to describe the effect observed. Therefore, the smallest radius of the graphite particles that are still pre-

dominantly crystalline after a milling time $\tau = 70$ min should be less than r_c. Otherwise, the quantum model of the absorption of elastic energy in graphite particles cannot be used. Since $r_c = (v_s c)^{1/2}/\alpha \cdot \omega_D$ (where v_s is the speed of sound in graphite, and c is the speed of light),[3] for $\alpha = 0.25$ we have $r_c = 4.2 \times 10^{-6}$ cm, while the mean radius of the graphite particles for $S_{max} = 72$ m^2/g ($\tau = 70$ min) is $r_{min} \simeq 1.9 \times 10^{-6}$ cm. Hence it follows that the absorption of elastic energy in graphite particles does, in fact, have a quantum character and the use of the quantum model is well-founded when $\tau = 70$ min. The substitution of the value of α into Eq. (2) gives an estimate of the CME energy in the graphite particles. When $\alpha = 0.25$, $W_S' = 9$ MeV/nucleon, which significantly exceeds the deuteron binding energy (2.22 MeV). Therefore, it can be assumed with a great degree of certainty that splitting of a deuteron does not occur when it interacts with a CME in the case under consideration.

It was shown in Ref. 2 that the condition $\langle E_{elas}\rangle = \alpha W_S'$, where $W_S = N_S W_S'$, holds in the case of the ''hard'' concentrating of elastic energy in a crystal (the concentrating of a CME in one or more nucleons in a nucleus). Here W_S is the total CME energy in a unit of mass of the crystal per unit time, and N_S is the number of such excitations. Hence with consideration of the equality $\langle E_{elas}\rangle = \beta E_t$ we have

$$N_S = \beta E_t/\alpha W_S'. \tag{6}$$

Substituting the known values of α, β, E_t, and W_S' into Eq. (6), we obtain $N_S = 2.4 \times 10^{11}$ CME/g·s. When Eq. (6) is taken into account, expression (1) takes on the form

$$\dot{N}(^{14}\text{C}) = (\beta E_t/\alpha W_S')P_d P_{13C} P_n. \tag{7}$$

To find $\dot{N}(^{14}\text{C})$ we now need only to evaluate P_d, defect P_{13C}, and P_n. The probabilities P_d and P_{13C} are determined trivially as the ratios of the atomic concentrations of deuterium and ^{13}C (when the concentrations by weight are known) to the atomic concentration of carbon (graphite). Hence $P_d = 6 \times 10^{-2}$, and $P_{13C} = 10^{-2}$. The value of P_n can be estimated approximately on the basis of the following arguments. Since the milled material is ''spread'' uniformly throughout the chamber (drum) during vibratory milling, its volume per gram will amount to 1/10 of the volume of the chamber or $v_1 \simeq 10$ cm^3, which corresponds to a surface $s_n = 20$ cm^2. According to measurements with a neutron detector, a flux of thermal neutrons $i_0 = 5 \times 10^{-2}$ n/s intercepts 1 cm^2 of surface of the chamber in the working zone. Therefore, a flux approximately equal to $I_n = i_0 S_n = 1$ n/g·s pervades 1 g of graphite under the conditions of our experiment. The number of neutrons interacting with a CME is determined by the number of CMEs and by the sojourn time τ_n of a thermal neutron in a unit mass of graphite ($m = 1$ g) and the flux of thermal neutrons entering a unit mass of graphite (I_n), i.e., $N_n = i_0 s_n \tau_n$ [n/CME]. When the linear dimensions of the sample (R_0) are sufficiently large, i.e., when $R_0 \gg L_D$ (the diffusion length of a neutron in graphite), τ_n will be close in order of magnitude to the diffusion time τ_D of the thermal neutrons or $\tau_n \simeq \tau_D$. In fact, since there are $\sim 10^{15}$ particles of diameter $\sim 4 \times 10^{-6}$ cm (Fig. 3, curve *3*),
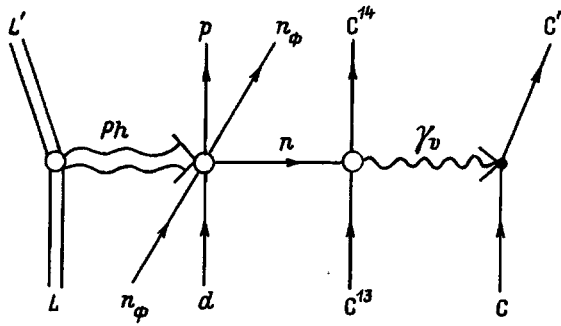
FIG. 6. Feynman diagram[6] of the proposed process leading to the formation of ^{14}C from ^{13}C.

the total linear dimension of the system exceeds 10^9 cm, while for graphite $L_D \simeq 60$ cm.[11] Therefore, $\tau_n \simeq \tau_D \simeq 1 \times 10^{-2}$ s. Thus, the probability of finding a background thermal neutron in the zone of action of a CME is $P_n \simeq N_n = 10^{-2}$.

Thus, substituting the value of N_S and of the probabilities P_d, P_{13C}, and P_n into expression (7), we obtain the rate of generation of the ^{14}C isotope $\dot{N}(^{14}C) \simeq 1.4 \times 10^6$ ^{14}C atoms/g·s. This value, which was calculated theoretically on the basis of theories regarding the quantum character of the absorption of elastic energy in graphite microparticles coincides in order of magnitude with the experimentally observed value $\langle \dot{N}(^{14}C) \rangle = 2 \times 10^6$ ^{14}C atoms/g·s. Thus, the change in the isotopic composition of graphite (the generation of ^{14}C) during vibratory milling in the presence of a deuterium-containing medium can be described with a fair degree of accuracy using the model of the quantum absorption[1–3] of a flux of elastic energy in disperse graphite particles in the presence of background thermal neutrons.

This situation allows us to present a possible scenario for the mechanotransmutation of ^{13}C into ^{14}C (Fig. 6). A coherent multiphonon lattice excitation ($L \rightarrow L' \rightarrow Ph$), which is a giant fluctuation of the elastic energy density supplied to a disperse graphite particle, brings about the splitting of a deuteron d into a proton p and a virtual neutron n. Since the resonance time of a background thermal neutron n_b near a deuteron $\tau \simeq 10^{-18}$ s is compatible with the lifetime of the composite nucleus $[n + ^{13}C]$, the presence of the background neutron leads to materialization of the virtual neutron formed when the deuteron d split. The role of the background thermal neutron thus reduces to a "catalyst" of the deuteron splitting and the transfer of the neutron to the neutron-acceptor nucleus (^{13}C) in the $n + ^{13}C \rightarrow ^{14}C + \gamma_v$ reaction (where γ_v is a virtual γ photon). Then the background neutron continues to move within the particle and interacts with other CMEs until it leaves the crystal. The neutron formed as a result of capture in the ^{14}C nucleus transfers energy equal to the binding energy of the neutron in the ^{14}C nucleus to the graphite lattice in the form of a virtual γ photon (the inverse

Mössbauer effect[12,13]) with the resultant formation of excited carbon atoms (C→C*) and defects in the graphite structure. We note again that the process described in Fig. 6 can be effective only for graphite with a high degree of crystallinity, in which the absorption of the flux of elastic energy by its particles has an essentially quantum coherent character. When the crystalline structure is lost, CMEs cannot form because of the lack of long-range order in the particles of the solid. Therefore, after long milling times ($\tau \geq 160$ min), when strong amorphization of the carbon particles is observed (Fig. 5), we should expect a sharp drop in the rate of the generation of ^{14}C, which, in turn, is limited by the drop in the rate of formation of the CMEs.

The increase discovered in the concentration of ^{14}C in the C + D$_2$O system during mechanical treatment, regardless of its true nature, can also take place in seismically active sites in the Earth's crust in the presence of deep groundwater, in which variation of the isotopic composition of carbon is observed.[14,15] We note that the error of the radiocarbon method for determining the age of fossil and organic systems will then increase, since the ^{14}C concentration at the time of measurement is determined not only by decay, but also by any alternative process leading to its accumulation.

To confirm the conclusions drawn, the performance of similar experiments at different values of the background of thermal neutrons, as well as with the use of other elements with a small neutron binding energy instead of deuterium, is planned for the future.

[1] V. A. Kuznetsov, A. G. Lipson, and D. M. Sakov, Zh. Fiz. Khim. **67**, 204 (1993).
[2] A. G. Lipson and V. A. Kuznetsov, Dokl. Akad. Nauk **332**, 172 (1993) [Phys. Dokl. **38**, 406 (1993)].
[3] V. A. Kuznetsov, A. G. Lipson, E. I. Saunin et al., Zh. Tekh. Fiz. **66** (3), 75 (1996) [Tech. Phys. **41**, 272 (1996)].
[4] G. S. Khodakov, *The Physics of Grinding* [in Russian], Nauka, Moscow (1972).
[5] J. Schwinger, Prog. Theor. Phys. **85**, 711 (1991).
[6] P. L. Hagelstein, Trans. Fusion Technol. **26** (4T), 461 (1994).
[7] A. G. Lipson, E. I. Saunin, S. I. Ushakov et al.., Pis'ma Zh. Tekh. Fiz. **21** (23), 19 (1995) [Tech. Phys. Lett. **21**, 957 (1995)].
[8] A. G. Lipson, D. M. Sakov, and E. I. Saunin, JETP Lett. **62**, 828 (1995).
[9] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 2nd ed., Pergamon Press, Oxford (1969).
[10] A. B. Volyntsev and A. N. Shilov, Dokl. Akad. Nauk **328**, 691 (1993) [Phys. Dokl. **38**, 89 (1993)].
[11] I. I. Gurevich and L. V. Tarasov, *Low-Energy Neutron Physics*, North-Holland, Amsterdam (1968).
[12] H. Frauenfelder, *The Mossbauer Effect*, Benjamin, New York (1962).
[13] H. Kozima, Nuovo Cimento A **27**, 1781 (1994).
[14] Yu. P. Konyushaya, in *Discoveries of Soviet Scientists, Part 1: Physico-technical Sciences* [in Russian], Izd. Mosk. Univ., Moscow (1988), p. 117.
[15] R. Sukumar, R. Ramesh, R. K. Pant, and G. Rajagopalan, Nature **364**, 703 (1993).

Translated by P. Shelnitz

Influence of partitioning on the dynamic stability of the superconducting state and the electrical losses in superconducting helicoids

A. N. Balev and S. L. Kruglov

Kurchatov Institute Russian Science Center, 123182 Moscow, Russia

The influence of partitioning of the winding of a superconducting helicoid on the electrical losses and dynamic stability of the superconducting state is investigated. A model of a helicoid in the form of an infinite isotropic hollow cylinder with an azimuthal transport current is used in the calculations. It is shown in the case of division of the helicoid into two sections that the appearance of a shielding current profile in the inner section initially increases the losses and lowers the dynamic stability. However, after the shielding current profile comes into contact with the principal current profile, the situation reverses. Ultimately, when the transverse section of the helicoid is completely filled by the current, the total electrical loss in the partitioned helicoid is approximately one-half of that in the unpartitioned helicoid. The highest permissible rate of current input that guarantees the absence of jumps in the magnetic flux in the winding (in other words, the dynamic stability boundary of the superconducting state) is three times higher in the partitioned helicoid than in the unpartitioned helicoid. Cases of a number of partitions greater than two are also considered. © *1997 American Institute of Physics.* [S1063-7842(97)02206-X]

INTRODUCTION

The purpose of partitioning superconducting magnetic systems is to optimize the winding with respect to losses in the superconductor. The winding is divided into sections with a decrease in the fraction of the superconductor in them as the magnetic field diminishes.

A superconducting helicoid with flat turns[1] is a fundamentally new design for superconducting windings and is regarded as an alternative to conventional windings in large stationary facilities, particularly in the superconducting magnetic systems of tokamaks.[2] The simplest helicoid is a set of separate parallel composite superconductors, which are galvanically and mechanically connected to form a single unit along a helical spiral (Fig. 1), for example, by soldering according to the novel technology described in Ref. 3.

A superconducting helicoid has several advantages over ordinary superconducting windings: a far higher mechanical stiffness and natural optimization with respect to the current density. The latter is due to the fact that at any point on a cross section of a helicoid the current flows either with the critical density or with zero density in accordance with the ''critical-state model.'' This empirical model has been tested experimentally many times and forms the basis of the electrodynamics of superconductors.

The deficiencies of a helicoid as a bulk superconducting composite structure include its high level of electrical energy losses to magnetization and the low dynamic stability of the superconducting state (which, for brevity, will be referred to below as dynamic stability). The influence of partitioning of the winding on these characteristics requires thorough study. The shielding of magnetic fields from the outer sections by the inner sections may increase the electrical losses and lower the limit of dynamic stability.

1. A ONE-SECTION (UNPARTITIONED) HELICOID

Let us consider the geometric model of a helicoid in the form of an infinite isotropic hollow cylinder with an azimuthal transport current (Fig. 2).

To visualize the situation we shall use the extremely simple critical-state model proposed by Bean, in which the critical current density does not depend on the magnetic field. The calculation of the electrical loses and the determination of the dynamic stability limit require knowledge of the distribution of the magnetic and electric fields in the cross section of the helicoid. For this reason we must solve the system of Maxwell's equations in cylindrical coordinates (Fig. 2)

$$\frac{dB}{dr} = -\mu_0 j_c, \quad B(r_1) = B_0,$$

$$\frac{1}{r}\frac{d(Er)}{dr} = -\dot{B}, \quad E(\delta_1) = 0.$$

Its solution is

$$B(r) = B_0 + \mu_0 j_c (r_1 - r),$$

$$E(r) = \frac{\dot{B}}{2r}(\delta_1^2 - r^2),$$

$$\delta_1 = \frac{B_0}{\mu_0 j_c} + r_1. \tag{1}$$

The coefficient k, which relates the transport current I of the helicoid and the magnetic induction at the center of the opening B_0 ($B_0 = kI$), is written in the model representation used as
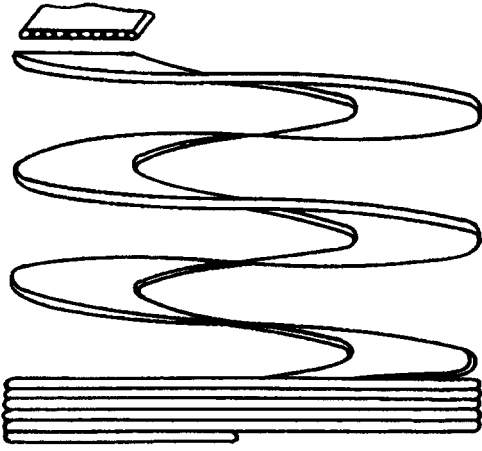
$$k = \frac{\mu_0}{\Delta},$$

FIG. 1. Superconducting helicoid with flat turns.

where Δ is the thickness of a flat turn of the helicoid.

In real helicoids it is, of course, smaller and is determined by their design.

The rate of the electrical losses per unit length of the helicoid in the model under consideration is

$$\frac{dQ}{dt} = 2\pi \int_{r_1}^{\delta_1} j_c E(r) r \, dr.$$

Using (1), we calculate the integral

$$\frac{dQ}{dt} = \pi j_c \dot{B}_0 \left(\frac{2}{3} \delta_1^3 - \delta_1^2 r_1 + \frac{r_1^3}{3} \right). \tag{2}$$

The total losses as the field increases monotonically from zero to B_m is

$$Q = \int_0^t \left(\frac{dQ}{dt} \right), \quad dt = \int_0^{B_m} \left(\frac{dQ}{dt} \right) \frac{dB_0}{\dot{B}_0}. \tag{3}$$

Substituting (2) into (3), we obtain

$$Q = \frac{\pi \mu_0 j_c^2 r_1^4}{3} \left(\frac{B_m^3}{\mu_0^3 j_c^3 r_1^3} + \frac{B_m^4}{2\mu_0^4 j_c^4 r_1^4} \right). \tag{4}$$
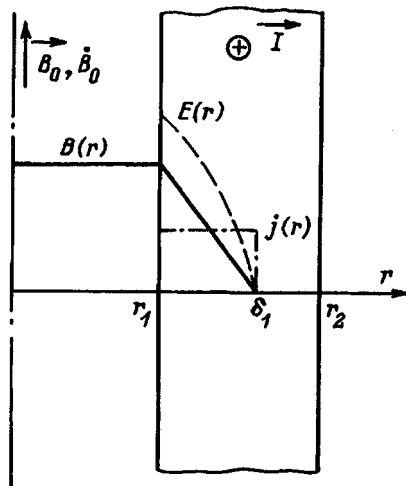
FIG. 2. Model representation of a helicoid by an infinite hollow cylinder.

For convenience we introduce the dimensionless magnetic field

$$b = \frac{B_m}{\mu_0 j_c r_1}$$

and the dimensionless electrical losses

$$q = \frac{Q}{\mu_0 j_c^2 r_1^4}.$$

As a result, (4) can be rewritten in the dimensionless form

$$q = \frac{\pi}{3} \left(b^3 + \frac{b^4}{2} \right). \tag{5}$$

The maximum achievable magnetic field in the helicoid is

$$B_m = \mu_0 j_c (r_2 - r_1);$$

therefore, the maximum value of b is

$$b = \frac{r_2 - r_1}{r_1}.$$

The criterion for the dynamic stability of composite superconductors in the dynamic approximation $D_t / D_m \to \infty$ (D_t and D_m are the thermal and magnetic diffusion coefficients) with the exponential form of the current–voltage characteristics taken into account was first obtained in Ref. 4:

$$\int_V E \frac{j_c}{j_0} \left| \frac{dj_c}{dT} \right| dV \leq \int_S h \, dS, \tag{6}$$

where j_0 and h are the growth parameter of the electrical resistance of the superconductor with respect to the current density and the coefficient of heat transfer from the superconductor surface.

Using the relationship between the broadening parameters of the current–voltage characteristics with respect to the current density j_0 and the temperature T_0 (Ref. 5),

$$T_0 = \frac{j_0}{\left| \frac{dj_c}{dT} \right|},$$

we obtain the expression for the dynamic stability limit (the absence of jumps in the magnetic flux as the magnetic induction varies from 0 to B_m at the rate $\dot{B}_0$)

$$\frac{db}{d\tau} \left(2b^2 + \frac{4b^3}{3} \right) \leq 1, \tag{7}$$

where τ is the dimensionless time:

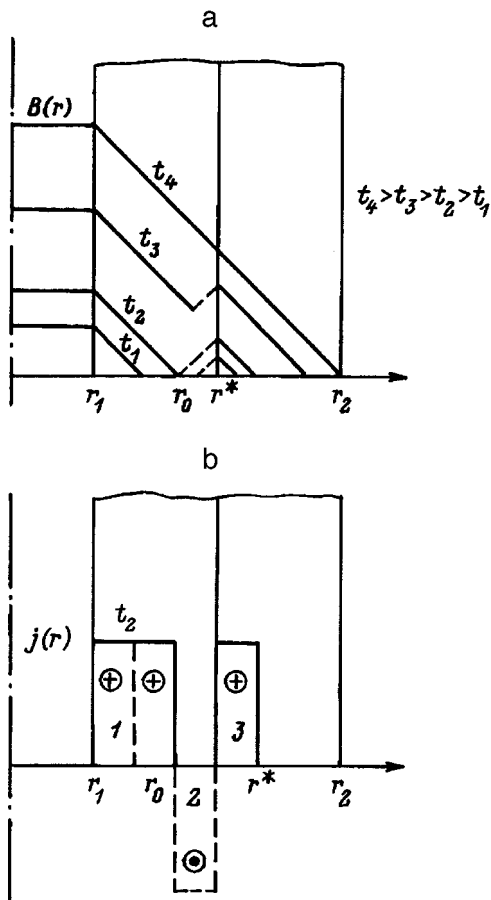$$\tau = \frac{4(r_1 + r_2) h T_0}{r_1^4 j_c^2 \mu_0} t.$$

FIG. 3. Dynamics of the distribution of the magnetic field (a) and the current (b) over a transverse section of a partitioned helicoid. Solid curves — principal profile, dashed curves — shielding profile.

2. A PARTITIONED HELICOID

Let us examine the influence of partitioning on the electrical losses and the dynamic stability in the case of a two-section helicoid with sections connected electrically in series.

The lack of a dependence of the critical current density on the magnetic field under the condition of equality between the transport currents in the sections dictates their equal thicknesses.

The fundamental difference between a partitioned and an unpartitioned helicoid is the shielding of the magnetic field from the outer sections by the inner sections, although the final distributions of the magnetic field and the current are identical and do not depend on the partitioning. Figure 3 shows the dynamics of the variation of the distributions of the magnetic field and the current from the beginning of current input ($t=0$) to the time t_4, when the current fills the entire transverse section of the partitions. Two temporal regimes can be distinguished during the increase in current: I – up to the moment when the principal and shielding profiles in the inner section come into contact (t_2 in Fig. 3a); II – from t_2 until complete filling of both partitions by the current (t_4 in Fig. 3a). Three regions in space can also be distinguished: 1 — the principal profile in the inner partition, 2 —

the shielding profile in the inner partition, 3 — the outer partition (Fig. 3b).

