Dynamics of the skin effect during the discharging of a capacitor through straight busbars

M. A. Shakirov and R. P. Kiyatkin

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
(Submitted December 18, 1995; resubmitted April 18, 1996)
Zh. Tekh. Fiz. **67**, 1–7 (July 1997)

The manifestations of the skin effect during transient processes in systems with straight busbars having cross sections of arbitrary shape and a general method for investigating this effect are considered. The method permits direct observation of the dynamics of the variation of the current density in busbars for any degree of manifestation of the skin effect in two-dimensional problems. © *1997 American Institute of Physics.* [S1063-7842(97)00107-4]

INTRODUCTION

The development of the technology for using strong pulsed magnetic fields raises the problem of creating effective methods for calculating the transient processes in electromagnetic fields with consideration of the skin effect in massive conductors, including straight cables. An analytical solution of this problem (generally on the basis of the Laplace transformation) is possible only for one-dimensional systems with straight conductors having simple cross sections: round and annular cross sections or cross sections in the form of a half plane.[1–4] The application of finite-element analysis requires large amounts of computer memory (since not only the cross sections of the conductors, but also the insulators surrounding them are discretized), and the machinery for integral equations is in the development stage and requires further research.[5] This attests to the inadequacy of the general methodology for analyzing the skin and proximity effects in pulsed systems.

The present work examines an effective and convenient method for analyzing the dynamics of the skin effect in straight conductors with busbars and shields of arbitrary cross section, which is based on the use of approximate equivalent circuits of massive conductors obtained by dividing the busbars and shields into elementary conductors under the following commonly employed assumptions: 1) the fields are plane-parallel and two-dimensional; 2) the bias currents in the busbars and shields are negligibly small in comparison with the conduction currents, i.e., the transient processes in the conductors are quasistationary; 3) the permeabilities of the busbars and shields are equal to the permeability of a vacuum μ_0, and the conductivities γ of the conducting materials are functions only of the coordinates; 4) the number of busbars and shields is not restricted, and the shapes of their cross sections are arbitrary.

MODELS OF MASSIVE CONDUCTORS

Let us consider the construction of the proposed equivalent circuits in the example of a system (Fig. 1a) consisting of a power source $e(t)$ and a capacitor C, which are connected to an $R-L$ load through busbars A and B in the presence of shield S. The busbars and the shield have the same length l. We divide the busbars and the shield into elementary conductors with very small cross sections S_k, assuming that the current $i_k(t)$ within the kth cross section is uniformly distributed and that its density equals

$$\delta_k(t)=\frac{i_k(t)}{S_k}. \tag{1}$$

Clearly, the smaller are the S_k, the smaller is the error of assumption (1). The active resistance of the kth elementary conductor is

$$R_k=\frac{l}{\gamma_k S_k}, \tag{2}$$

where γ_k is the conductivity of the kth elementary conductor.

It follows from general physical arguments that if the system were such that wires carrying forward and reverse currents could be identified *a priori* in its busbars and shields, the elementary conductors could be paired to form elementary turns. In this case the analysis of the field would be reduced to a calculation of the transient processes in a fairly simple circuit-diagram model with magnetic couplings (mutual inductance) between all the elementary turns. The problem is that such wires do not exist in the general case.

The main idea is to represent the busbars and shields by a set of elementary turns with one common reverse (base) elementary conductor for each electrically isolated subsystem. The system considered in Fig. 1a contains two subsystems that are not electrically coupled. The base elementary conductor in the subsystem consisting of the busbars A and B is denoted by O_B. The base elementary conductor in the shield S is denoted by O_S. If the number of elementary conductors in busbar A equals n_A, and the number in busbar B equals n_B, the number of elementary turns in the first subsystem will be $n_A+(n_B-1)$. The number of equivalent turns in the shield equals n_S-1, where n_S is the number of elementary conductors in the shield. There are magnetic couplings between all the elementary turns. The total number of elementary turns in the approximate model of the system (Fig. 1a) equals