It is easy to show that the principal and shielding profiles first come into contact when the magnetic induction achieves the value

$$B_0 = \frac{2}{3}\mu_0 j_c (r^* - r_1),$$

where r^* is the boundary between the sections,

$$r^* = \frac{r_2 + r_1}{2}.$$

Equating the magnetic fields of the principal and shielding profiles, we can obtain the dependence of the coordinates of their contact point r_0 on the magnetic induction

$$r_0 = \frac{B_0}{4\mu_0 j_c} + \frac{r_1 + r^*}{2}.$$

Omitting the cumbersome calculations, we write out the final expressions for the loss rates in regimes I and II:

$$\frac{dQ^{\mathrm{I}}}{dt} = \frac{\pi j_c B_0}{2}\left(\frac{4}{3}\delta_1^3 + \frac{2}{3}(\delta_2^3 + \delta_3^3) - 2\delta_1^2 r_1 \right.$$

$$\left. - (\delta_2^2 + \delta_3^2)r^* + \frac{2}{3}(r^{*3} + r_1^3)\right),$$

$$\frac{dQ^{\mathrm{II}}}{dt} = \frac{\pi j_c B_0}{2}\left(2r_0^3 - 2r_0^2 r_1 - r_0^2 r^* - \delta_3^2 r^* \right.$$

$$\left. + \frac{2}{3}r_1^3 + \frac{2}{3}\delta_3^3 + \frac{2}{3}r^{*3}\right).$$

With consideration of the cooling of the helicoid surface we can rewrite (6) in the form

$$\frac{dQ}{dt} \leq 2\pi(r_1 + r_2)hT_0.$$

Introducing the dimensionless radius $\rho^* = r^*/r_1$, we can represent the dependence of the total losses and the dynamic stability limit on the magnetic induction in the different regimes in the following form: regime I $[0 \leq b \leq 2/3(\rho^* - 1)]$

$$q = \frac{\pi}{3}\left(b^3 + \frac{b^4}{2} + \rho^* \frac{b^3}{2}\right), \tag{8}$$

$$\frac{db}{d\tau}\left(2b^2 + \frac{4}{3}b^3 + \frac{\rho^* b^2}{2}\right) \leq 1, \tag{9}$$

regime II $[(2/3)(\rho^* - 1) \leq b \leq (r_2/r_1 - 1)]$

$$q = q^* + \frac{\pi}{81}\left(\frac{20}{3}\rho^* + \frac{16}{3}\right)(\rho^* - 1)^3, \tag{10}$$

where

$$q^* = \frac{\pi}{2}\int^b\left(\frac{1}{3}\rho^{*3} + \frac{5}{12} + \frac{1}{8}b\rho^{*3} + \frac{11}{96}b^3 + \frac{3}{8}\rho^* b^2\right.$$

$$\left. - \frac{1}{4}\rho^{*2} + \frac{b^2}{16} - \frac{\rho^*}{2} - \frac{b}{8}\right)db\frac{2}{3}(\rho^* - 1),$$
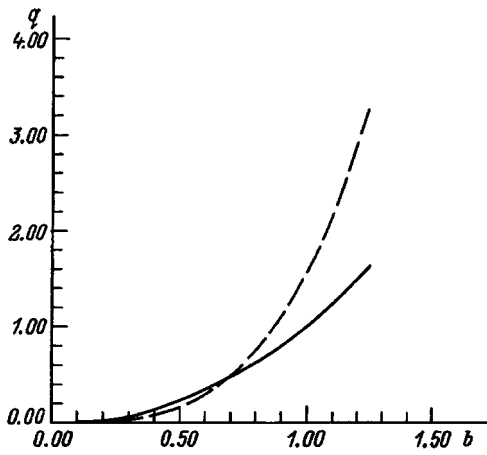
FIG. 4. Dependence of the electrical loss on the magnetic induction (in relative units) for partitioned (solid line) and unpartitioned (dashed line) helicoids.

$$\frac{db}{d\tau}\left(\frac{1}{3}\rho^{*^3}+\frac{5}{12}+\frac{1}{8}b\rho^{*^3}+\frac{11}{96}b^3+\frac{3}{8}\rho^*b^2\right.$$

$$\left.-\frac{1}{4}\rho^{*^2}+\frac{b^2}{16}-\frac{\rho^*}{2}-\frac{b}{8}\right)\leqslant 1. \qquad (11)$$

The dependence of the total loss on the magnetic induction (in dimensionless form) for unpartitioned [Eq. (5)] and partitioned [Eqs. (8) and (10)] helicoids is shown in Fig. 4 (the value of ρ^* equal 1.625 for the laboratory-scale helicoid investigated in Ref. 6 was used, and the range of ρ^* for real helicoids is generally $1.4-1.8$). It is seen that the total loss for the sectioned helicoid is at first higher due to the existence of the shielding profile, but in regime II the loss in the partitioned helicoid increases more slowly due to the weaker mean electric field. When the helicoid is completely filled by the current, the ratio between the total loss in the partitioned and unpartitioned helicoids equals 0.49.

The dependence of the dynamic stability limit of a helicoid on the magnetic induction for partitioned [Eqs. (9) and (11)] and unpartitioned [Eq. (7)] helicoids is presented in Fig. 5.

As in the case with the loss, in regime I the unpartitioned helicoid exhibits better dynamic stability, but in regime II the partitioned helicoid is considerably more stable. At the maximum achievable field the ratio between the safe rates of variation of the magnetic field (the transport current) for the partitioned and unpartitioned helicoids equals 3.

The processes for a large number of partitions can be considered using the same arguments. Without presenting the very cumbersome calculations, in Fig. 6 we show as an example the dependence of the ratio of the total loss (before complete filling of the cross section by the current) in the partitioned helicoid to the loss in an unpartitioned helicoid on the number of partitions.

3. MAIN RESULTS

The geometric representation of a superconducting helicoid by an infinite isotropic hollow cylinder into which a transport current is fed at a constant rate in an azimuthal direction was used in the calculations. Expressions for the electrical energy losses to magnetization and for the dynamic stability limit of the superconducting state in unpartitioned and partitioned (for the case of two sections) helicoid windings have been obtained in Bean's critical-state model. It has been shown that the appearance of a shielding current profile in the inner section initially leads to an increase in the electrical losses and lowering of the dynamic stability in the partitioned winding in comparison to the unpartitioned winding. However, after the principal and shielding current profiles come into contact, the situation reverses because of the lower mean electric field. Ultimately, when the transverse section of the helicoid winding is completely filled by the current, the total electrical losses in the partitioned variant are two times smaller and the dynamic stability limit is three times higher than in the unpartitioned winding. Cases of more than two partitions have also been considered.

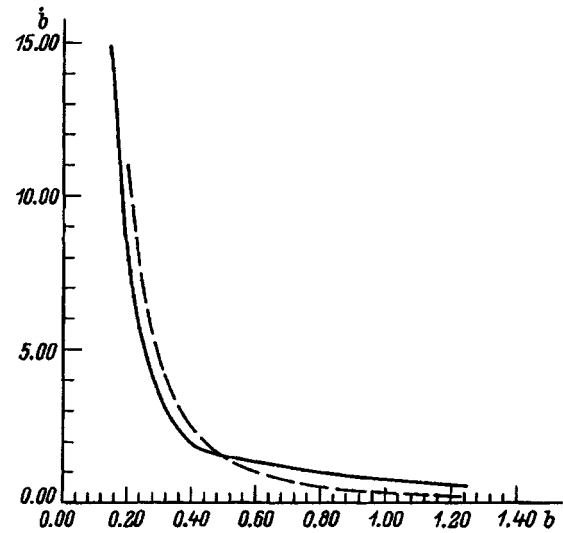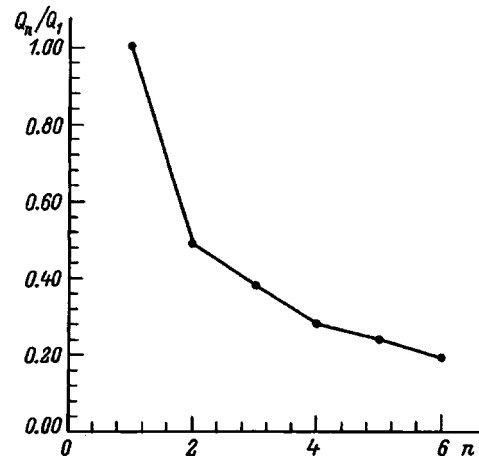This work was performed with support from the

FIG. 5. Dependences of the maximum permissible rate of increase in the magnetic field on the induction (in relative units) for partitioned (solid line) and unpartitioned (dashed line) helicoids.

FIG. 6. Dependence of the ratio between the total electrical losses in the partitioned and unpartitioned helicoids on the number of partitions.

[1] V. E. Keĭlin, I. A. Kovalev, S. L. Kruglov *et al.*, Dokl. Akad. Nauk SSSR **303**, 1366 (1988) [Sov. Phys. Dokl. **33**, 959 (1988)].

[2] V. E. Keilin, I. A. Kovalev, N. F. Kopeikin *et al.*, J. Fusion Energ. **11**, 1 (1992).

[3] V. E. Keĭlin, I. A. Kovalev, S. L. Kruglov *et al.*, USSR Inventor's Certificate (Patent) No. 1325587; Byull. Izobret. No. 27 (1987).

[4] R. G. Mints and A. L. Rakhmanov, J. Phys. D **15**, 2297 (1982).

[5] G. L. Dorofejev, A. B. Jmenitov, and E. Ju. Klimenko, Cryogenics. **20**, 307 (1980).

[6] I. A. Kovalev and S. L. Kruglov, Zh. Tekh. Fiz. **65**(11), 150 (1995) [Tech. Phys. **40**, 1170 (1995)].

Translated by P. Shelnitz

Sputtering of Al$_x$Ga$_{1-x}$As semiconductor targets by Ar$^+$ ions with energies of 2−14 keV

N. A. Berg and I. P. Soshnikov

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

Investigations of the sputtering of Al$_x$Ga$_{1-x}$As semiconductor solid solutions by Ar$^+$ ions with energies of 2−14 keV are performed. The dependence of the sputtering yield on the energy and angle of incidence of the ions are determined and the character of the surface relief formed during the sputtering is investigated. A comparison with theory shows that the best agreement between theory and experiment is achieved when the Haff–Switkowski formula is used together with Yudin's stopping cross section. It is shown that the surface binding energies obtained differ from the atomization energies by an amount approximately equal to the amorphization energy. © *1997 American Institute of Physics.* [S1063-7842(97)02306-4]

INTRODUCTION

Perfecting the ion-sputtering methods used in analytical research and the industrial processing of semiconductor structures and devices would be unfeasible without detailed quantitative experimental data on the processes imitated by ion bombardment. In addition, investigations of the sputtering of III–V materials by ion bombardment are of fundamental interest, since the features and laws governing the sputtering of multicomponent semiconductors can be studied in the case of these isovalent materials.

The published experimental data provide dissimilar pictures of the laws governing the sputtering of different III–V materials. The binary compounds GaAs and InP have been investigated relatively completely: the dependences of the sputtering yields on the angle of incidence, energy, and species of the ions[1–9] and the character of the topographic surface relief[1,2,4,5] formed as a result of ion bombardment have been studied, and the changes in the surface composition[10–13] caused by the preferential sputtering effects have been determined. At the same time, the sputtering of Al$_x$Ga$_{1-x}$As solid solutions has been studied to only a small extent. The data in Refs. 14–16, where the dependence of the sputtering yield on the target composition was measured using Ar$^+$ and Xe$^+$ ions with an energy $E=2$ keV at an angle of incidence $\Theta=60°$, are clearly insufficient for either a meaningful analysis of the phenomenon studied or for practical use.

The purpose of the present work was to investigate the main laws governing the sputtering of Al$_x$Ga$_{1-x}$As solid solutions upon irradiation by Ar$^+$ ions in the energy range from 2 to 15 keV.

EXPERIMENTAL

The target samples investigated were cut from plates having a layered solid-solution/substrate (Al$_x$Ga$_{1-x}$As/GaAs) structure. The samples were subjected to preliminary testing using electron probe x-ray microanalysis and electron microscopy, during which it was determined that the thickness of the Al$_x$Ga$_{1-x}$As layer was no less than 4 μm and that the composition within each sample differed across its area from the nominal composition by less than 1%

and had a gradient across its thickness no greater than $dx/dz \leqslant 0.02\%/\mu$m. After the preliminary testing, the samples were irradiated by Ar$^+$ ions in a regime chosen in accordance with the experimental program on a Balzers 1EU-100 ion-etching system or a specially constructed ion-sputtering module consisting of an ion source of the duoplasmatron type[17] and a sample holder. The ion sources of the sputtering systems were designed for an accelerating voltage from 1 to 10 kV and an ion current density of the order of $j \sim 10-500$ μA/cm^2. In addition, the design of the stage employed in the sputtering module made it possible to shift the target potential without disrupting the operating regime of the ion beam and to vary the energy of the ions accordingly to $E=15$ keV.

The sputtering yield Y was determined by measuring the thickness of the sputtered layer, i.e., from the height h of the step which forms on the sample surface during ion bombardment when part of it is masked. A detailed description of the method for determining the sputtering yield was given in the preceding reports.[1,2]

The topography of the surfaces of the irradiated samples was investigated by scanning electron microscopy.

RESULTS AND DISCUSSION

During the investigations of samples irradiated by Ar$^+$ ions with an energy $E=5$ keV in doses $\Phi \sim 10^{18}-10^{19}$ ion/cm^2, local relief features of the ''cone in a crater'' type (Fig. 1) with typical dimensions of the order of 0.1–5 μm and a density as high as 10^3 cm^{-2} were observed on the target surface. The highest density of these features was observed on the GaAs samples, and it decreased as the aluminum content increased. The formation of such a surface morphology has been observed repeatedly on various amorphized materials[1–3,5] and is usually attributed to the presence of random local inhomogeneities on the target surface.[18,19]

It is generally known that the state of an irradiated surface can vary as a function of the target temperature during ion bombardment: at $T<T_c$ the surface layer amorphizes, and at $T>T_c$ the amorphization process is suppressed by the annealing of defects. The recrystallization temperature of
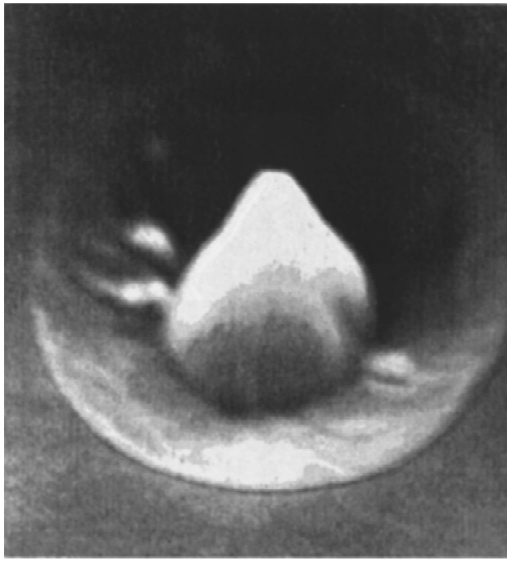
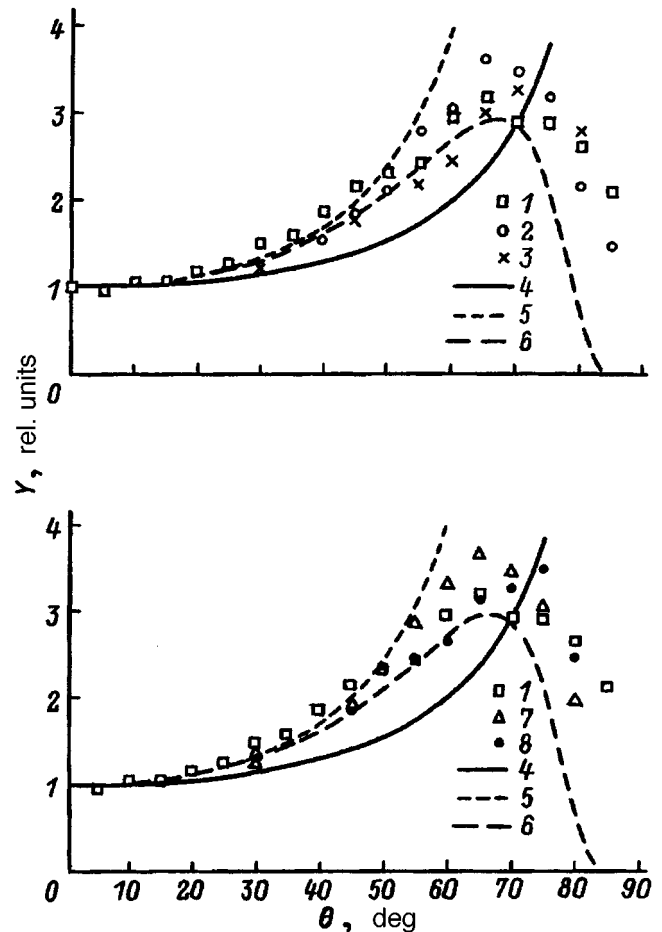FIG. 1. Electron-microscope image of the surface of a GaAs sample.

FIG. 2. Dependence of the sputtering yield on the angle of incidence of Ar^+ ions with an energy $E = 5$ keV. Experimental data: *1* — GaAs, *2* — $Al_{0.4}Ga_{0.6}As$, *3* — $Al_{0.8}Ga_{0.2}As$, *7* — $Al_{0.2}Ga_{0.8}As$, *8* — $Al_{0.6}Ga_{0.4}As$; calculation: *4* — $f \sim 1/\cos\Theta$, *5* — $f \sim 1/\cos^2\Theta$, *6* — GaAs according to Yamamura's theory.

gallium arsenide is $T_c = 125$ °C.[4] Taking into account the principle of the correlation of the temperatures for the crystallization and annealing of vacancies[7,20] for III–V materials, we can expect that the recrystallization temperature of Al-GaAs solid solutions is not less than 125 °C.

Since the temperature of the samples in the control experiments did not exceed $T = 75$ °C, it should be expected that the surface layer is amorphized. The results of the tests to determine the state of the surface (the absence of orientational effects in the ejection patterns of the "Wehner spots") and the agreement between the sputtering yields of the samples with the (100) and (111) crystallographic orientations of the surface support the assumption that the surface is amorphized under the experimental conditions chosen. Thus, Sigmund's model for the sputtering of amorphous bodies[21,22] is applicable for analyzing the experimental results.

The results of the measurements of the angular dependence of the sputtering yield $Y_n(\Theta)$ are presented in Fig. 2. Such curves are typical of the sputtering of amorphous targets, i.e., without a fine structure and with a single maximum in the range $\Theta = 65 - 75°$, confirming the conclusion that the surface of the targets is amorphized during ion bombardment. A comparison of the experimental data with the asymptotic dependence $Y \sim 1/\cos^h\Theta$ reveals that the exponent depends on the target composition and varies from $h = 2.2$ for $x = 0$ to $h = 1.25$ for $x = 0.8$. In addition, the variation of the composition of the target leads to displacement of the position of the maximum from $\Theta_{max} = 65°$ for $x = 0$ to $\Theta = 72°$ for $x = 0.8$. Such variation of the plots of the angular dependence of the sputtering yield is apparently attributable to a decrease in the reflection efficiency of the incident particles and an increase in the mean free path of the ions in targets containing the light element.

A comparison of the experimental data with the results of a calculation based on Yamamura's formula[23]

$$Y \sim t^f \exp[g(t-1)], \tag{1}$$

where $t = 1/\cos\Theta$, $g = f/\cos\Theta_{max}$, and $f = h/(1 - 1/\cos\Theta_{max})$, reveals good agreement in the range $\Theta < \Theta_{max}$. However in the range of glancing angles $\Theta > \Theta_{max}$ the experimental results differ from the results based on Yamamura's model. This can be attributed either to the influence of surface microrelief or to the incorrectness of using the model in this range.

Figures 3 and 4 present the results of measurements of the sputtering yield for the normal incidence of the Ar^+ ions as a function of the energy and the composition of the target. A comparison of the plots obtained of the dependence of the sputtering yield on the composition with the results in Refs. 14 and 16 shows that the plots obtained under different vacuum conditions have a similar character and that, more specifically, variation of the aluminum content leads to variation of the sputtering yield only in the range $x \leq 0.3$. The sputtering yield of the targets with an aluminum content greater than 0.3 depends weakly on the composition. Since the experiments in Refs. 14 and 15 were carried out with a residual pressure in the vicinity of the sample no greater than 10^{-8} Torr, which eliminates oxidation of the most active component, aluminum, the agreement between the results of the present work and the literature data attests to the com-
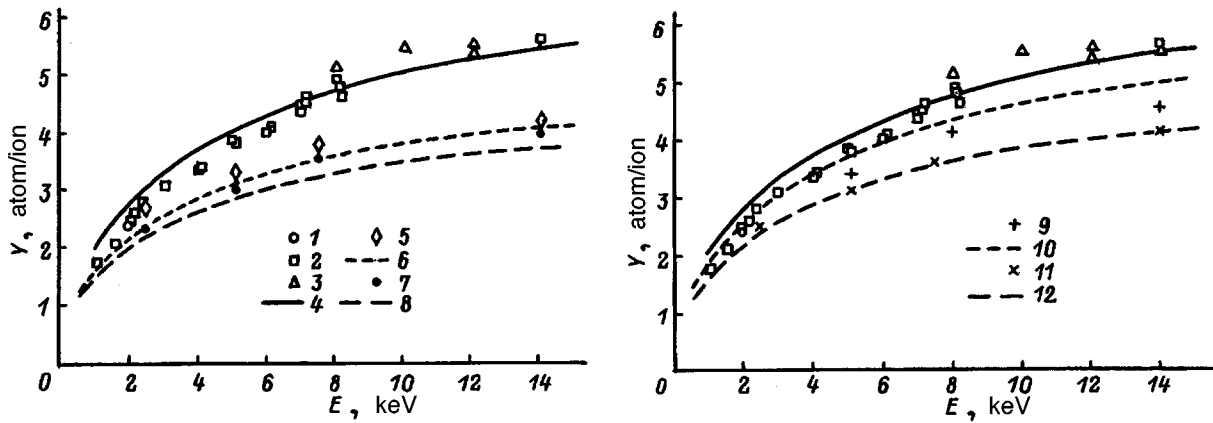
FIG. 3. Dependence of the sputtering yield on the energy of Ar^+ for normal incidence. Data from the literature: 1 — from Ref. 6, 3 — from Ref. 3 for GaAs; experimental data: 2 — GaAs, 5 — $Al_{0.4}Ga_{0.6}As$, 7 — $Al_{0.8}Ga_{0.2}As$, 9 — $Al_{0.2}Ga_{0.8}As$, 11 — $Al_{0.6}Ga_{0.4}As$; calculation using Eq. (5) and Yudin's model : 4 — GaAs, 6 — $Al_{0.4}Ga_{0.6}As$, 8 — $Al_{0.8}Ga_{0.2}As$, 10 — $Al_{0.2}Ga_{0.8}As$, 12 — $Al_{0.6}Ga_{0.4}As$.

plete absence or very weak occurrence of oxidation in the range of vacuum conditions chosen.

To describe the sputtering of multicomponent materials, Sigmund[22] obtained the formula

$$Y = \sum_i C_i(\Lambda_i)F_D(E,\Theta,0), \qquad (2)$$

where Y is the total sputtering yield, C_i is the concentration of target atoms of type i, Λ_i is a ''material'' constant for the atoms of species i, and $F_D(E,\Theta,0)$ is the distribution function of the absorbed energy.

Performing a series of transformations in analogy to Refs. 2, 22, and 24, we obtain the expression

$$Y = (1/U_0)\sum_i [C_i^s \alpha_i S_i], \qquad (3)$$

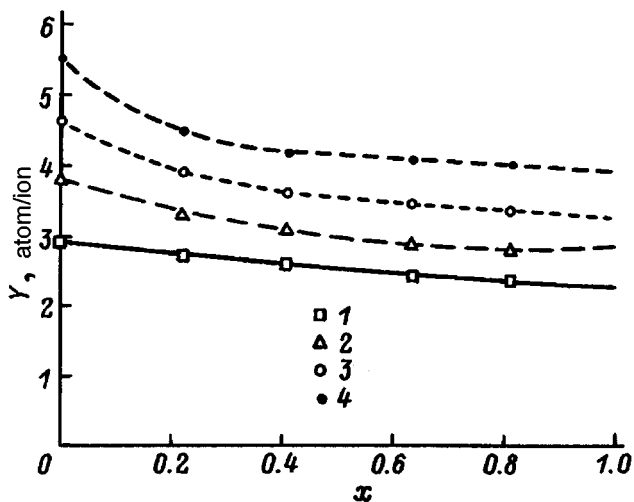where S_i is the nuclear stopping cross sections of the atoms of type i, U_0 is the reduced surface binding energy, α_i is a

FIG. 4. Dependence of the sputtering yield for Ar^+ ions at normal incidence on the target composition. E, keV: 1 — 2.5, 2 — 5, 3 — 7.5, 4 — 14.

dimensionless parameter that depends on the mass ratio M_i/M_{ion} and the angle of incidence of the ions, and C_i^s is the surface concentration of the atoms of type i.

We note that Eq. (3) is similar to the Haff–Switkowski formula for binary compounds.[24] The values of the surface concentrations were determined from the solution of the system of equations

$$Y = ax^2 + bx + c,$$

$$Y_{Al}/Y_{Ga} = C_{Al}^\nu/C_{Ga}^\nu,$$

$$(Y_{Al} + Y_{Ga})/Y_{As} = (C_{Al}^\nu + C_{Ga}^\nu)/C_{As}^\nu,$$

$$C_{Al}^s|_{x=0} = 0, \quad (C_{Al}^s/C_{As}^s)|_{x=1} = 1.5,$$

$$C_{Ga}^s|_{x=1} = 0, \quad (C_{Ga}^s/C_{As}^s)|_{x=0} = 1.4. \qquad (4)$$

The ratio between the surface concentrations of the elements for GaAs in this system of equations was taken from the literature,[7–10] and the ratio for AlAs was calculated from the Patterson–Shirn–Sigmund formula[22,25]

$$C_A^s/C_B^s = C_A^\nu/C_B^\nu(U_{0B} - U_{0A})(S_B/S_A), \qquad (5)$$

where the C_i^ν ($i=A,B$) are the bulk concentrations of the atoms.

The values of the energies for the monoelemental materials $(U_0)_A$ and $(U_0)_B$ (Ref. 26) were taken as approximate values for U_{0A} and U_{0B}.

The reduced surface binding energies U_0 appearing in expression (4) were determined by curve fitting. The properties of the model approximations of the nuclear stopping

TABLE I. Regions of applicability of approximations of the stopping cross sections.

Model	Reference	Working range
LNS	Ref. 24	$E \geqslant 0.1E_{TF}$
Biersack	Ref. 25	$E \sim 1$ keV
Yudin	Ref. 26	$E = 0.1 - 40$ keV
Matsunami	Ref. 27	$E = 0.1 - 40$ keV

TABLE II. Surface binding energies U_0 obtained by curve fitting.

Target	Atomization energy D, eV	Amorphization energy U_m, eV	Binding energy U_0, eV			
			LNS	Biersack	Yudin	Matsunami
GaAs	3.28	0.5	2.75	3.2	2.8	2.75
$Al_{0.2}Ga_{0.8}As$	3.4	0.6	2.7	3.0	2.9	2.9
$Al_{0.4}Ga_{0.6}As$	3.5	0.7	2.7	3.0	2.8	2.8
$Al_{0.6}Ga_{0.4}As$	3.6	0.8	2.7	2.9	2.8	2.8
$Al_{0.8}Ga_{0.2}As$	3.7	0.9	2.7	2.9	2.7	2.7

cross section[26–29] (Table I) were taken into account in the calculations. An analysis of the values of the binding energy U_0 obtained, which are presented in Table II, shows that they differ from the atomization energies by an amount equal to the amorphization energy (the heat of fusion).[31,32] This finding is apparently due to the influence of the amorphization processes on the sputtering of the materials studied.

CONCLUSIONS

As the composition of the $Al_xGa_{1-x}As$ targets changes from $x=0$ to $x=0.8$, a change in the parameters of the angular dependence is observed: the exponent of the cosine changes from $h=2.2$ to $h=1.25$; b) the position of the maximum shifts from $\Theta=66$ to $73°$. These changes are apparently attributable to a decrease in the stopping efficiency of the particles in collision cascades and an increase in the mean free path of the ions in targets with a large content of the light element (aluminum).

In the range of angles $\Theta \leqslant \Theta_{max}$ the Yamamura model is well suited to describing the angular dependences. This theory, however, does not provide satisfactory results for the range of glancing angles $\Theta > \Theta_{max}$.

The analog of the Haff–Switkowski formula for multicomponent materials [Eq. (3) in the text] can be used to describe the energy dependence of the sputtering yield of $Al_xGa_{1-x}As$ targets.

The reduced surface binding energies obtained differ from the atomization energies by an amount approximately equal to the amorphization energy.

In conclusion, we wish to thank I. A. El'yashevich for the research samples supplied and S. G. Konnikov for support of this work.

[1] N. A. Bert, K.Yu. Pogrebitskiĭ, I. P. Soshnikov, and Yu. N. Yur'ev, Zh. Tekh. Fiz. **62**(4), 162 (1992) [Sov. Phys. Tech. Phys. **37**, 449 (1992)].
[2] N. A. Bert and I. P. Soshnikov, Fiz. Tverd. Tela (St. Petersburg) **35**, 2501 (1993) [Phys. Solid State **35**, 1239 (1993)].
[3] J. Farren and W. J. Scaife, Talanta **15**, 1217 (1968).
[4] S. R. Bhattacharya, D. Ghose, and D. Basu, Indian J. Pure Appl. Phys. **25**, 328 (1987).
[5] S. R. Bhattacharya, D. Ghose, and D. Basu, Nucl. Instrum. Methods Phys. Res. Sect. B **47**, 253 (1990).
[6] W.-X. Chen, L. M. Walpita, C. C. Sun, and W. S. C. Chang, J. Vac. Sci. Technol. B **4**, 701 (1986).
[7] J. Linders, H. Niedrig, T. Sebald et al., Nucl. Instrum. Methods Phys. Res. Sect. B. **13**, 374 (1986).
[8] O. Wada, J. Phys. D **17**, 2429 (1984).
[9] H. Tubo, S. Namba, Y. Yuba et al., Jpn. J. Appl. Phys. **22**, 1206 (1983).
[10] I. L. Singer, J. S. Murday, and L. R. Cooper, Surf. Sci. **108**, 7 (1981).
[11] S. Valeri, A. di Bona, and E. Nava, Vuoto: Sci. Technol. **20**, 77 (1990).
[12] P. H. Holloway, Appl. Surf. Sci. **26**, 550 (1986).
[13] H. J. Kang, T. W. Kang, J. J. Lee et al., J. Vac. Sci. Technol. A **7**, 3251 (1989)
[14] S. W. Downey, A. B. Emerson, and R. F. Kopf, Nucl. Instrum. Methods Phys. Res. B **62**, 456 (1992).
[15] H. Kinoshita, T. Isida, and K. Kaminishi, Appl. Phys. Lett. **49**, 204 (1986).
[16] J. R. Arthur and J. J. LePore, J. Vac. Sci. Technol. **14**, 979 (1977).
[17] V. T. Barchenko and A. Sokolovskiĭ, Izv. Leningr. Elektrotekh. Inst., No. 303, pp. 42–47 (1982).
[18] G. Carter and M. J. Nobes, Proc. R. Soc. London A **407**, 405 (1986).
[19] L. B. Begrambekov, Itogi Nauki Tekh., Ser. Puchki Zaryazhen. Chastits Tverd. Telo, No. 7, pp. 4–53 (1993).
[20] D. V. Lang and L. C. Kimerling, in *Point Defects in Semiconduters, 1974*, Institute of Physics, London (1974), p. 581.
[21] P. Sigmund, Phys. Rev. **184**, 384 (1969).
[22] P. Sigmund and N. Q. Lam, Mat.-Fys. Medd. K. Dan. Vidensk. Selsk. **43**, 255 (1993).
[23] Y. Yamamura, Radiat. Eff. **80**, 57 (1984).
[24] P. K. Haff and Z. E. Switkowski, Appl. Phys. Lett. **29**, 549 (1976).
[25] W. L. Patterson and G. A. Shirn, J. Vac. Sci. Technol. **4**, 343 (1967).
[26] K. Konaya, K. Hujou, K. Koga et al., Jpn. J. Appl. Phys. **12**, 1297 (1973).
[27] J. Lindhard, V. Nielsen, and M. Scharfi, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd. **36** No. 10 (1968).
[28] J. P. Biersack, Nucl. Instrum. Methods Phys. Rev. Sect. B. **27**, 21 (1987).
[29] V. V. Yudin, Elektron. Tekh., Ser. Poluprovodn. Prib. 6(172), 3 (1984).
[30] N. Matsunami, Y. Yamamura, N. Itoh et al., Radiat. Eff. Lett. **68**(3), 83 (1982).
[31] *Thermodynamic Constants of Materials*, No. 5 [in Russian], V. P. Glushko (ed.), Izd. Akad. Nauk SSSR, Moscow (1971).
[32] *Physicochemical Properties of Semiconductor Substances* [in Russian], A. V. Novoselova (ed.), Nauka, Moscow (1978).

Translated by P. Shelnitz

BRIEF COMMUNICATIONS

Investigation of the axisymmetric buckling of round plates

M. A. Khusainov

Novgorod State University, 173003 Novgorod, Russia
(Submitted February 8, 1996)

The stress–strain behavior of a round plate during thermal cycling through the range for martensitic phase transformations is investigated. It is shown that the plate loses stability and recovers its original shape with a clap when a certain temperature $A_{in} < T_{ls} \leqslant A_{fin}$ is reached. The results of measurements of the impact and reactive forces are presented. The conditions for the occurrence of a clap are determined. © *1997 American Institute of Physics.*
[S1063-7842(97)02406-9]

The round plates in measurement devices usually act as sensitive elements (membranes) which respond to changes in pressure and temperature. If the elements are not joined along their perimeter to a rigid ring and receive only a transverse pressure as the temperature varies, they buckle. It was shown in Ref. 1 that a compressed zone forms when the deflection of the plate is large and that this zone is the focus of the loss of stability by the plate. The buckling of bimetal plates varies in proportion to the temperature, while the deflection of shape-memory alloys changes suddenly at a characteristic temperature. A loss of stability of a plate accompanied by instantaneous buckling (a clap) was observed in the stage in which the shape-memory effect was realized.[2] The temperature corresponding to sudden axisymmetric buckling with a clap coincided in some cases with the temperature of completion of the reverse martensitic phase transformation (A_{fin}). When the thermal cycles were repeated, a degree of spontaneous shape recovery in the heating stage and incomplete repair of the strain after repeated claps were noted.

Further study of the features of the thermomechanical behavior of round nickel titanium plates would be of interest in connection with the sudden repair of the strain imparted to the martensitic phase.

The experiments were carried out on plates of various dimensions made from TiNi of equiatomic composition. To assign ''shape memory'' to the material, a plate of diameter D was depressed in a chuck to a depth f in the martensitic state. The radius of the depression R was selected so as to satisfy the condition $f/h \geqslant 2$, where h is the thickness of the plate.[1] The restrained plate in the chuck was annealed at 500 °C (20 min).

The investigative procedure was as follows: a plate was strained in the martensitic state to a distance $2f$ under a load P applied in the transverse direction so as to provide for axisymmetric buckling due to transition of the midplane of the plate into a surface of revolution (Fig. 1). Then the plate was heated to the temperature A_{fin}. In this stage significant compressive reactive stresses appeared on the perimeter, which led to a loss of stability and instantaneous shape recovery.

Figure 2 presents characteristic curves which graphically depict the recovery of the shape of the membrane on heating. Straight lines *1* and *1′* illustrate the strains caused by the loads P and P', which create deflections of the membrane equal to $2f$ and $2f'$ in the martensitic state. The beginning of the heating in the first cycle is depicted by curves *2* and *2′*. The repair of the strain accompanied by a clap at the temperatures T_{ls} and T'_{ls} is illustrated by linear segments *3* and *3′*. Complete recovery of the preassigned shape is completed at the temperatures A_{fin} and A'_{fin}.

The data presented provide evidence that the repair of the strain with a clap takes place mainly at T_{ls} (the point of the loss of stability), which differs from A_{fin} of the original sample. The displacement of T_{ls} and A_{fin} toward increased temperatures is related to the magnitude of the strain (the deflection) imparted to the martensitic phase. The degree of nonrecovery of the shape Δf_{nr} when the clap occurs in the first cycle is insignificant. Repetition of the thermal cycles, i.e., deflection in the martensitic state, heating to A_{fin}, and cooling to M_{fin}, caused increases in Δf and Δf_{nr}, as well as displacement of the temperatures T_{ls} and A_{fin} (linear segments *4* and *4′*). The original hysteresis loop is represented here by curve *5*.

It was noted in Ref. 2 that an impulse arises during the instantaneous buckling of a round plate. However, no quantitative data were presented in that work. Here we present the results of experimental investigations of the thermomechanical behavior of plates (disks) under the conditions of a specified repairable strain at the temperature T_{ls} followed by rigid restraining to generate reactive stresses.

To obtain systematic results the dynamometer was set at an assigned value using a micrometer screw (Fig. 3). The impact force was determined from the deflection of the arrow
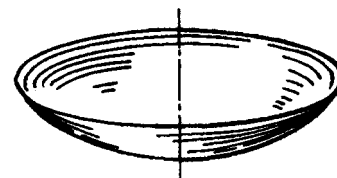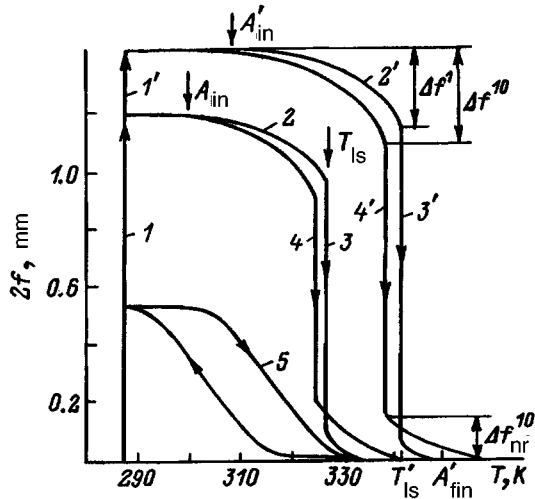

FIG. 1. Axisymmetric buckling of a round plate.

FIG. 2. Typical diagrams of the thermomechanical hysteresis of nickel titanium disks. Here Δf^1 is the deflection before the loss of stability in the first cycle, and Δf^{10} is the deflection in the tenth cycle.

on the scale of the instrument, and the reactive force P_r and the temperature were recorded on an LKD $X-Y$ potentiometer. The experiments showed that the dynamometer set at a distance f_i^{rd} maintains an impact force P_{im} and ensures rigid restraining of the plate, preventing free shape recovery. As a result, reactive stresses are generated in the disk material. Their magnitude for an assigned rigidity of the system is determined by the degree of preliminary deflection (the strain). The kinetics of the development of the reactive stresses in the regime leading to a clap and restricted strain are shown in Fig. 4. It is seen that a jump in the stresses occurs at a temperature close to A_{fin}. This event is accompanied by a clap and impact against the opposing body (the dynamometer). As the heating is continued in the compressed state, the reactive force P_r, whose level depends on the assigned magnitude of the free deflection f_i^{rd}, appear in the plate material.

The sudden shape recovery with impact at a certain temperature T_{ls} still requires study. However, several factors that definitely initiate a clap can be cited. The martensitic lattice is known to acquire instability near the temperature A_{in} with the resultant appearance of chemical and mechanical restoring forces. At the same time, reactive stresses, which act in the transverse direction, are generated in the plate material. Under the influence of these forces, a compressed zone forms
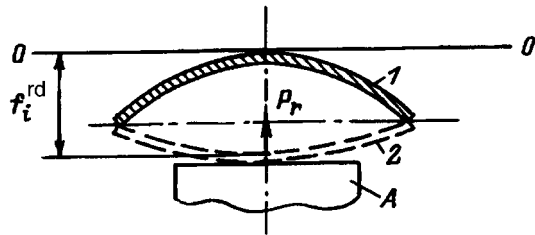

FIG. 3. Position of a round plate (disk) relative to the opposing body: 1 — original position of the plate, 2 — position at the time of impact against the opposing body A during heating. Here f_i^{rd} is the repairable deflection of the plate from the center of curvature.
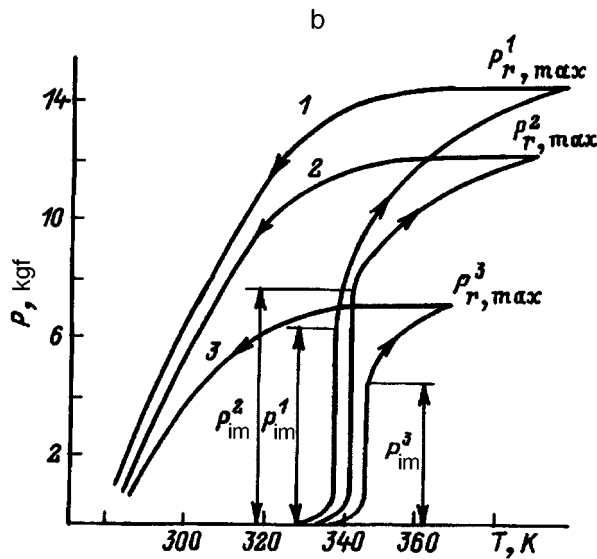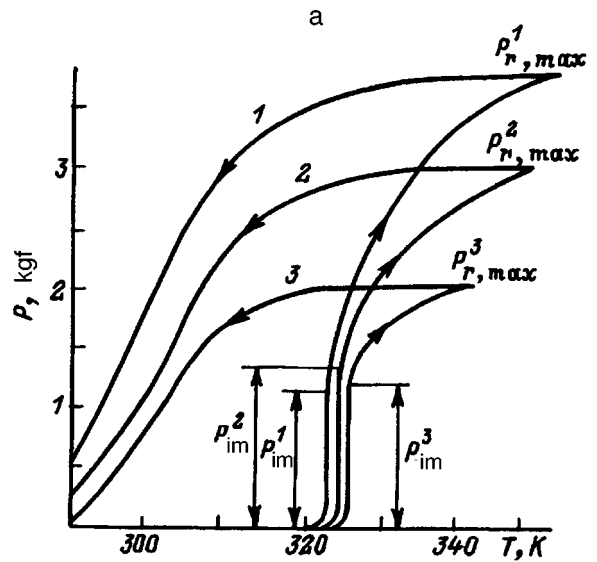

FIG. 4. Characteristic plots of the dependence of the impact force P_{im} and the reactive force P_r on the temperature for different values of the repairable deflection of the round plate from the center of curvature (f_i^{rd}) and a counteracting stiffness $K \to \infty$. a — $D=20$ mm, $2f=1.2$ mm; f_1^{rd}, mm: 1 — 0.48, 2 — 0.64, 3 — 0.83; b — $D=11$ mm, $2f=1.4$ mm; f_1^{rd}, mm: 1 — 0.78, 2 — 0.93, 3 — 1.2.

at the perimeter of the round plate and serves as the focus of the loss of stability. It must also be taken into account that, if the high-temperature pseudoelasticity is most pronounced near A_{fin}, the loss of stability is possibly also attributable to this effect, since the condition $A_{in} < T_{ls} \leqslant A_{fin}$ always holds.