$$n=n_A+n_B+n_S-2. \tag{3}$$

 1063-7842/97/070717-07$10.00

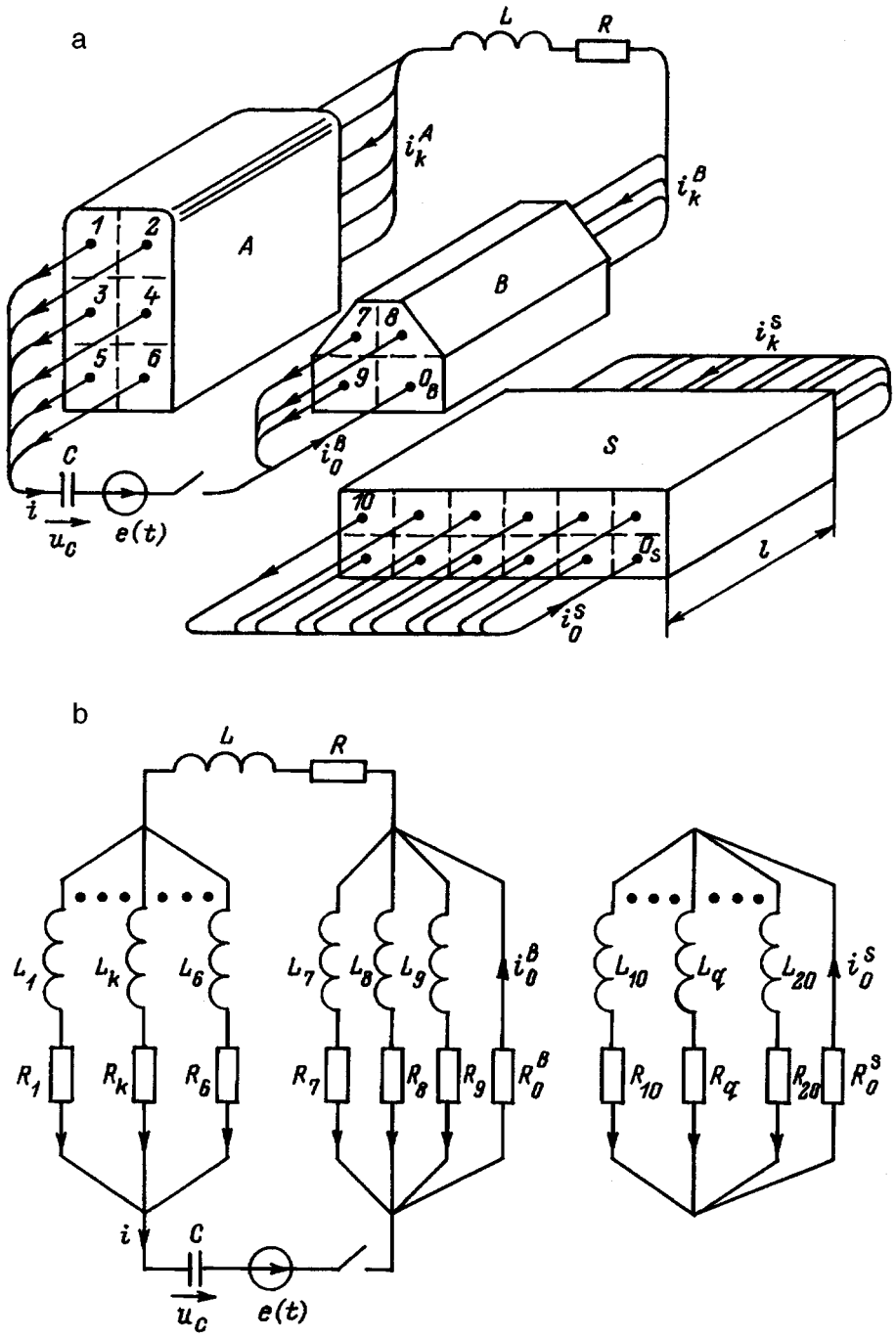

FIG. 1. Discretization of the busbars and shield of a conductor (a) and its equivalent circuit (b) for analyzing transient processes.

In Fig. 1a $n_A=6$, $n_B=4$, and $n_S=12$; therefore, $n=20$. The numbering of the elementary turns coincides with the numbering of the elementary conductors. The set of elementary turns corresponds to an equivalent circuit (Fig. 1b), in which all the turns are inductively coupled.

To write down the equations of the circuit-diagram model (Fig. 1b) we introduce the following notation: $\mathbf{i}_A$ is the column (of dimension n_A) of the currents in the elementary currents of busbar A; $\mathbf{i}_B$ is the column of the (n_B-1) currents in the elementary conductors of busbar B; $\mathbf{i}_S$ is the column of the (n_S-1) currents in the elementary conductors of the shield S; $\mathbf{R}_A$ is the diagonal $(n_A \cdot n_A)$ matrix of the active resistances of all the elementary conductors of busbar A; $\mathbf{R}_B$ is the diagonal $((n_B-1)\cdot(n_B-1))$ matrix of the active resistances of all the elementary conductors in busbar B, except the base conductor; $\mathbf{R}_S$ is the diagonal $((n_S-1)\cdot(n_S-1))$ matrix of the active resistances of all the elementary conductors in the shield, except the base conductor; $\mathbf{R}_0^B$ is a column (of dimension n_A+n_B-1), all of whose elements are equal to the resistance R_0^B of the base elementary conductor in busbar B; $\mathbf{R}_0^S$ is a column (of dimension n_S-1), all of whose elements are equal to the resistance R_0^S of the base elementary conductor in the shield; **1** is a column (of dimension n_A), all of whose elements are equal to unity. According to the diagram in Fig. 1b, the following system of Kirchhoff's equations can be written in the state variables $\mathbf{i}_A$, $\mathbf{i}_B$, $\mathbf{i}_S$, and u_C and the ''algebraic'' variables i, i_0^B, and i_0^S (i