An analysis of the plots of $P = f(f_i^{rd})$ showed (Fig. 5) that the clapping force, or the impact force of the disk, gradually increases, reaching its maximum value at a certain value of f_i^{rd} and then decreasing sharply as the latter increases. The distance f_i^{rd} corresponding to the strongest clap (the maximum impact force) is $f_{max}^{rd} = (0.65 - 0.75)2f$, where $2f$ is the deflection of the plate in the martensitic state. It is noteworthy that both the impact force P_{im} and the reactive force P generated in the material after a clap are determined by the geometric parameters of the disk (D, h, f, and R) and
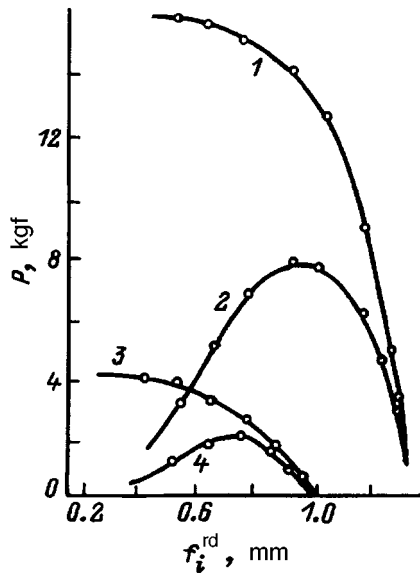
FIG. 5. Dependence of the impact force (2, 4) and the reactive force (1, 3) on the repairable deflection of the plate at the time of the clap. $D = 11$ (1, 2) and 20 mm (3, 4).

depend on the composition and structural state of the alloy.

Careful observations of the variation of the shape of the plates during thermal cycling showed that reversible shape memory is created. It must be assumed that the active deflection of a round plate in the martensitic state to a specified value of $2f$ promotes the appearance of oriented nuclei of the phase being formed and the appearance of microstresses in the direction of the force. Subsequent heating to T_{ls} causes

axisymmetric buckling of the plate with a clap, which is associated, as must be assumed, with the action of the oriented microstresses and partial repair of the strain. These events gives rise to a stable reversible shape memory effect, which makes possible the repeated alteration of the shape of the disk as the deflection is reduced.

Samples which underwent thermal conditioning during 160–200 cycles exhibited spontaneous buckling upon cooling and shape recovery with a clap in the heating stage.

CONCLUSIONS

1. The recovery of the shape of a plate preliminarily strained in the martensitic state occurs with a clap as a consequence of the loss of stability near the temperature A_{fin}.

2. A clap occurs when a definite relationship among the geometric parameters of the round plate (D, h, f, and R) is attained.

3. Thermal cycling through the range for martensitic phase transformations promotes the formation of repeatable reversible memory in the disk material.

4. The impact force of the disk increases as the fraction of the repairable strain increases and reaches its maximum value when $f_{max}^{rd} = (0.65 - 0.75)2f$.

[1] A. S. Vol'mir, *Stability of Deformable Systems* [in Russian], Nauka, Moscow (1987).

[2] M. A. Khusainov, V. N. Belyakov, and A. F. Pazgalov, in *Materials with Complicated Functional-Mechanical Properties* [in Russian], Novgorod (1994), Part 2, pp. 152–158.

Translated by P. Shelnitz

Nonuniformity of the transverse distribution of the resonant field excited by a beam of electromagnetic waves at the critical surface of a radially nonuniform plasma sphere

N. S. Bukhman

Michurinsk State Agricultural Academy, 393740 Michurinsk, Tambov Region, Russia

Simple approximate formulas are obtained for the characteristic length of the transverse amplitude–phase nonuniformity of the resonant field excited by a beam of electromagnetic fields at the critical surface of a spherical plasma body. © *1997 American Institute of Physics.*
[S1063-7842(97)02506-3]

This paper is devoted to the study of the transverse nonuniformity of the high-frequency resonant electric field excited at the critical surface of the body of a dense, nearly collisionless plasma by an electromagnetic beam.[1–3] This question is of interest, in partiuclar, in the study of the heating of a target by a light beam in laser fusion. Complete information on the transverse (at the critical surface) nonuniformity of the resonant field can be obtained by performing a complete calculation of the transverse distribution of the resonant field (for example, in accordance with the method described in Ref. 4). Unfortunately, employment of the results in Ref. 4 presumes the use of information on the refraction of the beam in the plasma corona and is impossible without the use of information (which is not always available) regarding the radial distribution of the plasma density far from the critical surface and the phase characteristics of the beam. Utilization of the results in Ref. 5, which are less demanding in regard to the original information (only the radius of curvature of the critical surface and the characteristic radial length for the nonuniformity of the plasma near the critical surface are used, and no data on the incident wave are employed), yields only information on the smallest possible (for an arbitrary electromagnetic wave) nonuniformity of the transverse distribution of the resonant field: the actual distribution of the resonant field can be significantly more uniform, depending on the characteristics of the incident wave.

The formulas presented in this communication permit finding the mean length for the amplitude–phase nonuniformity of the resonant field at the critical surface of a plasma sphere with the use of a minimum amount of information on the plasma sphere (the radius of curvature of the critical surface and the characteristic length for the nonuniformity of the plasma near the critical surface) and the electromagnetic beam (the vacuum distribution of the intensity of the beam field near the center of the plasma sphere).

Let us move on to a description of the results obtained. Let a paraxial beam of electromagnetic waves strike a plasma sphere with a density $n(r)$, the radius of curvature of the critical surface being equal to r_0 and the characteristic length for the nonuniformity of the plasma near the critical surface being equal to L (the plasma profile can differ from a linear one far from the critical surface). Let all the geometric dimensions be large in comparison with the wavelength ($k_0 L \gg 1$, and $k_0 r_0 \gg 1$, where $k_0 = 2\pi/\lambda_0$), and, in addition, let the condition

$$k_0 r_0 \gg (k_0 L)^{1/3}, \tag{1}$$

which is more rigid than the traditional condition $k_0 r_0 \gg 1$, hold. As was shown in Ref. 5, when condition (1) is violated, the estimates presented in Ref. 5 lose their "minimum" (with respect to the parameters of the incident wave) character, and the characteristic length for the transverse nonuniformity of the resonant field is equal to r_0, regardless of the characteristics of the beam.

Then, for the characteristic length $\Lambda^{(\mathrm{res})}$ for the nonuniformity of the resonant field excited by the beam at the critical surface of the plasma sphere, it is not difficult to obtain (in analogy to Ref. 6)

$$\Lambda^{(\mathrm{res})} = \frac{r_0 \lambda_0}{2} \frac{\iint I_\rho(\boldsymbol{\rho}) \widetilde{Q}(\rho/r_1) dS}{\iint I_\rho(\boldsymbol{\rho}) \widetilde{Q}(\rho/r_1) \rho \, dS}, \tag{2}$$

where

$$r_1 = r_0 (k_0 L)^{-1/3} \ll r_0, \tag{3}$$

and $\widetilde{Q}(\tau)$ is the repeatedly tabulated[7,8] resonant absorption coefficient of a plane electromagnetic wave at the flat critical surface of a one-dimensional nonuniform plasma layer, which is related to the Denisov function $\Phi_D(\tau)$ (Refs. 1 and 3) by the expression

$$\widetilde{Q}(\tau) = \Phi_D^2(\tau)/2. \tag{4}$$

The integration in Eq. (2) should be performed in the $0xy$ plane, which passes through the center of the plasma sphere perpendicularly to the vacuum axis of the beam (with consideration of the paraxial character of the beam, any of its "midlines" can be selected as the axis of the beam). The function $I(\boldsymbol{\rho})$ is the distribution of the vacuum intensity of the wave beam in the aforementioned plane, and $I_\rho(\boldsymbol{\rho})$ is the distribution of the vacuum intensity of the radially polarized component of the beam field in that plane

$$I_\rho(\boldsymbol{\rho}) = |\mathbf{E}_{\mathrm{vac}}(\boldsymbol{\rho}) \cdot \hat{\boldsymbol{\rho}}|^2, \quad I_\phi(\boldsymbol{\rho}) = |\mathbf{E}_{\mathrm{vac}}(\boldsymbol{\rho}) \cdot \hat{\boldsymbol{\phi}}|^2,$$

$$I(\boldsymbol{\rho}) = |\mathbf{E}_{\mathrm{vac}}(\boldsymbol{\rho})|^2 = I_\rho(\boldsymbol{\rho}) + I_\phi(\boldsymbol{\rho}), \tag{5}$$

where $\boldsymbol{\rho}=x\hat{\mathbf{x}}+y\hat{\mathbf{y}}$, $\rho^2=x^2+y^2$, $\hat{\boldsymbol{\rho}}=\boldsymbol{\rho}/\rho$, $\hat{\boldsymbol{\phi}}=(-x\hat{\mathbf{x}}+y\hat{\mathbf{y}})/\rho$, $\mathbf{E}_{vac}(\boldsymbol{\rho})$ is the amplitude of the vacuum field of the beam in the $0xy$ plane, and $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the unit vectors of the $0xy$ rectangular Cartesian coordinate system.

It should be noted that the vacuum parameters of the beam field are used in Eqs. (2) and (5), because these parameters specify the weight coefficients in the expansion of the incident wave in vector multipoles.[6]

Equalities (2)–(5) show that in the general case of a nonaxisymmetric beam of arbitrary polarization the characteristic length $\Lambda^{(res)}$ for nonuniformity depends only on the polarization and the distribution of the vacuum intensity of the incident field in the plane passing through the center of the plasma sphere perpendicular to the beam axis. It is also seen that only the vacuum field in the circle $\rho\lesssim r_1=r_0(k_0L)^{-1/3}\ll r_0$ is significant to the value of this parameter, because the function $\widetilde{Q}(\rho/r_1)$ is exponentially small at $\rho\gg r_1$. We note that the radius of this circle r_1 is small in comparison to the radius of curvature of the critical surface r_0.

Let us analyze the relation (2) obtained. We confine ourselves to an examination of the simplest case of an axisymmetric (with respect to the intensity), elliptically polarized beam, whose axis passes through the center of the plasma sphere $[I(\boldsymbol{\rho})=I(\rho)]$. In this case

$$\Lambda^{(res)}=\frac{r_0\lambda_0}{2}\frac{\int_0^\infty I(p)\widetilde{Q}(\rho/r_1)\rho d\rho}{\int_0^\infty I(p)\widetilde{Q}(\rho/r_1)\rho^2 d\rho}. \tag{6}$$

The integrals in Eq. (6) can be calculated exactly for each concrete beam profile $I(\rho)$, but we shall confine ourselves to an approximate evaluation of $\Lambda^{(res)}$ for various values of w/r_1, where w is the vacuum radius of the beam at the center of the plasma sphere.

For example, in the case of $w\ll r_1$, where the vacuum radius of the beam at the center of the sphere is small in comparison to r_1, we have

$$\Lambda^{(res)}\approx\pi r_0/(k_0 w). \tag{7}$$

In the case of $w\approx r_1$ or $w>r_1$, where the vacuum radius of the beam at the center of the sphere is of the same order as r_1 or greater than r_1 (even for a plane wave), we have

$$\Lambda^{(res)}\approx\lambda_0(k_0L)^{1/3}\gg\lambda_0. \tag{8}$$

In is not difficult to verify that in the case of $w\approx r_1$ estimates (7) and (8) coincide. Thus, they can be combined in the form

$$\Lambda^{(res)}\approx\max\{\pi r_0/(k_0 w), \quad \lambda_0(k_0L)^{1/3}\}. \tag{9}$$

As a result of the examination just performed, it can be concluded that sufficiently sharp vacuum focusing of the electromagnetic beam in the center of the plasma sphere ensures an increase in the uniformity of the angular distribution of the resonant field at the critical surface, regardless of the character of the reflection of the beam far from the critical surface of the plasma sphere. Since expression (9) gives an ''unbiased'' estimate for the amplitude–phase nonuniformity of the transverse distribution of the resonant field, it is a lower estimate of the characteristic length for the amplitude nonuniformity of the resonant field excited by the corresponding electromagnetic beam. One advantage of this estimate is the absence of a dependence on the character of the reflection of the electromagnetic beam far from the critical surface.

[1] V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas*, 2nd ed., Pergamon Press, Oxford–New York (1970).

[2] V. E. Golant and A. D. Piliya, Usp. Fiz. Nauk **104**, 413 (1971) [Sov. Phys. Usp. **14**, 413 (1971)].

[3] V. P. Silin, Usp. Fiz. Nauk **145**, 225 (1985) [Sov. Phys. Usp. **28**, 136 (1985)].

[4] N. S. Bukhman, Pis'ma Zh. Tekh. Fiz. **21**(4), 51 (1995) [Tech. Phys. Lett. **21**, 153 (1995)].

[5] V. S. Bukhman and N. S. Bukhman, Zh. Tekh. Fiz. **65**(2), 206 (1995) [Tech. Phys. **40**, 226 (1995)].

[6] N. S. Bukhman, Zh. Tekh. Fiz. **65**(2), 30 (1995) [Tech. Phys. **40**, 133 (1995)].

[7] O. Ya. Omel'chenko, V. I. Panchenko, and K. N. Stepanov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **14**, 1484 (1971).

[8] N. S. Bukhman, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **33**, 912 (1990).

Expansion of an arbitrary function in products of spherical functions

I. P. Skal'skaya, N. I. Zlatina, and I. B. Suslova

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

A formula for expanding an arbitrary function defined in the interval $(0,1)$ into a series in products of spherical Legendre functions of the first and second kinds is presented. Examples of expansions of this type are given. © *1997 American Institute of Physics.*
[S1063-7842(97)02606-8]

It was shown in Ref. 1 that a function defined in the interval $(0,1)$ can be expanded into a series in squares of Legendre polynomials:

$$\frac{1}{\sqrt{1-x^2}} \int_x^1 \frac{f(y)dy}{\sqrt{1-y^2}}$$

$$= 2 \sum_{n=0}^{\infty} (2n+1)P_n^2(x) \int_0^1 f(y)P_n(y)Q_n(y)dy, \qquad (1)$$

where $P_n(x)$ and $Q_n(x)$ are spherical Legendre functions of the first and second kind.

Conditions which are sufficient for this expansion to be correct were established, and several particular examples, which are of interest for the theory of special functions were considered. The purpose of the present work is to investigate an expansion, which, in a certain sense, can be regarded as symmetric with respect to Eq. (1), viz.,

$$\frac{1}{\sqrt{1-x^2}} \int_0^x \frac{f(y)dy}{\sqrt{1-y^2}}$$

$$= 2 \sum_{n=0}^{\infty} (2n+1)P_n(x)Q_n(x) \int_0^1 f(y)P_n^2(y)dy. \qquad (2)$$

Several results established in Ref. 1 are used during the investigation. To prove equality (2) we assume that $f(x) \in L(0,1)$, and we consider the sum

$$F_N(x) = 2 \sum_{n=0}^{N} (2n+1)P_n(x)Q_n(x) \int_0^1 f(y)P_n^2(y)dy, \qquad (3)$$

$$0 < x < 1, \quad N = 0, 1, 2, \ldots.$$

Since $|f(y)P_n^2(y)| \leq |f(y)|$ for $0 < y < 1$, the integral in Eq. (3) converges absolutely, and expression (3) has meaning. The sum under consideration can be written as

$$F_N(x) = \int_0^1 f(y)S_n(x,y)dy, \qquad (4)$$

where

$$S_N(x,y) = 2 \sum_{n=0}^{N} (2n+1)P_n(x)Q_n(x)P_n^2(y), \qquad (5)$$

$$0 < x < 1.$$

According to Eq. (32) in Ref. 1, the sum $S_N(x,y)$ converges as $N \to \infty$ to the finite limit

$$\lim_{N \to \infty} S_N(x,y) = 2 \sum_{n=0}^{\infty} (2n+1)P_n(x)Q_n(x)P_n^2(y)$$

$$= \begin{cases} \dfrac{1}{\sqrt{1-x^2}\sqrt{1-y^2}}, & y < x, \\ 0, & y > x. \end{cases} \qquad (6)$$

If admissibility of the limiting transition under the integral sign in Eq. (4) is postulated, we obtain

$$\lim_{N \to \infty} F_N(x) = \int_0^1 f(y) \lim_{N \to \infty} S_N(x,y)dy$$

$$= \frac{1}{\sqrt{1-y^2}} \int_0^x \frac{f(y)dy}{\sqrt{1-y^2}}$$

or, according to definition (3),

$$\frac{1}{\sqrt{1-x^2}} \int_0^x \frac{f(y)dy}{\sqrt{1-y^2}} = 2 \sum_{n=0}^{\infty} (2n$$

$$+1)P_n(x)Q_n(x) \int_0^1 f(y)P_n^2(y)dy. \qquad (7)$$

Thus, to complete the proof of the theorem it is sufficient to substantiate the permissibility of the limiting transition cited. For this purpose we utilize the asymptotic formulas for spherical functions obtained in Ref. 1:

$$P_n(\cos \theta) = \sqrt{\frac{2}{\pi(n+1/2)\sin \theta}} \sin[(n+1/2)\theta + \pi/4]$$

$$+ \frac{O(1)}{(n+1/2)^{9/10}(\sin \theta)^{9/10}},$$

$$Q_n(\cos \theta) = \sqrt{\frac{\pi}{2(n+1/2)\sin \theta}} \cos[(n+1/2)\theta + \pi/4]$$

$$+ \frac{O(1)}{(n+1/2)^{9/10}(\sin \theta)^{9/10}} \ln \frac{2}{1-\cos \theta}, \qquad (8)$$

where $0 < \theta \leq \pi/2$; $n = 0,1,2,\ldots$; and $O(1)$ does not depend on n and θ.

We write the product

$$4(n+1/2)P_n(\cos\theta)Q_n(\cos\theta)P_n^2(\cos\phi)=4(n+1/2)$$

$$\times\left\{\frac{\cos 2(n+1/2)\theta}{2(n+1/2)\sin\theta}+\frac{O(1)}{(n+1/2)^{7/5}(\sin\theta)^{9/5}}\ln\frac{2}{1-\cos\theta}\right\}\left\{\frac{1+\sin 2(n+1/2)\phi}{\pi(n+1/2)\sin\phi}+\frac{O(1)}{(n+1/2)^{7/5}(\sin\phi)^{9/5}}\right\}$$

$$=4(n+1/2)\left\{\frac{\cos 2(n+1/2)\theta}{2\pi(n+1/2)^2\sin\theta\sin\phi}+\frac{2(n+1/2)\phi\cos 2(n+1/2)\theta}{2\pi(n+1/2)^2\sin\theta\sin\phi}+\frac{O(1)[1+\sin 2(n+1/2)\phi]}{\pi(n+1/2)^{12/5}(\sin\theta)^{9/5}\sin\phi}\ln\frac{2}{1-\cos\theta}\right.$$

$$\left.+\frac{O(1)\cos 2(n+1/2)\theta}{2(n+1/2)^{12/5}\sin\theta(\sin\phi)^{9/5}}+\frac{O(1)}{(n+1/2)^{14/5}(\sin\theta)^{9/5}(\sin\phi)^{9/5}}\ln\frac{2}{1-\cos\theta}\right\}=\frac{2\cos 2(n+1/2)\theta}{\pi(n+1/2)\sin\theta\sin\phi}$$

$$+\frac{\sin 2(n+1/2)(\phi+\theta)+\sin 2(n+1/2)(\phi-\theta)}{\pi(n+1/2)\sin\theta\sin\phi}+\frac{O(1)[1+\sin 2(n+1/2)\phi]}{(n+1/2)^{7/5}(\sin\theta)^{9/5}\sin\phi}\ln\frac{2}{1-\cos\theta}$$

$$+\frac{O(1)\cos 2(n+1/2)\theta}{(n+1/2)^{7/5}\sin\theta(\sin\phi)^{9/5}}+\frac{O(1)}{(n+1/2)^{9/5}(\sin\theta)^{9/5}(\sin\phi)^{9/5}}\ln\frac{2}{1-\cos\theta}=\frac{2\cos 2(n+1/2)\theta}{\pi(n+1/2)\sin\theta\sin\phi}$$

$$+\frac{\sin 2(n+1/2)(\phi+\theta)+\sin 2(n+1/2)(\phi-\theta)}{\pi(n+1/2)\sin\theta\sin\phi}+\frac{O(1)}{(n+1/2)^{7/5}(\sin\theta)^{9/5}(\sin\phi)^{9/5}}\left\{(\sin\phi)^{4/5}[1\right.$$

$$\left.+\sin 2(n+1/2)\phi]\ln\frac{2}{1-\cos\theta}+(\sin\theta)^{4/5}\cos 2(n+1/2)\theta+\frac{1}{(n+1/2)^{2/5}}\ln\frac{2}{1-\cos\theta}\right\}=\frac{2\cos 2(n+1/2)\theta}{\pi(n+1/2)\sin\theta\sin\phi}$$

$$+\frac{\sin 2(n+1/2)(\phi+\theta)+\sin 2(n+1/2)(\phi-\theta)}{\pi(n+1/2)\sin\theta\sin\phi}+\frac{O(1)}{(n+1/2)^{7/5}(\sin\theta)^{9/5}(\sin\phi)^{9/5}},\qquad(9)$$

where $x=\cos\theta,\ 0\le x<1,\ 0<\theta\le\pi/2,\ \theta$ is fixed, and

$$\ln\frac{2}{1-\cos\theta}=O(1).$$

Thus,

$$4(n+1/2)P_n(\cos\theta)Q_n(\cos\theta)P_n^2(\cos\phi)$$

$$=\frac{2\cos 2(n+1/2)\theta}{\pi(n+1/2)\sin\theta\sin\phi}$$

$$+\frac{\sin 2(n+1/2)(\phi+\theta)+\sin 2(n+1/2)(\phi-\theta)}{\pi(n+1/2)\sin\theta\sin\phi}$$

$$+\frac{O(1)}{(n+1/2)^{7/5}(\sin\theta)^{9/5}(\sin\phi)^{9/5}},\qquad(10)$$

where $n=0,1,2,\ldots;\ 0<\theta\le\pi/2;\ 0<\phi\le\pi/2;$ and O does not depend on $\theta,\ \phi,$ and n.

Let us consider sums of the form $(x=\cos\theta,\ y=\cos\phi)$

$$S_N(x,y)$$

$$=2\sum_{n=0}^{N}(2n+1)P_n(\cos\theta)Q_n(\cos\theta)$$

$$\times P_n^2(\cos\phi).\qquad(11)$$

From (10) it follows that

$$S_N(x,y)=\frac{4}{\pi\sin\theta\sin\phi}\sum_{n=0}^{N}\frac{\cos(2n+1)\theta}{2n+1}$$

$$+\frac{2}{\pi\sin\theta\sin\phi}\sum_{n=0}^{N}\frac{\sin(2n+1)(\phi+\theta)}{2n+1}$$

$$+\frac{2}{\pi\sin\theta\sin\phi}\sum_{n=0}^{N}\frac{\sin(2n+1)(\phi-\theta)}{2n+1}$$

$$+\frac{O(1)}{(\sin\theta)^{9/5}(\sin\phi)^{9/5}}\sum_{n=0}^{N}\frac{1}{(2n+1)^{7/5}}$$

or

$$S_N(x,y)=\frac{4}{\pi\sin\theta\sin\phi}\sum_{n=0}^{N}\frac{\cos(2n+1)\theta}{2n+1}$$

$$+\frac{2}{\pi\sin\theta\sin\phi}\sum_{n=0}^{N}\frac{\sin(2n+1)(\phi+\theta)}{2n+1}$$

$$+\frac{2}{\pi\sin\theta\sin\phi}\sum_{n=0}^{N}\frac{\sin(2n+1)(\phi-\theta)}{2n+1}$$

$$+\frac{O(1)}{(\sin\theta)^{9/5}(\sin\phi)^{9/5}},\qquad(12)$$

where $0<\theta\le\pi/2,$ and $0<\phi\le\pi/2$.

From the previously obtained estimates

$$\sum_{n=0}^{N}\frac{\cos(2n+1)\theta}{2n+1}=O(1)\ln\frac{2}{1-\cos\theta}=O(1),\qquad(13)$$

$$\sum_{n=0}^{N}\frac{\sin(2n+1)(\phi+\theta)}{2n+1}=O(1)$$

for all $0<\phi\le\pi/2,\qquad(14)$

$$\sum_{n=0}^{N}\frac{\sin(2n+1)(\phi-\theta)}{2n+1}=O(1)$$

for all $0<\phi\leq\pi/2$. $\qquad(15)$

It follows from Eqs. (12)–(15) that

$$S_N(x,y)=\frac{4}{\pi\sin\theta\,\sin\phi}O(1)+\frac{O(1)}{\sin\theta\,\sin\phi}+\frac{O(1)}{\sin\theta\,\sin\phi}$$

$$+\frac{O(1)}{(\sin\theta)^{9/5}(\sin\phi)^{9/5}}=\frac{O(1)}{\sin\phi}+\frac{O(1)}{(\sin\phi)^{9/5}}$$

$$=\frac{O(1)}{(\sin\phi)^{9/5}}[1+(\sin\phi)^{4/5}]=\frac{O(1)}{(\sin\phi)^{9/5}},$$

$$0<\phi\leq\pi/2 \qquad(16)$$

or

$$S_N(x,y)=\frac{O(1)}{(1-y)^{9/10}},\quad 0\leq y<1, \qquad(17)$$

where x is a fixed number, and $0<x<1$.

If we assume that $f(y)(1-y)^{-9/10}\in L(0,1)$, the sequence $f(y)S_N(x,y)$, where $N=0,1,2,\ldots$, will be majorized, and the limiting transition under the integral sign will be permissible. The previously found condition $f(y)\in L(0,1)$ will then be satisfied. Thus, the following theorem has been proved. Let $f(x)(1-x)^{-9/10}\in L(0,1)$, then

$$\frac{1}{\sqrt{1-x^2}}\int_0^x\frac{f(y)dy}{\sqrt{1-y^2}}$$

$$=2\sum_{n=0}^{\infty}(2n+1)P_n(x)Q_n(x)\int_0^1 f(y)P_n^2(y)dy. \qquad(18)$$
$$\hspace{2cm}{\scriptstyle 0<x<1}$$

The following relations can serve as examples of the expansion of the type under consideration:

$$x=-2\sum_{n=0}^{\infty}\frac{P_n(x)Q_n(x)}{(2n-1)(2n+3)},\quad 0<x<1, \qquad(19)$$

$$x^3=-2\cdot3\sum_{n=0}^{\infty}$$

$$\times\frac{2n^2+2n-3}{(2n-1)(2n-3)(2n+3)(2n+5)}P_n(x)Q_n(x),$$

$$0<x<1, \qquad(20)$$

$$xF(-m,1,3/2,x^2)$$
$${\scriptstyle 0<x<1}$$

$$=\sum_{n=0}^{N}\frac{2n+1}{n+m+3/2}\cdot\frac{(-m-1/2)_n}{(m+3/2)_n}P_n(x)Q_n(x). \qquad(21)$$

The following more general formula is also valid

$$xF(1/2-\nu,1,3/2,x^2)$$

$$=\sum_{n=0}^{\infty}\frac{2n+1}{n+\nu+1}\frac{(-\nu)_n}{(1+\nu)_n}P_n(x)Q_n(x), \qquad(22)$$

where $\nu>-1$, $0<x<1$; $F(a,b,c,x)$ is a hypergeometric function;

$$1=-\sum_{n=0}^{\infty}(2n+1)\,_3F_2(-n,n+1,1/2;1,2,1)P_n(x)Q_n(x),$$

$$0<x<1; \qquad(23)$$

$_3F_2(a,b,c;e,f,x)$ is a generalized hypergeometric function;

$$x^2=1/2-1/\pi\sum_{n=1}^{\infty}\frac{(4n+1)\Gamma^2(n+1/2)}{(2n-1)(2n+2)(n!)^2}P_{2n}(x)Q_{2n}(x),$$

$$0<x<1; \qquad(24)$$

$$-\ln(1-x^2)=P_0(x)Q_0(x)+1/\pi\sum_{n=0}^{\infty}\frac{4n+3}{(n+1)(2n+1)}$$

$$\times\,_3F_2(n+1,-n-1/2,1/2;1,3/2,1)$$

$$\times P_{2n+1}(x)Q_{2n+1}(x),\quad 0<x<1; \qquad(25)$$

$$\frac{\arcsin x}{\sqrt{1-x^2}}=2\sum_{n=0}^{\infty}P_n(x)Q_n(x),\quad 0<x<1; \qquad(26)$$

$$K(x)=2\sum_{n=0}^{\infty}(-1)^nP_n(x)Q_n(x),\quad 0<x<1; \qquad(27)$$

$$E(x)=4\sum_{n=0}^{\infty}(-1)^n\frac{2n^2+2n-1}{(2n-1)(2n+3)}P_n(x)Q_n(x),$$

$$0<x<1; \qquad(28)$$

and $K(x)$ and $E(x)$ are complete elliptic integrals of the first and second kinds.

A more general formula is also valid:

$$\frac{\sqrt{\pi}}{2}\frac{\Gamma(1+\nu)}{\Gamma(3/2+\nu)}x^{2\nu+1}F(1/2,1+\nu,3/2+\nu,x^2)$$

$$=\sum_{n=0}^{\infty}(-1)^n\frac{2n+1}{n+\nu+1}\frac{(-\nu)_n}{(1+\nu)_n}P_n(x)Q_n(x),$$

$$\nu>-1,\quad 0<x<1. \qquad(29)$$

The coefficients in these expansions are calculated using the integral representation [see Eq. (22) in Ref. 1]

$$P_n^2(x)=\frac{1}{\pi}\int_{2x^2-1}^{1}\frac{P_n(t)dt}{\sqrt{(t-2x^2+1)(1-t)}},$$

$$-1<x<1,\quad n=0,1,2,\ldots, \qquad(30)$$

as well as the known expansions of a function in Legendre polynomials. Some of the formulas remain valid even at the ends of the interval considered.

[1] N. N. Lebedev and I. P. Skal'skaya, Differ. Uravn. **28**, 262 (1992).

Translated by P. Shelnitz

Calculation of the quantum yield of nonequilibrium carriers in an insulator under ionizing radiation

A. I. Nazarov and V. V. Sergeev

Petrozavodsk State University, 185640 Petrozavodsk, Russia

INTRODUCTION

Consideration of the quantum yield of free carriers when materials used in electronics are irradiated is of both theoretical and practical importance, particularly in connection with the correct simulation of the electronic processes leading to the formation of the radiation-induced charge in the gate insulator of a field-effect transistor. The solution of this problem[1,2] calls for taking into account the geminate recombination processes, which determine the final concentrations of the thermalized nonequilibrium carriers in the insulator.

GEMINATE RECOMBINATION

After completion of the exchange of energy, a thermalized electron and the corresponding hole are at a certain distance r_0 from one another. Two scenarios are then possible. One of them involves dissociation of the pair and the appearance of two thermalized nonequilibrium carriers, viz., a free electron and a hole. The other involves recombination of the electron and the hole with one another. Such recombination is generally called geminate or pair recombination.[2]

The probability of each of these processes depends on the ratio between the thermalization length r_0 and the Coulomb capture radius r_k. The Coulomb capture radius (or the Onsager radius) r_k can be determined by assuming that the Coulomb interaction energy of an electron and a hole will be equal to the thermal energy kT at that distance. Then we obtain

$$r_k = \frac{e^2}{4\pi\varepsilon\varepsilon_0 kT}, \qquad (1)$$

where e is the charge of the electron, ε is the permittivity of free space, ε_0 is the dielectric constant of the medium, k is Boltzmann's constant, and T is the temperature.

If $r_0 > r_k$, the electron-hole pair is free. If $r_0 < r_k$, dissociation of the pair is possible with a certain probability $f(r)$, and geminate recombination is possible with a probability $1 - f(r)$. According to Ref. 2, the probability of dissociation in the presence of an electric field F is

$$f(r) = \exp(r_k/r)(1 + \beta_0 E), \qquad \beta_0 = \frac{qr_k}{2kT} \qquad (2)$$

for weak fields and

$$f(r) = \exp\left(\frac{r_k}{r} \frac{\exp(-2\beta_0 Fr/r_k) - 1}{2\beta_0 Fr/r_k}\right) \qquad (3)$$

for strong fields.

Knowing the function $f(r)$ and the radial spatial distribution function $g(r)$ of the electrons, we can obtain the quantum yield $Y(E)$ of free carriers. Clearly, $Y(F)$ can be found by integrating the product of $f(r)$ and $g(r)$

$$Y(F) = \int_0^\infty f(r)g(r)dr. \qquad (4)$$

Several semiempirical approximations are used to calculate $g(r)$. Our goal was to find $g(r)$ and then to determine $Y(F)$ from first principles using the Monte Carlo method.

APPLICATION OF THE MONTE CARLO METHOD TO THE CALCULATION OF g(r)

Since $g(r)$ is the probability density of finding an electron near the ''original'' hole, it can be found numerically after determining the distance at which electrons escape from these holes and achieve complete thermalization. It is understood that we are interested in low-energy electrons, since they would otherwise traverse a distance significantly greater than r_k, and it would be meaningless to discuss geminate recombination.

Performing the Monte Carlo calculations requires knowledge of the characteristics of the electron scattering processes in the solid. The main processes for low-energy electrons include scattering on phonons. In these processes the electrons interact with polar optical phonons (LO phonons) and with nonpolar acoustic phonons. The intensity of the electron–phonon interaction is characterized by the scattering rate of the carriers. In the former case the scattering rate is determined from the following formula[3]

$$f_{LO}^{\pm} = \left(n + \frac{1}{2} \pm \frac{1}{2}\right) \frac{1}{\hbar^2} \left(\frac{m^*}{2E}\right)^{1/2} \frac{e^2}{4\pi\varepsilon_0}$$

$$\times \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon}\right) \hbar\omega_{LO} Ln \frac{1 + \left(1 \mp \frac{\hbar\omega_{LO}}{E}\right)^{1/2}}{\pm 1 \mp \left(1 \mp \frac{\hbar\omega_{LO}}{E}\right)^{1/2}}, \qquad (5)$$

where the plus sign refers to the emission, and the minus sign refers to the absorption (annihilation) of a LO phonon with an energy $\hbar\omega_{LO}$; m^* is the effective mass; ε_0, ε, and ε_∞ are the absolute, static, and optical dielectric constants; and n is the energy distribution function of the electrons

$$n = \left(\exp\left(\frac{\hbar\omega_{LO}}{kT}\right) - 1\right)^{-1}.$$
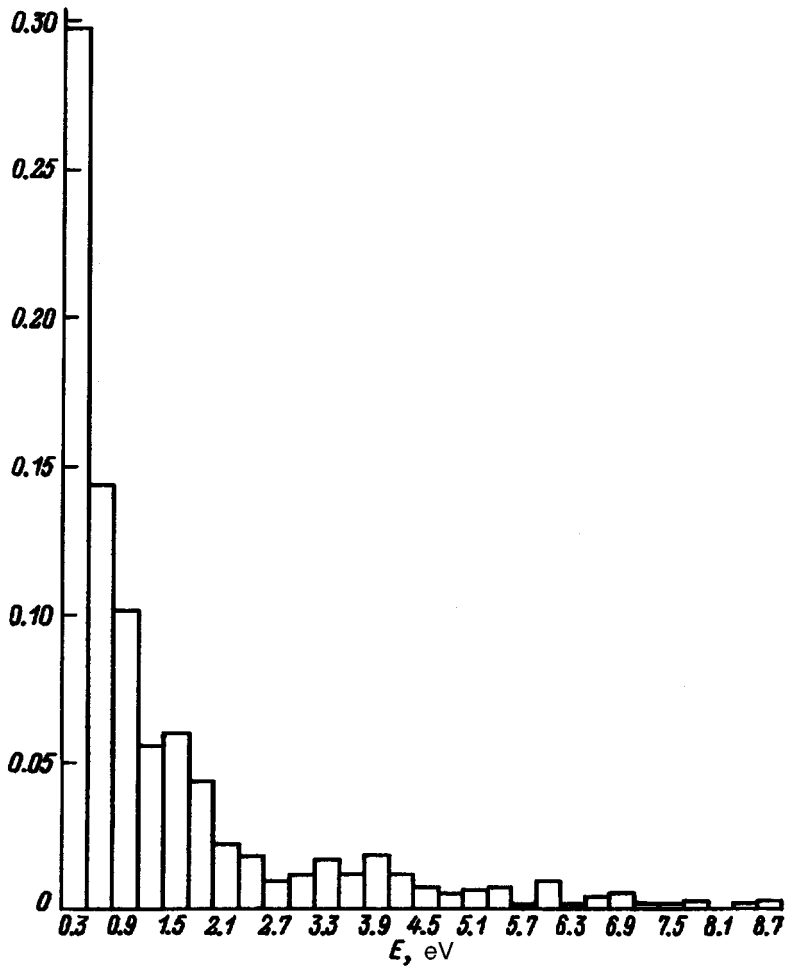
FIG. 1. Discrete energy distribution used to randomly select the initial electron energy. It was obtained from the results of the Monte Carlo simulation of the trajectories of electrons with an initial energy of 10 keV injected into an insulator.

In the case of scattering on acoustic phonons the scattering rate depends on the energy of the electrons on the Brillouin-zone edge ε_{BZ}. When $E \ll \varepsilon_{BZ}$, the rate of nonpolar acoustic scattering equals[4]

$$f^{\pm} \approx \frac{3m^{*3/2}C_1^2 kT}{\sqrt{2}\,\pi\rho C_s^2 \hbar^4}\sqrt{E}. \tag{6}$$

When $E > \varepsilon_{BZ}$,

$$f^{\pm} \approx \frac{8\pi^3\hbar^2 N^2\xi}{m^* M \omega_{BZ}}\left(\frac{E}{\varepsilon_{BZ}}\right)^{3/2}\left(n_{BZ}+\frac{1}{2}\pm\frac{1}{2}\right), \tag{7}$$

where C_1 is the deformation potential constant, ρ is the density, C_s is the velocity of sound, N is the number density of the lattice atoms, M is the mass of the heaviest atom in the unit cell, n_{BZ} and ω_{BZ} are the occupation numbers and frequencies of the acoustic phonons, and $\xi = 3.5 \times 10^{-15}$ cm^2 for silicon dioxide, a material that often serves as the gate insulator in transistors.

In our calculations we used the following values of the parameters: $m^* \approx m_0$, $T = 300$ K, $\hbar\omega_{LO} = 0.15$ and $\hbar\omega_{LO} = 0.06$ eV for different LO phonon modes, $\varepsilon = 3.8$, $\varepsilon_\infty = 2.25$, $C_1 = 3.5$ eV, $\rho = 2.65$ g/cm^3, $C_s = 4.03 \times 10^5$ cm/s, $N = 7.95 \times 10^{22}$ cm^{-3}, $M = 45.8 \times 10^{-27}$ kg, $\omega_{BZ} = 4.8 \times 10^{13}$ c^{-1}, $\varepsilon_{BZ} = 5.5$ eV, and $n_{BZ} = 0.816$.

The simulation scheme based on the general theory and the Monte Carlo method consisted of the following steps.

1. Selection of the initial parameters of the carrier: the coordinates x_0, y_0, and z_0, the velocity $\nu = (2E/m^*)^{1/2}$, and the energy E. For convenience all the trajectories were simulated from a single center. The angle of the initial motion of the electron was assumed to be uniformly distributed in the range from 0 to 2π. The initial energy was randomly selected from a distribution obtained from the results of a Monte Carlo simulation of the trajectories of moderate-energy electrons in SiO$_2$ on the basis of the data on scattering processes presented in Ref. 5. The number of electrons N having an energy below the ionization threshold ($\leqslant 8.9$ eV) after completion of the last inelastic collision was determined. This number was used to determine the number n of electrons having an energy in the range from E to $E + dE$. Then the probability p of the escape of an electron from a center with an energy in the range from E to $E + dE$ equals n/N. An energy-probability table served as a discrete distribution for randomly selecting the initial energy of the electron formed from the "initial" atom (Fig. 1).

2. Calculation of the probability densities of the electron–phonon scattering processes and determination of the time of flight Δt before the next collision on the basis of Eqs. (5)–(7).

3. Calculation of the coordinates of the electron[3]

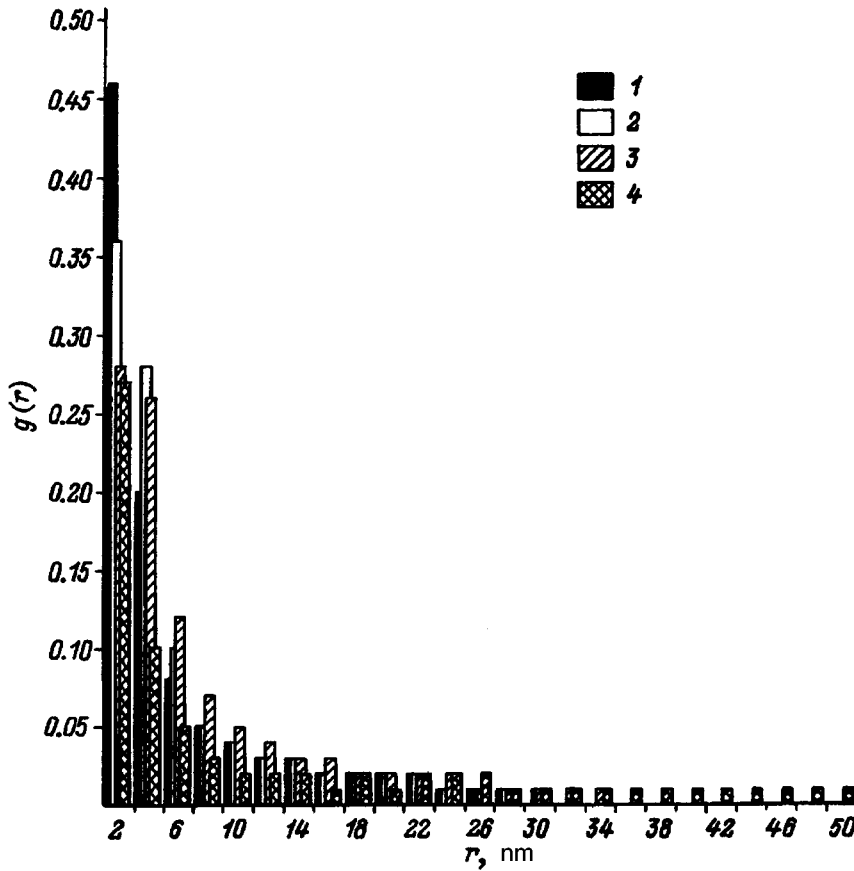$$x_n = x_{n-1} + \nu_{n-1}\sin\vartheta_{n-1}\cos\varphi_{n-1}\Delta t,$$

FIG. 2. Dependence of $g(r)$ on the distance r for various values of the electric field strength. E, V/cm: 1 — $1e+5$, 2 — $5e+5$, 3 — $1e+6$, 4 — $5e+6$.

$$y_n = y_{n-1} + \nu_{n-1} \sin \vartheta_{n-1} \sin \varphi_{n-1} \Delta t,$$

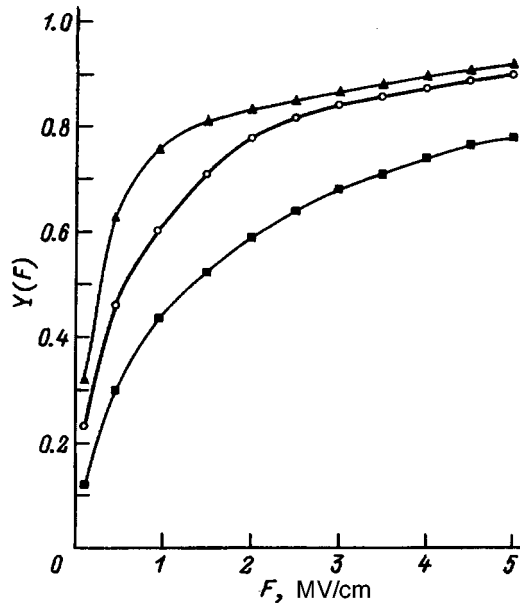$$z_n = z_{n-1} + \nu_{n-1} \cos \vartheta_{n-1} \Delta t + \frac{qF}{2m^*} \Delta t^2, \tag{8}$$

FIG. 3. Dependence of the quantum yield of free charge carriers on the electric field strength according to the results of our calculations and the experimental data in Ref. 6. E, keV: $\bigcirc$ — 10 (our calculations); $\blacksquare$ — 10 (Ref. 6), x rays; $\blacktriangle$ — Co-60 (Ref. 6).

where ϑ is the angle between the direction of motion and the z axis, and φ is the angle between the y axis and the projection of the direction of motion onto the xy plane.

The angle ϑ should henceforth be corrected[3]

$$\vartheta_n^F = \text{arccot}\left(\cot \vartheta_{n-1} + \frac{qF}{2E_{n-1}\sin^2 \vartheta_{n-1}} \right.$$

$$\left. \times \sqrt{(x_n - x_{n-2})^2 + (y_n - y_{n-1})^2} \right). \tag{9}$$

4. Determination of the energy after the interaction: consideration of the energy change due to the electric field

$$E_n^F = E_{n-1} + qF(z_n - z_{n-1}), \tag{10}$$

and calculation of the energy after the interaction

$$E_n = E_n^F \pm \hbar \omega. \tag{11}$$

5. Finding the scattering angles.
The polar scattering angle is given by the expression[3]

$$\cos \alpha_n = \frac{E_n^F + E_n}{2\sqrt{E_n^F E_n}}(1 - B^R) + B^R, \tag{12}$$

where R is a random number that is uniformly distributed in the interval $(0,1)$, and

$$B = \frac{E_n^F + E_n + 2\sqrt{E_n^F E_n}}{E_n^F + E_n - 2\sqrt{E_n^F E_n}}. \tag{13}$$

The azimuthal angle β is uniformly distributed in the interval $(0, 2\pi)$ and is given by the expression

$$\beta = 2\pi R, \tag{14}$$

where R is a random number.

The new angles ϑ and φ were calculated from formulas which were presented in Ref. 3:

$$\cos\vartheta_n = \cos\alpha_n\cos\vartheta_n^F + \sin\alpha_n\cos\beta_n\sin\vartheta_n^F, \tag{15}$$

$$\varphi_n = \varphi_{n-1} + \arctan\frac{\sin\alpha_n\sin\beta_n}{\cos\alpha_n\sin\vartheta_n^F}. \tag{16}$$

The procedure described above was repeated from step 2 until the electron energy dropped to less than 0.06 eV. The functions $g(r)$ were found from the results of these calculations (Fig. 2), and the quantum yields of the free carriers were determined for various fields (Fig. 3).

[1] V. A. Gurtov, A. I. Nazarov, and I. V. Travkov, Fiz. Tekh. Poluprovodn. **24**, 969 (1990) [Sov. Phys. Semicond. **24**, 611 (1990)].
[2] A. V. Vannikov, V. P. Matveev, V. L. Sichkar', and A. P. Tyunev, *Radiation Effects in Polymers. Electrical Properties* [in Russian], Nauka, Moscow (1982).
[3] H.-J. Fitting and J.-U. Frieman, Phys. Stat. Sol., 349 (1982).
[4] G. V. Gadiyak, S. P. Sinitsa, and I. V. Travkov, Mikroelektronika **17**, 448 (1988).
[5] J. C. Ashley and V. E. Anderson, J. Electron Spectrosc. Relat. Phenom. **24**, 127 (1981).
[6] J. M. Benedetto and H. E. Boesch, IEEE Trans. Nucl. Sci. **NS-33**, 1318 (1986).
[7] I. V. Travkov and V. A. Shveĭgert, *Autometry* [in Russian], Novosibirsk, (1988), pp. 67–73.
[8] N. Ozturk and W. Williamson, J. Appl. Phys. **74**, 1 (1993).
[9] D. G. Esaev and S. P. Sinitsa, Pis'ma Zh. Tekh. Fiz. **6**, 1063 (1986) [Sov. Tech. Phys. Lett. **12**, 440 (1986)].

Translated by P. Shelnitz

Spin polarization of helium 2^3S_1 atoms in a Na−He gas-discharge plasma with optically oriented sodium atoms

S. P. Dmitriev and D. É. Denisov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted June 3, 1996)
Zh. Tekh. Fiz. **67**, 131–133 (June 1997)

The polarization of helium 2^3S_1 atoms is achieved by collisional processes involving optically oriented sodium atoms in a Na−He gas-discharge plasma, and the conditions for observing the signal of the magnetic resonance excited by the 2^3S_1 state of the He atoms are investigated. © *1997 American Institute of Physics.* [S1063-7842(97)02806-7]
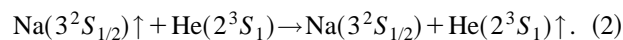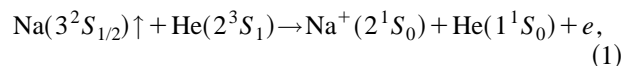
Investigations of atomic collisions involving spin-polarized particles under the conditions of an alkali-metal/helium gas-discharge plasma are of interest both for the physics of the spin-dependent processes occurring in atomic collisions and for applied problems of quantum magnetometry. The dependence of the probability of Penning ionization during the collisions of alkali metal atoms and metastable helium atoms on the mutual orientation of the electron spin magnetic moments of the partners was demonstrated experimentally and investigated in Refs. 1 and 2. An influence of the spin orientation of the atoms under the conditions of an alkali-metal/helium plasma on its electrical conductivity was discovered. There is special interest in research on the transfer of spin orientation from optically oriented alkali-metal atoms to the products of an alkali-metal/helium plasma, viz., helium 2^3S_1 metastable atoms and free electrons.[3,4] The possibility of orienting the spin magnetic moments of helium 2^3S_1 atoms in the absence of direct optical pumping by resonant helium light permits elimination of the large photoinduced frequency shifts of the magnetic resonance excited in the system of Zeeman sublevels of the 2^3S_1 state of the He atoms, which is important for solving the quantum magnetometry problems. A novel type of quantum device, viz., a helium magnetometer with optically oriented alkali-metal atoms or an alkali-metal/helium magnetometer, was proposed in Refs. 5 and 6. Helium 2^3S_1 atoms have been polarized by transferring spin orientation from optically oriented alkali-metal atoms, and devices have been created under the conditions of Cs–He, Rb–He, and K–He gas-discharge plasmas.[5,7,8] The high accuracy and weak orientational dependence of the readings place them among the most promising devices for measuring the geomagnetic field and its variation.

In the present work we polarized helium 2^3S_1 atoms in collisional processes involving optically oriented sodium atoms in a Na–He gas-discharge plasma and investigated the conditions for observing the signal due to the magnetic resonance excited in the 2^3S_1 state of the He atoms.

The experimental procedure was as follows. Metastable helium atoms were excited in the Na–He plasma of a radio-frequency (rf) pulsed discharge. The pulse repetition rate was 1–10 kHz, the pulse duration was 1–10 μs, and the filling frequency was 45 MHz. The discharge was excited in a cylindrical glass chamber (with a length of 50 mm and a diam-

eter of 50 mm) with an internal glass-coated electrode. (The rf voltage on the electrode was $\simeq 100$ V). The chamber contained metallic sodium and gaseous ^{4}He at a pressure of 1 Torr. The chamber was placed in a glass thermostat that provided for heating to 150 °C, which enabled us to achieve a Na vapor density corresponding to an atom number density $\simeq 10^{11}$cm^{-3}. The thermostat together with the chamber was placed in a multilayer magnetic shield, within which a field $H_0 = 20-50$ mOe directed along the axis of the shield was created using Helmholtz coils. The Na atoms were optically oriented by circularly polarized light from a resonant sodium lamp (the D$_1$ and D$_2$ spectral lines of the Na atoms), which propagated parallel to the field H_0. To excite the magnetic resonance in the systems of Zeeman sublevels of the $3^2S_{1/2}$ ground state of the Na atoms and the 2^3S_1 metastable state of the He atoms, an amplitude modulated rf magnetic field $H_1 = A \sin \Omega t \sin \omega t$ (the modulation frequency was $\Omega = 70$ Hz, and the frequency of the rf field was $\omega = \gamma_A H_0$, where γ_A is the gyromagnetic ratio: $\gamma_{Na} \simeq 0.7$ Hz/Oe and $\gamma_{He} \simeq 2.8$ Hz/Oe, respectively) was applied perpendicular to H_0. The magnetic resonance (for both the sodium $3^2S_{1/2}$ atoms and the helium 2^3S_1 atoms) was detected in the variation of the intensity of the circularly polarized light of the sodium pump lamp passing through the chamber. A silicon photodiode, a narrow-band amplifier, a synchronous detector, and an $X-Y$ recorder were employed for this purpose. To reduce the noise in the optical system, a broad-band glass filter, which cut off the spectral lines lying far from the region of the sodium doublet, was placed in front of the photodiode.

As a result of the optical orientation of the Na atoms and the collisions of these atoms with the helium 2^3S_1 metastable atoms, the latter become spin-polarized. The most important mechanism for transferring orientation are spin-dependent Penning ionization of Na atoms (1) and spin-exchange collisions (2):

$$\text{Na}(3^2S_{1/2})\uparrow + \text{He}(2^3S_1) \rightarrow \text{Na}^+(2^1S_0) + \text{He}(1^1S_0) + e, \quad (1)$$

$$\text{Na}(3^2S_{1/2})\uparrow + \text{He}(2^3S_1) \rightarrow \text{Na}(3^2S_{1/2}) + \text{He}(2^3S_1)\uparrow. \quad (2)$$

In reaction (1) the spin polarization of the helium 2^3S_1 atoms appears because of the removal of the atoms for which Penning ionization is allowed by the total spin magnetic moment conservation law from the ensemble of 2^3S_1 atoms.[1,5] A significant contribution to the polarization of the helium
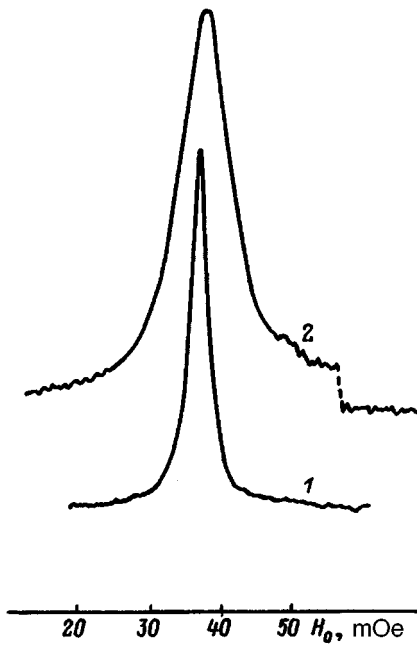
FIG. 1. Signals of the magnetic resonance excited in the $3^2S_{1/2}$ state of Na atoms (*1*) and in the 2^3S_1 state of He atoms (*2*): $\omega/2\pi = 25$ (*1*), 100 kHz (2).

FIG. 2. Dependence of the amplitude of the signals S_{Na} (*1*) and S_{He} (*2*) observed on optical orientation of the Na atoms in a Na–He gas-discharge plasma ($f_{rep} = 2$ kHz, $\tau_{pul} = 3$ μs) on the chamber temperature.

2^3S_1 atoms can be made by the collisions of these atoms with electrons polarized during spin exchange with optically oriented sodium atoms.[3,8]

Disorientation of the helium 2^3S_1 atoms by the magnetic resonance excited in the system of Zeeman sublevels of the metastable He atoms leads to variation of the equilibrium polarization of the ensemble of spin-polarized plasma particles, including sodium atoms, which can be detected in the variation of the absorption of the circularly polarized pump light from the resonant Na lamp.

Figure 1 presents the signals of the magnetic resonance excited in the $3^2S_{1/2}$ state of the Na atoms (S_{Na}) and in the 2^3S_1 state of the He atoms (S_{He}). The width of the signals is governed mainly by the nonuniformity of the field H_0 and amounts to $\simeq 1$ (Na) and $\simeq 4$ kHz (He). These values correspond to the ratio between γ_{Na} and γ_{He}. The maximum value of S_{He} determined as the amplitude of the first harmonic (with the frequency Ω) with a 50 kΩ load on the photodiode, amounted to $\simeq 1$ mV when the signal-to-noise ratio was $\simeq 50$ (in a 1 Hz band). The amplitude of S_{Na} was approximately 30 times greater than that of S_{He}.

The dependence of the amplitudes of the observed signals S_{Na} and S_{He} on the temperature and the parameters of the pulsed discharge was investigated in this work (Fig. 2). The optimal conditions for observing S_{Na} and S_{He} with the chamber under consideration approximately coincide and fall in range $110-130$ °C. The increase in S_{Na} and S_{He} with the temperature is due to an increase in the concentration of optically oriented Na atoms. The decrease in S_{Na} and S_{He} at temperatures above $\simeq 120$ °C is due to a decrease in the intensity of the detecting light as a result of an increase in the optical density of the Na vapor in the chamber. The dependence of the amplitude of the helium signal on the parameters of the pulse di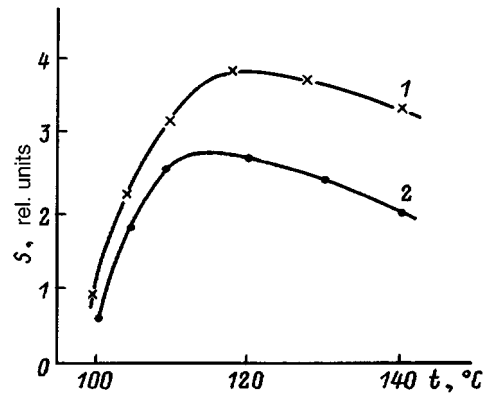scharge (on the pulse repetition rate f_{rep} in the $1-10$ kHz range and the pulse duration τ_{pul} in the $1-10$ μs range) was investigated at $\simeq 120$ °C. Typical curves are presented in Fig. 3. The highest amplitudes of S_{He} were observed for a pulse duration equal to $3-7$ μs and a pulse repetition rate equal to $2-3$ kHz. No significant influence of the temperature and the parameters of the pulsed discharge on the width of S_{He} was observed under our conditions due to the large inhomogeneous broadening of the signals and a signal-to-noise ratio that is insufficient for exact measurements.

Additional investigations are needed to determine the influence of the temperature, the parameters of the pulsed discharge, and the helium pressure in the chamber on the width of S_{He}. In addition, a significant increase in the amplitude of S_{He} should be expected when the chamber is filled with helium to a pressure of $5-10$ Torr (see, for example, Ref. 8).
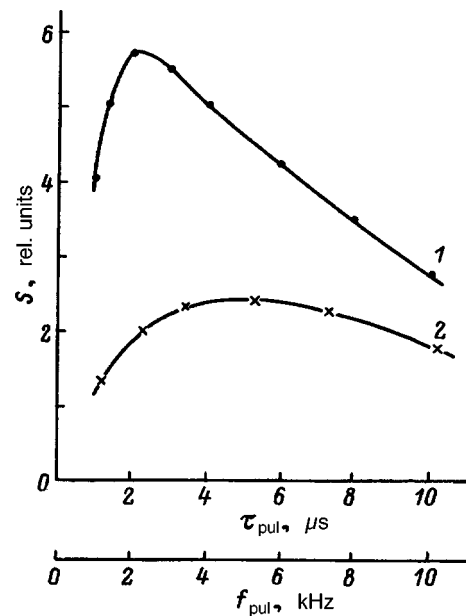

FIG. 3. Dependence of the amplitude of the helium signal on the parameters of the pulsed discharge: *1* — dependence of S_{He} on the rf pulse repetition rate ($\tau_{pul} = 3$ μs), *2* — dependence of S_{He} on the pulse duration ($f_{rep} = 2$ kHz). The chamber temperature was $\simeq 120$ °C.

However, in that case the influence of collisional mixing in the excited states of the alkali-metal atoms leads to a need to filter the D_1 and D_2 lines in the pump spectrum.[9] When K, Rb, and Cs are employed, this is easily accomplished using interference filters. However, the D_1 and D_2 lines of the sodium doublet are close to one another [$\lambda(D_1) = 589.0$ nm and $\lambda(D_2) = 589.6$ nm], and they cannot be separated by ordinary filters without a significant loss of intensity. The use of a tunable dye laser, which permits the performance of experiments with separate high-intensity D_1 and D_2 lines, seems promising for the optical orientation of sodium atoms.

Thus, the spin polarization of helium 2^3S_1 metastable atoms has been achieved in a Na–He gas-discharge plasma with optical orientation of the sodium atoms in this work, and the optimal temperature and parameters of the pulsed discharge for the signal amplitude have been determined.

[1] S. P. Dmitriev, R. A. Zhitnikov, and A. I. Okunevich, Zh. Éksp. Teor. Fiz. **70**, 69 (1976) [Sov. Phys. JETP **43**, 35 (1976)].

[2] S. P. Dmitriev, R. A. Zhitnikov, and A. I. Okunevich, in *Abstracts of the 20th Congress AMPÈRE*, Tallin, 1978, p. 418.

[3] G. M. Keiser, H. G. Robinson, and C. E. Jonson, Phys. Lett. A **51**, 5 (1975).

[4] E. V. Blinov, R. A. Zhitnikov, and P. P. Kuleshov, Pis'ma Zh. Tekh. Fiz. **2**, 305 (1976) [Sov. Tech. Phys. Lett. **2**, 117 (1976)].

[5] E. V. Blinov, R. A. Zhitnikov, and P. P. Kuleshov, Zh. Tekh. Fiz. **49**, 588 (1979) [Sov. Phys. Tech. Phys. **24**, 336 (1979)].

[6] E. V. Blinov, R. A. Zhitnikov, and P. P. Kuleshov, USSR Inventor's Certificate (Patent) No. 578 630; Byull. Izobret. No. 40 (1977).

[7] E. V. Blinov, B. I. Ginzburg, R. A. Zhitnikov, and P. P. Kuleshov, Zh. Tekh. Fiz. **54**, 287 (1984) [Sov. Phys. Tech. Phys. **29**, 168 (1984)].

[8] E. V. Blinov, B. I. Ginzburg, R. A. Zhitnikov, and P. P. Kuleshov, Zh. Tekh. Fiz. **54**, 2315 (1984) [Sov. Phys. Tech. Phys. **29**, 1362 (1984)].

[9] M. Elbel and F. Naumann, Z. Phys. **204**, 501 (1967).

Translated by P. Shelnitz

Experimental investigation of beam auto-oscillations in a section of a backward-wave linear accelerator

V. V. Kozlyuk

Institute of Chemical Kinetics and Combustion, Russian Academy of Sciences, Siberian Branch,
630090 Novosibirsk, Russia

Auto-oscillations of the current of an accelerated beam were discovered experimentally relatively recently in the accelerating structure of a one-section backward-wave linear accelerator.[1,2] These auto-oscillations are undesirable for the operation of the accelerator, since they lower its efficiency and adversely affect the output beam quality.

The physical mechanism of the beam auto-oscillations beam in a section of the backward-wave linear accelerator was discussed in Ref. 1 and 2. It would be useful to focus once again on this question. The longitudinal distribution of the accelerating field in a section of a backward-wave linear accelerator has an exponentially decaying character. The particles are injected toward the flow of rf power from the side where it leaves the accelerating structure. When the section is loaded with an insignificant beam current, the distribution of the field is similar to the calculated distribution, under which synchronous acceleration of the particles is ensured. If the section is overloaded by the injection current (the auto-oscillation regime), the rf power is converted almost entirely into beam power in the first cells of the accelerating structure along the course of the beam. The level of rf power remaining after this conversion is not sufficient for ensuring synchronous acceleration of the next particles, and they escape the acceleration ''bucket''. After the synchronous particles leave the section, the original distribution of the accelerating field begins to be restored in it. When the field strength is close to the calculated value, the next portion of particles will be caught in the accelerating ''bucket''. This process then continues periodically until the injection of particles is stopped. The beam current is composed of a sequence of two groups of particles, viz., accelerated and unaccelerated particles. In other words, the time intervals during which particle acceleration occurs alternate with intervals during which the particles escape from the accelerating ''bucket''. The period of this process[1–3] is equal to the sum of the time-of-flight of the particles caught in the accelerating ''bucket'' and the time needed for the rf power to fill the accelerating structure after the synchronous particles have left it:

$$T = L(1/V_{ph} + 1/V_g), \tag{1}$$

where L is the length of the section, and V_{ph} and V_g are, respectively, the mean values of the phase and group velocities along the length of the accelerating structure.

A diagram describing the acceleration of particles in a backward-wave linear accelerator was presented in Ref. 4. An expression relating the conditions for the appearance of auto-oscillations to the parameters of the accelerator can be derived from a treatment of this diagram under the following assumptions: 1) the same portion of rf power is absorbed in each cell of the section, 2) there are no beam losses in the acceleration process, and 3) the output rf power is close to zero in the auto-oscillation regime.[1,2] Under these conditions

$$\frac{\Delta I}{I^0} = \frac{Pe/(I^0 \Delta W)}{N^{-1} \Sigma_{i=1}^{N} \Pi_{j=i+1}^{N} \exp(-2\alpha_j l_j)}, \tag{2}$$

where $\Delta I = I - I^0$ is the deviation of the current of the accelerated beam relative to its design value I^0; P is the output rf power corresponding to loading the accelerating structure with the current I^0; ΔW is the energy gained by the particles in the section; e is the electron charge; α_j is the damping of the rf power in the jth cell of the accelerating structure over the distance l_j; N is the number of accelerating cells; and $i = 1, 2, \ldots, N$.

The denominator in the fraction clearly has lower and upper bounds defined by the inequalities

$$N^{-1} < N^{-1} \sum_{i=1}^{N} \prod_{j=i+1}^{N} \exp(-2\alpha_j l_j) < 1.$$

Then, for a fixed value of I_0 and $I^0 \Delta W / e \gg P$ it is seen that the requirements placed on the stability of the beam current increase significantly as the output energy of the accelerator increases.

Expression (2) was tested experimentally using a section of the backward-wave linear proton accelerator that was developed and constructed in the Institute of Chemical Kinetics and Combustion of the Siberian Branch of the Russian Academy of Sciences. The section was fabricated on the basis of an accelerating structure of the counterpivot type. Its period increases from approximately 2 to 3 mm. The aperture of the accelerating channel is ~ 3 mm. The working frequency equals 1818 MHz. The total length of the accelerating structure is 1.25 m, in which the relative damping varies from the injection end to the exit from 0.56 to 0.21 m^{-1}, respectively. Some of the design parameters of the section are presented in Table I.

This section was investigated using a pulsed beam by electron modeling.[5] The design parameters of the electron model are presented in Table 1. The strength of the focusing magnetic field in the model was set at a level ~ 4 times greater than the design value in order to overcome the transverse deflection of the electrons in the Earth's magnetic field and to ensure good passage of the current through the aperture of the accelerating structure. The output current was

TABLE I.

Parameter name	Design parameters of the section	
	Proton accelerator	Electron model
Injection energy	385 keV	210 eV
Output energy	1.9 MeV	1035 eV
Accelerated current	64 mA	35 μA
Input rf power	180 kV	54 mV
Output rf power	30 kV	9 mV
Focusing field	8.5 T	46 G
Mean phase velocity	1.38×10^7 m/s	1.38×10^7 m/s
Mean group velocity	1.54×10^7 m/s	1.54×10^7 m/s

FIG. 2. Dependence of the accelerated beam current on the input rf power for various values of the injection voltage. U_{inj}/U^0: $\bigcirc$ — 1.05, $\triangle$ — 1.1, $\square$ — 1.15, + — 1.2; P_0 and U^0 are the design values of the input rf power and the injection voltage, respectively, and P_{in} is the input rf power.

measured using a secondary-electron multiplier.[6] A stopping potential of ~ 1000 V was supplied to the guard ring of the secondary-electron multiplier.

When the injection currents are significantly higher than the design values, steady beam auto-oscillations appear in the section (Fig. 1). The period of these auto-oscillations is ~ 160 ns. An estimate of the period based on Eq. (1) gives $T \sim 170$ ns, which differs only slightly from the experimentally observed value. To determine the critical beam currents at which steady auto-oscillations appear in the section, we performed a series of measurements of these currents for various values of the input rf power and the injection energy. The results of these measurements are presented in the form of plots in Fig. 2. It is seen that for values of the rf power close to the design values the value of the critical current is 47 ± 6 μA, which is greater than its design value by $34 \pm 15\%$. An estimate of the critical beam current based on Eq. (2) gives $I/I^0 = 1.46$, which agrees closely with the experimentally obtained values of the currents. It follows from the material presented here and the results published in Refs. 1 and 2 that auto-oscillations appear when the beam current
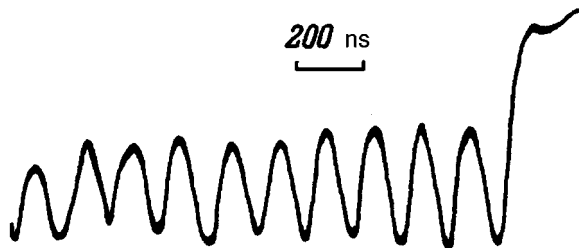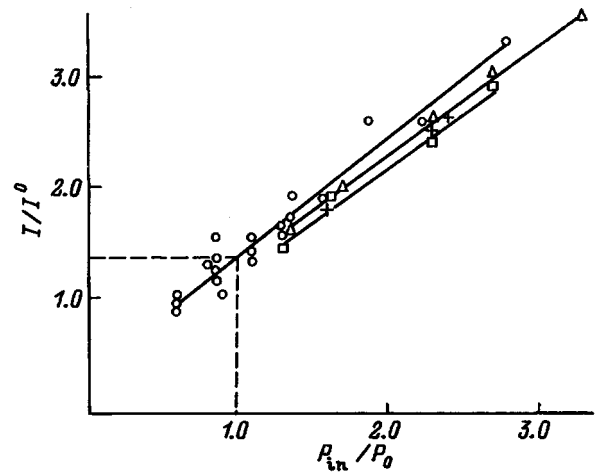
exceeds a certain critical limit, which is characteristic of each of the accelerating structures investigated.

Thus, the following conclusions can be drawn from this work.

1. The beam auto-oscillations in a section of a backward-wave linear accelerator can be characterized as a regular phenomenon for the backward-wave acceleration of charged particles at beam currents above the critical values.

2. The simplified mathematical expression (2) agrees well with the available experimental material and can be used to estimate the critical beam currents.

3. The results obtained from investigating the auto-oscillations in a section with $\Delta W = 500$ keV (Refs. 1 and 2) and the results of the present work allow us to state that the requirements placed on the stability of the beam current increase as the output energy of a one-section backward-wave linear accelerator rises.

[1] V. V. Kozlyuk, Vopr. At. Nauki Tekh., Ser. Yad.-Fiz. Issled. (6/6), 67 (1989).
[2] V. V. Kozlyuk, Pis'ma Zh. Tekh. Fiz. **17** (1), 75 (1991) [Sov. Tech. Phys. Lett. **17**, 29 (1991)].
[3] G. A. Kirpichnikov and A. N. Tarasovskiĭ, Preprint No. 26 of the Institute of Chemical Kinetics and Combustion, Russian Academy of Sciences, Siberian Branch, Novosibirsk (1989).
[4] A. S. Bogomolov, Dokl. Akad. Nauk SSSR **208**, 1328 (1973) [Sov. Phys. Dokl. **18**, 152 (1973)].
[5] V. V. Kozlyuk, Candidate's Dissertation, Moscow (1992).
[6] A. N. Borisenko, V. V. Kozlyuk, and V. V. Perov, Prib. Tekh. Éksp. (2), 133 (1988).

Translated by P. Shelnitz

200 ns

FIG. 1. Oscillogram of the auto-oscillations of the accelerated beam current.

Modification of the polarization-holographic method for partial polarization of a field of electromagnetic waves

Sh. D. Kakichashvili and B. N. Kilosanidze

Institute of Cybernetics, Georgian Academy of Sciences, 380028 Tbilisi, Georgia

[S1063-7842(97)03006-7]

In Refs. 1 and 2 we considered some cases of the application of the Jones method modified for the partial polarization of light in holographic problems.[3,4] In the present paper the polarization-holographic recording and reconstruction of images using partially polarized radiation are analyzed theoretically. The state and degree of polarization of the undiffracted beam and the imaginary and real images formed by a polarization hologram are also considered.

Let there be a partially elliptically polarized wave propagating along the z axis. The modified Jones vector of this wave can be represented in the form of the orthogonal elliptical polarization basis[2,5]

$$\mathbf{E}_{\mathrm{op}} = E_{AX} \exp i(\omega t + \varphi) \begin{pmatrix} 1 \\ i\varepsilon \end{pmatrix}$$
$$\oplus E_{BY} \exp i\left(\omega t + \Psi - \frac{\pi}{2} \right) \begin{pmatrix} i\varepsilon \\ 1 \end{pmatrix}, \qquad (1)$$

where

$$\varepsilon = \frac{E_{AY}}{E_{AX}} = \frac{E_{BX}}{E_B} \quad (0 \leqslant \varepsilon \leqslant 1),$$

the symbol $\oplus$ was introduced to denote incoherent summation of the amplitudes (the corresponding rules for working with $\oplus$ are defined in Ref. 2), $\mathbf{E}_A \exp i\varphi$ is the complex amplitude of a component of one basis, and $\mathbf{E}_B \exp i\Psi$ is the complex amplitude of a component of another basis, which is orthogonal and incoherent to the first.

The passage of (1) through an arbitrary (anisotropic gyrotropic) object results in the formation of an object wave, whose modified Jones vector is written in the form

$$\mathbf{E}_{\mathrm{obj}} = E_{AX} \exp i(\omega t + \varphi + \delta) M_{\mathrm{obj}} \begin{pmatrix} 1 \\ i\varepsilon \end{pmatrix}$$
$$\oplus E_{BY} \exp i\left(\omega t + \Psi - \frac{\pi}{2} + \delta \right) M_{\mathrm{obj}} \begin{pmatrix} i\varepsilon \\ 1 \end{pmatrix}, \qquad (2)$$

where

$$M_{\mathrm{obj}} = \begin{pmatrix} \hat{m}_{11} & \hat{m}_{12} \\ \hat{m}_{21} & \hat{m}_{22} \end{pmatrix}$$

is the complex Jones matrix of the object,[6] and δ is the phase difference caused by the oblique propagation of the object wave.

In the polarization-holographic recording process, each orthogonal component interferes independently with the component that is coherent to it, and the resultant fields are combined incoherently and additively. The total field in the plane of the hologram has the form

$$\mathbf{E}_\Sigma = E_{AX} \exp i(\omega t + \varphi)[1 + \exp i\delta \cdot M_{\mathrm{obj}}] \begin{pmatrix} 1 \\ i\varepsilon \end{pmatrix}$$
$$\oplus E_{BY} \exp i\left(\omega t + \Psi - \frac{\pi}{2} \right)[1 + \exp i\delta \cdot M_{\mathrm{obj}}] \begin{pmatrix} i\varepsilon \\ 1 \end{pmatrix}. \qquad (3)$$

The real part of expression (3) describes the intensity of the electric vector[7]

$$\mathrm{Re}(\mathbf{E}_\Sigma) = \mathbf{p} \cos \omega t + \mathbf{q} \sin \omega t, \qquad (4)$$

where the parameters of the summed ellipse $\mathbf{p}$ and $\mathbf{q}$ are defined in terms of components of the polarization ellipse of each of the bases A and B according to the rules[2]

$$\mathbf{p} = \mathrm{Re}(\mathbf{E}_\Sigma)_A \oplus \mathrm{Re}(\mathbf{E}_\Sigma)_B = \mathbf{p}_A \oplus \mathbf{p}_B,$$

$$\mathbf{q} = \mathrm{Im}(\mathbf{E}_\Sigma)_A \oplus \mathrm{Im}(\mathbf{E}_\Sigma)_B = \mathbf{q}_A \oplus \mathbf{q}_B,$$

whence

$$\mathbf{p}_A = \frac{1}{2} E_{AX}$$

$$\times \begin{pmatrix} \hat{a} \exp i(\varphi + \delta) + \hat{a}^* \exp -i(\varphi + \delta) \\ + (\exp i\varphi + \exp -i\varphi) \\ \hat{b} \exp i(\varphi + \delta) + \hat{b}^* \exp -i(\varphi + \delta) \\ + i\varepsilon(\exp i\varphi - \exp -i\varphi) \end{pmatrix},$$

$$\mathbf{p}_B = \frac{1}{2} E_{BY}$$

$$\times \begin{pmatrix} \hat{c} \, \exp i\left(\Psi - \dfrac{\pi}{2} + \delta\right) - \hat{c}^* \, \exp{-i}\left(\Psi - \dfrac{\pi}{2} + \delta\right) \\ + i\varepsilon\left[\exp i\left(\Psi - \dfrac{\pi}{2}\right) - \exp{-i}\left(\Psi - \dfrac{\pi}{2}\right)\right] \\ \hat{d} \, \exp i\left(\Psi - \dfrac{\pi}{2} + \delta\right) + \hat{d}^* \, \exp{-i}\left(\Psi - \dfrac{\pi}{2} + \delta\right) \\ + \left[\exp i\left(\Psi - \dfrac{\pi}{2}\right) + \exp{-i}\left(\Psi - \dfrac{\pi}{2}\right)\right] \end{pmatrix},$$

$$\mathbf{q}_A = \frac{i}{2} E_{AX} \begin{pmatrix} \hat{a} \, \exp i(\varphi + \delta) - \hat{a}^* \, \exp{-i}(\varphi + \delta) \\ + (\exp i\varphi - \exp{-i}\varphi) \\ \hat{b} \, \exp i(\varphi + \delta) - \hat{b}^* \, \exp{-i}(\varphi + \delta) \\ + i\varepsilon(\exp i\varphi + \exp{-i}\varphi) \end{pmatrix},$$

$$\mathbf{q}_B = \frac{i}{2} E_{BY}$$

$$\times \begin{pmatrix} \hat{c} \, \exp i\left(\Psi - \dfrac{\pi}{2} + \delta\right) - \hat{c}^* \, \exp{-i}\left(\Psi - \dfrac{\pi}{2} + \delta\right) \\ + i\varepsilon\left[\exp i\left(\Psi - \dfrac{\pi}{2}\right) + \exp{-i}\left(\Psi - \dfrac{\pi}{2}\right)\right] \\ \hat{d} \, \exp i\left(\Psi - \dfrac{\pi}{2} + \delta\right) + \hat{d}^* \, \exp{-i}\left(\Psi - \dfrac{\pi}{2} + \delta\right) \\ + \left[\exp i\left(\Psi - \dfrac{\pi}{2}\right) - \exp{-i}\left(\Psi - \dfrac{\pi}{2}\right)\right] \end{pmatrix}.$$

$$(5)$$

Here, to simplify the written forms of the expressions we have introduced the notation $\hat{a} = \hat{m}_{11} + i\varepsilon\hat{m}_{12}$, $\hat{b} = \hat{m}_{21} + i\varepsilon\hat{m}_{22}$, $\hat{c} = i\varepsilon\hat{m}_{11} + \hat{m}_{12}$, and $\hat{d} = i\varepsilon\hat{m}_{21}\hat{m}_{22}$. Photoanisotropy and photogyrotropy were first discovered in sensitive media under the action of linearly and circularly polarized light by Weigert[8] and by Zocher and Coper.[9] The law describing the relationship of the anisotropy and gyrotropy thus induced to the polarization characteristics of completely polarized inducing light was derived in Refs. 1 and 10. Complex photoinduced elliptic birefringence coefficients appear in this law, and functions of the isotropic ($\hat{s}$), anisotropic ($\hat{v}_L$), and gyrotropic ($\hat{v}_G$) responses were introduced to describe the vector photoresponse of a polarization-sensitive medium.

In the case of partial polarization the complex elliptic birefringence coefficient is represented as the result of the summation of the respective complex coefficients that are induced separately by the two independent, mutually incoherent components of the inducing radiation. The corresponding law for a two-dimensional medium is written in the form

$$\hat{n}_1^2 = \hat{n}_0^2 + \hat{s}(I_1 + I_2)_A + \hat{s}(I_1 + I_2)_B$$
$$+ \sqrt{[\hat{v}_L(I_1 - I_2)_A]^2 + [\hat{v}_G(2\sqrt{I_1 I_2} \, \sin \gamma)_A]^2}$$
$$+ \sqrt{[\hat{v}_L(I_1 - I_2)_B]^2 + [\hat{v}_G(2\sqrt{I_1 I_2} \, \sin \gamma)_B]^2},$$

$$\hat{n}_2^2 = \hat{n}_0^2 + \hat{s}(I_1 + I_2)_A + \hat{s}(I_1 + I_2)_B$$
$$- \sqrt{[\hat{v}_L(I_1 - I_2)_A]^2 + [\hat{v}_G(2\sqrt{I_1 I_2} \, \sin \gamma)_A]^2}$$
$$- \sqrt{[\hat{v}_L(I_1 - I_2)_B]^2 + [\hat{v}_G(2\sqrt{I_1 I_2} \, \sin \gamma)_B]^2},$$

$$\sin 2\theta_A = \frac{(2\sqrt{I_1 I_2} \, \sin \gamma)_A}{\sqrt{(I_1 - I_2)_A^2 + (2\sqrt{I_1 I_2} \, \sin \gamma)_A^2}},$$

$$\sin 2\theta_B = \sin 2\left(\theta_A + \frac{\pi}{2}\right),$$

$$\cos 2\theta_A = \frac{(I_1 - I_2)_A}{\sqrt{(I_1 - I_2)_A^2 + (2\sqrt{I_1 I_2} \, \sin \gamma)_A^2}},$$

$$\cos 2\theta_B = \cos 2\left(\theta_A + \frac{\pi}{2}\right), \qquad (6)$$

where $\hat{n}_1$ and $\hat{n}_2$ are the complex elliptic birefringence coefficients of the medium, $\hat{n}_0$ is the complex refractive index of the medium in the original unirradiated state, θ_A and θ_B are the angles specifying the orientation of the major axis of the polarization ellipse for the A and B components measured in the counterclockwise direction relative to the x axis, $(I_1 + I_2)_A$ and $(I_1 + I_2)_B$ are the first Stokes parameters, $(I_1 - I_2)_A$ and $(I_1 - I_2)_B$ are the second Stokes parameters, $(2\sqrt{I_1 I_2} \, \sin\gamma)_A$ and $(2\sqrt{I_1 I_2} \, \sin\gamma)_B$ are the fourth Stokes parameter for the A and B components, and $\gamma = \pm(\pi/2)$ with a plus sign for right-hand rotation of the polarization ellipse and a minus sign for left-hand rotation.

The light-induced anisotropy and gyrotropy of the medium can be described by Jones matrices.[6,1] On the basis of Eq. (6), the resultant Jones matrix of a polarization-sensitive medium is constructed from the Jones matrices corresponding to two structures of mutually orthogonal anisotropy and gyrotropy induced by two incoherent, mutually orthogonally polarized components of the partially polarized inducing radiation. The following rules are used here:

I: $\quad M = M_A M_B$;

II: $\quad M = \prod_{i=1}^{n} \prod_{j=1}^{n} M_{Ai} M_{Bj}, \quad M_A = \prod_{i=1}^{n} M_{Ai}, \quad M_B = \prod_{j=1}^{n} M_{Bj}$;

III: $\quad M(\theta) = S(-\theta) M S(\theta), \qquad (7)$

where $S(\theta)$ and $S(-\theta)$ are the forward and reverse rotation matrices.[11]

On the basis of Eqs. (6) and (7), for the resultant matrix in the linear approximation we obtain

$$M = M_A M_B \approx \exp{-2i\varkappa d\hat{n}_0} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \qquad (8)$$

where

$$M_{11,22} = 1 - \frac{i\varkappa d}{2\hat{n}_0}[\hat{s}(I_1+I_2)_A + \hat{s}(I_1+I_2)_B \pm \hat{v}_L \cos 2\theta_A$$

$$\times (I_1-I_2)_A \pm \hat{v}_L \cos 2\theta_B \cdot (I_1-I_2)_B],$$

$$M_{12,21} = -\frac{i\varkappa d}{2\hat{n}_0}[\hat{v}_L \sin 2\theta_A \cdot (I_1-I_2)_A + \hat{v}_L \sin 2\theta_B$$

$$\times (I_1-I_2)_B \mp i\hat{v}_G(2\sqrt{I_1 I_2} \sin \gamma)_A$$

$$\mp i\hat{v}_G(2\sqrt{I_1 I_2} \sin \gamma)_B],$$

$\varkappa = 2\pi/\lambda$, and d is the thickness of the recording medium.

We express the Stokes parameters appearing in expression (8) in terms of the parameters $\mathbf{p}_A$, $\mathbf{p}_B$, $\mathbf{q}_A$, and $\mathbf{q}_B$.[1] For the matrix of the hologram we then obtain

$$M = M_0 + M_{-1} + M_{+1},$$

where

$$M_0 \approx \exp{-2i\varkappa d\hat{n}_0} \begin{pmatrix} (M_0)_{11} & (M_0)_{12} \\ (M_0)_{21} & (M_0)_{22} \end{pmatrix}, \tag{9}$$

$$(M_0)_{11,22} = 1 - \frac{i\varkappa d}{2\hat{n}_0}[(\hat{s}\pm\hat{v}_L)(E_{AX}^2 + \varepsilon^2 E_{BY}^2) + (\hat{s}\mp\hat{v}_L)$$

$$\times (\varepsilon^2 E_{AX}^2 + E_{BY}^2)] - \frac{i\varkappa d}{2\hat{n}_0}\{(\hat{s}\pm\hat{v}_L)[(E_{AX}^2$$

$$+ \varepsilon^2 E_{BY}^2)\hat{m}_{11}\hat{m}_{11}^* + (\varepsilon^2 E_{AX}^2 + E_{BY}^2)\hat{m}_{12}\hat{m}_{12}^*$$

$$- i\varepsilon(E_{AX}^2 - E_{BY}^2)(\hat{m}_{11}\hat{m}_{12}^* - \hat{m}_{11}^*\hat{m}_{12})] + (\hat{s}\mp\hat{v}_L)$$

$$\times [(\varepsilon^2 E_{AX}^2 + E_{BY}^2)\hat{m}_{22}\hat{m}_{22}^* + (E_{AX}^2 + \varepsilon^2 E_{BY}^2)$$

$$\times \hat{m}_{21}\hat{m}_{21}^* + i\varepsilon(E_{AX}^2 - E_{BY}^2)(\hat{m}_{22}\hat{m}_{21}^* - \hat{m}_{22}^*\hat{m}_{21})]\},$$

$$(M_0)_{12,21} = -\frac{i\varkappa d}{2\hat{n}_0}\{2i\varepsilon[(\hat{v}_L \pm \hat{v}_G)E_{AX}^2 + (\hat{v}_L \mp \hat{v}_G)E_{BY}^2]\}$$

$$- \frac{i\varkappa d}{2\hat{n}_0}\{(\hat{v}_L \mp \hat{v}_G)[(E_{AX}^2 + \varepsilon^2 E_{BY}^2)\hat{m}_{11}\hat{m}_{21}^*$$

$$+ i\varepsilon(E_{AX}^2 - E_{BY}^2)(\hat{m}_{12}\hat{m}_{21}^* - \hat{m}_{11}\hat{m}_{22}^*) + (\varepsilon^2 E_{AX}^2$$

$$+ E_{BY}^2)\hat{m}_{12}\hat{m}_{22}^*] + (\hat{v}_L \pm \hat{v}_G)[(E_{AX}^2$$

$$+ \varepsilon^2 E_{BY}^2)\hat{m}_{11}^*\hat{m}_{21}^* - i\varepsilon(E_{AX}^2 - E_{BY}^2)(\hat{m}_{12}^*\hat{m}_{21}$$

$$- \hat{m}_{11}^*\hat{m}_{22}) + (\varepsilon^2 E_{AX}^2 + E_{BY}^2)\hat{m}_{12}^*\hat{m}_{22}]\}. \tag{10}$$

$$M_{-1} \approx -\frac{i\varkappa d}{2\hat{n}_0} \exp{-2i\varkappa d\hat{n}_0} \exp{i\delta} \begin{pmatrix} (M_{-1})_{11} & (M_{-1})_{12} \\ (M_{-1})_{21} & (M_{-1})_{22} \end{pmatrix},$$

$$(M_{-1})_{11,22} = (\hat{s}\pm\hat{v}_L)[(E_{AX}^2 + \varepsilon^2 E_{BY}^2)\hat{m}_{11} + i\varepsilon(E_{AX}^2$$

$$- E_{BY}^2)\hat{m}_{12}] + (\hat{s}\mp\hat{v}_L)[-i\varepsilon(E_{AX}^2 - E_{BY}^2)\hat{m}_{21}$$

$$+ (\varepsilon^2 E_{AX}^2 + E_{BY}^2)\hat{m}_{22}],$$

$$(M_{-1})_{12,21} = (\hat{v}_L \mp \hat{v}_G)[-i\varepsilon(E_{AX}^2 - E_{BY}^2)\hat{m}_{11} + (\varepsilon^2 E_{AX}^2$$

$$+ E_{BY}^2)\hat{m}_{12}] + (v_L \pm \hat{v}_G)[(E_{AX}^2 + \varepsilon^2 E_{BY}^2)\hat{m}_{21}$$

$$+ i\varepsilon(E_{AX}^2 - E_{BY}^2)\hat{m}_{22}], \tag{11}$$

$$M_{+1} \approx -\frac{i\varkappa d}{2\hat{n}_0} \exp{-2i\varkappa d\hat{n}_0} \exp{-i\delta}$$

$$\times \begin{pmatrix} (M_{+1})_{11} & (M_{+1})_{12} \\ (M_{+1})_{21} & (M_{+1})_{22} \end{pmatrix},$$

$$(M_{+1})_{11,22} = (\hat{s}\pm\hat{v}_L)[(E_{AX}^2 + \varepsilon^2 E_{BY}^2)\hat{m}_{11}^* - i\varepsilon(E_{AX}^2$$

$$- E_{BY}^2)\hat{m}_{12}^*] + (\hat{s}\mp\hat{v}_L)[i\varepsilon(E_{AX}^2 - E_{BY}^2)\hat{m}_{21}^*$$

$$+ (\varepsilon^2 E_{AX}^2 + E_{BY}^2)\hat{m}_{22}^*],$$

$$(M_{+1})_{12,21} = (\hat{v}_L \pm \hat{v}_G)[i\varepsilon(E_{AX}^2 - E_{BY}^2)\hat{m}_{11}^* + (\varepsilon^2 E_{AX}^2$$

$$+ E_{BY}^2)\hat{m}_{12}^*] + (\hat{v}_L \mp \hat{v}_G)[(E_{AX}^2 + \varepsilon^2 E_{BY}^2)\hat{m}_{21}^*$$

$$- i\varepsilon(E_{AX}^2 - E_{BY}^2)\hat{m}_{22}^*]. \tag{12}$$

Here the matrix M_0 is responsible for the formation of the undiffracted beam, and M_{-1} and M_{+1} are responsible for the formation of the imaginary and real images, respectively. In (10)–(12) we set

$$\hat{s}+\hat{v}_L \neq 0, \quad \hat{s}-\hat{v}_L = 0, \quad \hat{v}_L + \hat{v}_G = 0, \quad \hat{v}_L - \hat{v}_G \neq 0. \tag{13}$$

These conditions are equivalent to $\hat{s} = \hat{v}_L$ and $\hat{v}_L = -\hat{v}_G$. In most photoanisotropic media, for example, in mordant azo dyes,[1] they hold with a high accuracy and are physically meaningful. Under these conditions

$$M_0 \approx \exp{-2i\varkappa d\hat{n}_0}\left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{i\varkappa d\hat{v}_L}{\hat{n}_0}[P_0 \right.$$

$$\left. + M_{obj}P(M_{obj}^*)^T]\right\}, \tag{10'}$$

where we have introduced the notation

$$P_0 = \begin{pmatrix} E_{AX}^2 + \varepsilon^2 E_{BY}^2 & 2i\varepsilon E_{BY}^2 \\ 2i\varepsilon E_{AX}^2 & \varepsilon^2 E_{AX}^2 + E_{BY}^2 \end{pmatrix},$$

$$P = \begin{pmatrix} E_{AX}^2 + \varepsilon^2 E_{BY}^2 & -i\varepsilon(E_{AX}^2 - E_{BY}^2) \\ i\varepsilon(E_{AX}^2 - E_{BY}^2) & \varepsilon^2 E_{AX}^2 + E_{BY}^2 \end{pmatrix},$$

$$(M_{obj}^*)^T = \begin{pmatrix} \hat{m}_{11}^* & \hat{m}_{21}^* \\ \hat{m}_{12}^* & \hat{m}_{22}^* \end{pmatrix}$$

is the transposed conjugate matrix of the object,

$$M_{-1} \approx -\frac{i \varkappa d \hat{v}_L}{\hat{n}_0} \exp -2i \varkappa d \hat{n}_0 \, \exp i\delta M_{\mathrm{obj}} P, \qquad (11')$$

$$M_{+1} \approx -\frac{i \varkappa d \hat{v}_L}{\hat{n}_0} \exp -2i \varkappa d \hat{n}_0 \, \exp -i\delta P(M_{\mathrm{obj}}^*)^T. \quad (12')$$

When light is passed through the polarization hologram of the original wave (1), the wave passing through it is molded into three waves, among which the wave passing without diffraction is

$$\mathbf{E}_0' = M_0 \mathbf{E}_{\mathrm{op}} \approx \exp -2i\varkappa d \hat{n}_0 \left\{ E_{AX} \, \exp \, i(\omega t + \varphi) \right.$$

$$\times \left[\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) - \frac{i\varkappa d \hat{v}_L}{\hat{n}_0} (P_0 + M_{\mathrm{obj}} P(M_{\mathrm{obj}}^*)^T) \right] \left(\begin{array}{c} 1 \\ i\varepsilon \end{array} \right)$$

$$\oplus E_{BY} \, \exp \, i \left(\omega t + \Psi - \frac{\pi}{2} \right) \left[\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) - \frac{i\varkappa d \hat{v}_L}{\hat{n}_0} \right.$$

$$\left. \times (P_0 + M_{\mathrm{obj}} P(M_{\mathrm{obj}}^*)^T) \right] \left(\begin{array}{c} i\varepsilon \\ 1 \end{array} \right) \bigg\}. \qquad (14)$$

In addition, the imaginary and real images take on the following forms

$$\mathbf{E}_{-1}' = M_{-1} \mathbf{E}_{\mathrm{op}} \approx -\frac{i\varkappa d \hat{v}_L}{\hat{n}_0} (1 + \varepsilon^2) \exp -2i\varkappa d \hat{n}_0$$

$$\times \left\{ E_{AX}^3 \, \exp \, i(\omega t + \varphi + \delta) M_{\mathrm{obj}} \left(\begin{array}{c} 1 \\ i\varepsilon \end{array} \right) \right.$$

$$\left. \oplus E_{BY}^3 \, \exp \, i \left(\omega t + \Psi - \frac{\pi}{2} + \delta \right) M_{\mathrm{obj}} \left(\begin{array}{c} i\varepsilon \\ 1 \end{array} \right) \right\}, \quad (15)$$

$$\mathbf{E}_{+1}' = M_{+1} \mathbf{E}_{\mathrm{op}} \approx -\frac{i\varkappa d \hat{v}_L}{\hat{n}_0} \exp -2i\varkappa d \hat{n}_0$$

$$\times \left\{ E_{AX} \exp \, i(\omega t + \varphi - \delta) P(M_{\mathrm{obj}}^*)^T \left(\begin{array}{c} 1 \\ i\varepsilon \end{array} \right) \right.$$

$$\left. \oplus E_{BY} \, \exp \, i \left(\omega t + \Psi - \frac{\pi}{2} - \delta \right) P(M_{\mathrm{obj}}^*)^T \left(\begin{array}{c} i\varepsilon \\ 1 \end{array} \right) \right\}.$$

$$(16)$$

It follows from Eqs. (14)–(16) that the undiffracted beam does not contain any information on the phase of the object, since products of elements of the object matrix and the conjugates appear in the elements of M_0. Accordingly, the polarization of the beam which passed without diffraction is transformed. In the imaginary image the object wave forms with complete restoration of the state and degree of partial polarization to within a multiplier. Similarly, a pseudoscopic and polarization-transformed field of the object forms in the real image. Thus, in the most general case of an anisotropic gyrotropic object, the real image forms with reverse rotation of the polarization ellipse and maintenance of the orientation of its major axis.

Polarization-holographic recording in photosensitive media, whose response functions obey conditions different from conditions (13), leads to different transformations of the restored object fields, which we plan to analyze in the future.

[1] Sh. D. Kakichashvili, *Polarization Holography* [in Russian], Leningrad (1989).
[2] Sh. D. Kakichashvili, Zh. Tekh. Fiz. **65** (7), 200 (1995) [Tech. Phys. **40**, 743 (1995)].
[3] D. Gabor, Nature **161**, 777 (1948).
[4] Yu. N. Denisyuk, Opt. Spektrosk. **15**, 522 (1963) [Opt. Spectrosc. (USSR) **15**, 279 (1963)].
[5] R. C. Jones, J. Opt. Soc. Am. **31**, 488 (1941).
[6] H. Hurwitz Jr. and R. C. Jones, J. Opt. Soc. Am. **31**, 493 (1941).
[7] M. Born and E. Wolf, *Principles of Optics*, 3rd ed., Pergamon Press, Oxford (1966) [Russ. transl., Nauka, Moscow (1970)].
[8] F. Weigert, Verhandl. Deutsch. Phys. Ges. **21**, 479 (1919).
[9] H. Zocher and K. Coper, Z. Phys. Chem. **132**, 313 (1928).
[10] Sh. D. Kakichashvili, Opt. Spektrosk. **52**, 317 (1982) [Opt. Spectrosc. (USSR) **52**, 191 (1982)].
[11] W. A. Shurcliff, *Polarized Light: Production and Use*, Harvard University Press, Cambridge, Mass. (1962) [Russ. transl., Mir, Moscow (1965)].

The magnetic field of Langmuir oscillations

N. S. Bukhman

Michurinsk Agricultural Academy, Michurinsk, Russia

The small-scale quasistationary magnetic field appearing in a plasma exhibiting plasma oscillations is considered. It is shown that this magnetic field is capable of leading to dissipation of the electromagnetic waves propagating in the plasma. © *1997 American Institute of Physics.* [S1063-7842(97)03106-1]

Let there be a high-frequency monochromatic electric field $\mathbf{E}(\mathbf{r})\exp(i\omega t)$ in a uniform, nearly collisionless plasma. Now, let us consider (within the elementary plasma theory[1]) the motion of an individual electron in this field. After solving the equation of motion of the electron, finding its angular momentum (averaged over the oscillations in the high-frequency field), and utilizing the known gyromagnetic ratio for the orbital motion of an electron,[2] we can easily prove that the motion of plasma electrons in an rf electric field leads to the appearance in the plasma of a bulk stationary magnetic moment

$$\mathbf{M}=\frac{ne^3}{4m^2\omega^3 c}\,\mathrm{Re}\{i[\mathbf{E}\times\mathbf{E}^*]\},\tag{1}$$

where n is the plasma density.

In the case of a transverse electromagnetic wave, Eq. (1) describes an inverse Faraday effect.[2] In this paper we consider the case in which a high-frequency electric field in a plasma is associated with the existence of plasma oscillations in it. It is known[1] that plasma oscillations have an electrostatic character and can be described using the scalar potential Φ

$$\mathbf{E}(\mathbf{r})=-\mathrm{grad}\ \Phi(\mathbf{r}),\quad \Phi=A\ \exp(i\phi).\tag{2}$$

With consideration of expression (2), Eq. (1) can be rewritten in the form

$$\mathbf{M}=\frac{ne^3}{4m^2\omega^3 c}\,[\,\mathrm{grad}\ A^2\times\mathrm{grad}\ \phi],\tag{3}$$

the divergence of the bulk magnetic moment being equal to zero ($\mathrm{div}\mathbf{M}=0$) regardless of the form of the functions $A(\mathbf{r})$ and $\phi(\mathbf{r})$. Then, for the magnetic field excited in the plasma we have (after solving the system of magnetostatic equations[3]) $\mathbf{H}=0$, and

$$\mathbf{B}=4\pi\mathbf{M}=\frac{\pi ne^3}{m^2\omega^3 c}\,[\,\mathrm{grad}\ A^2\times\mathrm{grad}\ \phi].\tag{4}$$

It is seen from Eq. (4) that the magnetic field of the plasma oscillations is nonzero only where these oscillations exist (the bulk magnetic moment appearing in a plasma does not create an "external" magnetic field). It is seen, in addition, that the appearance of plasma oscillations (or electron plasma waves) in a plasma fails to lead to the appearance of a quasistationary magnetic field only if the gradients of the amplitude and phase of the scalar potential of the rf field are parallel to one another (i.e., when there is a single plane plasma wave, as often happens in theory, but rarely in experiment).

In the general case of directionally uncorrelated plasma waves ("Langmuir noise"), for the magnetic field appearing in the plasma we have the estimate $B\cong B_{Le}$, where

$$B_{Le}=\frac{\pi ne^3}{m^2\omega^3 c}\,E^2.\tag{5}$$

Here E^2 is the characteristic value of the intensity of the field of the plasma oscillations.

In principle, two plane waves propagating at an angle of the order of 90° relative to one another are sufficient for estimate (5) to be valid. In practice this is the case of a "gas of plasmons."

For the cyclotron frequency of electrons in the magnetic field B_L we have

$$\omega_c/\omega=(1/2)(n/n_{\mathrm{cr}})(W_E/mc^2),\tag{6}$$

where n_{cr} is the kinetic plasma density, W_E is the kinetic energy of the oscillations of an electron in an rf electric field, and $mc^2=0.5$ MeV is the rest energy of an electron.

It is seen from Eq. (6) that in the case of $n\cong n_{\mathrm{cr}}$ and $W_E<mc^2$ (when our treatment is well-founded), the magnetic field of the plasma oscillations has a weak influence on the motion of the electrons (justifying the performance of the treatment in the first nonvanishing order with respect to the effect).

Nevertheless, the effect under consideration can have a significant influence on the dissipation of electromagnetic waves in a plasma, since the magnetic field induced by plasma oscillations is small-scale (with a characteristic length for nonuniformity of the order of the Debye length) and chaotic (in the case of Langmuir chaos). Like Coulomb scattering on heavy ions, scattering on a small-scale magnetic field should lead to the dissipation of electromagnetic waves in a plasma with an "effective collision frequency" ω_c (6).

Comparing the cyclotron frequency (6) with the effective frequency of electron–ion collisions,[3] we have

$$\omega_c/\nu_{ei}=W_E/W_0,$$

$$W_0[eV]=1.2\{(Z\ln\Lambda)/(\lambda_0[m]T^{3/2}[eV])\},\tag{7}$$

where Z is the charge of the ions, $\ln\Lambda$ is the Column logarithm, T is the electron temperature, and λ_0 is the vacuum length of an electromagnetic wave whose frequency equals the frequency of the plasma oscillations.

In a typical (in investigations of the interaction of intense microwaves with a plasma) situation[4] $Z\cong1$, $\ln\Lambda\cong10$, $\lambda_0\cong0.1$ m, and $T\cong10$ eV we have $W_0\cong4$ eV. If the kinetic energy of the oscillations of electrons in a field of plasma oscillations exceeds that value, dissipation of the electromagnetic wave in such a plasma by means of scattering of the electrons on the small-scale magnetic field of the plasma waves is more extensive than ordinary dissipation by means of electron–ion collisions, and absorption of the electromagnetic wave will be controlled by the ''anomalous'' collision frequency ω_c (6).

[1] V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas*, 2nd ed., Pergamon Press, Oxford–New York (1970).
[2] M. M. Bredov, V. V. Rumyantsev, and I. N. Toptygin, *Classical Electrodynamics* [in Russian], Nauka, Moscow (1985).
[3] F. F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, 2nd ed., Plenum Press, New York (1984) [Russ. transl., Mir, Moscow (1987)].
[4] G. M. Batanov and V. A. Silin, in *The Dissipation of Electromagnetic Waves in Plasmas (Proceedings of the Lebedev Physics Institute, Academy of Sciences of the USSR, Vol. 92)*, N. G. Basov (ed.), Consultants Bureau, New York (1982) [Russ. original, Tr. Inst. Fiz. Inst. Akad. Nauk SSSR **92**, 3–34 (1977)].

ERRATA

Erratum: Modification of several physicochemical properties of a high-temperature superconducting yttrium ceramic upon doping with gold [Tech. Phys. 42, 152–154 (February (1997)].

T. V. Krachino, I. N. Zimkin, M. A. Mittsev, and Yu. P. Stepanov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
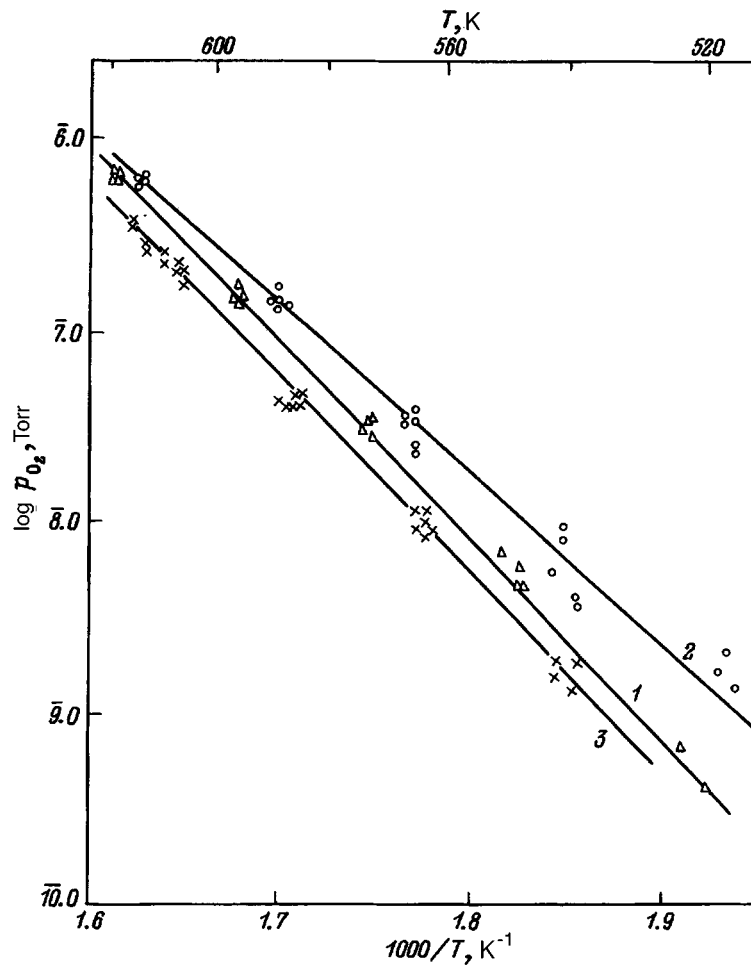
[S1063-7842(97)03206-6]

Figure 1 should read as follows:

FIG. 1.