is the current through the element C, which coincides with the load current, and i_0^B and i_0^S are the currents in the base elementary conductors of busbar B and the shield)

$$\begin{bmatrix} \mathbf{M}'_A & \mathbf{M}_{AB} & \mathbf{M}_{AS} & \mathbf{O} \\ \mathbf{M}_{BA} & \mathbf{M}_B & \mathbf{M}_{BS} & \mathbf{O} \\ \mathbf{M}_{SA} & \mathbf{M}_{SB} & \mathbf{M}_S & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & C \end{bmatrix} \cdot \frac{d}{dt} \begin{bmatrix} \mathbf{i}_A \\ \mathbf{i}_B \\ \mathbf{i}_S \\ u_C \end{bmatrix}$$

$$= \begin{bmatrix} -R'_A & \mathbf{O} & \mathbf{O} & -\mathbf{1} & \mathbf{O} & & \mathbf{O} \\ & & & & & -\mathbf{R}_0^B & \\ \mathbf{O} & -\mathbf{R}_B & \mathbf{O} & \mathbf{O} & \mathbf{O} & & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & -\mathbf{R}_S & \mathbf{O} & \mathbf{O} & \mathbf{O} & -\mathbf{R}_0^S \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\times \begin{bmatrix} \mathbf{i}_A \\ \mathbf{i}_B \\ \mathbf{i}_S \\ u_C \\ i \\ i_0^B \\ i_0^S \end{bmatrix} + \begin{bmatrix} \mathbf{1} \\ \mathbf{O} \\ \mathbf{O} \\ 0 \end{bmatrix} \cdot e(t). \qquad (4)$$

In (4) $\mathbf{R}'_A$ is a completely filled matrix formed from $\mathbf{R}_A$ by adding the load resistance R of all of its elements. This can be represented symbolically in the form

$$\mathbf{R}'_A = \mathbf{R}_A + R. \qquad (5)$$

The matrix $\mathbf{M}'_A$ is formed from $\mathbf{M}_A$ by adding the load inductance L to each of its elements ($\mathbf{M}'_A = \mathbf{M}_A + L$). The algebraic variables on the right-hand side of (4) are easily expressed in terms of the state variables

$$i = \sum_{k=1}^{n_A} i_k^A, \quad i_0^B = \sum_{k=1}^{n_A} i_k^A + \sum_{k=1}^{n_b-1} i_k^B, \quad i_0^S = \sum_{k=1}^{n_S-1} i_k^S. \qquad (6)$$

Using (6), it is not difficult to transform (4) into a system of equations in the state variables alone:

$$\begin{bmatrix} \mathbf{M}'_A & \mathbf{M}_{AB} & \mathbf{M}_{AS} & \mathbf{O} \\ \mathbf{M}_{BA} & \mathbf{M}_B & \mathbf{M}_{BS} & \mathbf{O} \\ \mathbf{M}_{SA} & \mathbf{M}_{SB} & \mathbf{M}_S & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & C \end{bmatrix} \cdot \frac{d}{dt} \begin{bmatrix} \mathbf{i}_A \\ \mathbf{i}_B \\ \mathbf{i}_S \\ u_C \end{bmatrix}$$

$$= \begin{bmatrix} & & \mathbf{O} & -\mathbf{1} \\ & -\mathbf{R}' & & \\ & & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & -\mathbf{R}'_S & \mathbf{O} \\ \mathbf{1}^T & \mathbf{O} & \mathbf{O} & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{i}_A \\ \mathbf{i}_B \\ \mathbf{i}_S \\ u_C \end{bmatrix} + \begin{bmatrix} \mathbf{1} \\ \mathbf{O} \\ \mathbf{O} \\ 0 \end{bmatrix} \cdot e(t), \qquad (7)$$

where [in analogy to the symbolic expression (5)]

$$\mathbf{R}' = \begin{bmatrix} \mathbf{R}'_A & \mathbf{O} \\ \mathbf{O} & \mathbf{R}'_B \end{bmatrix} + R_0^B, \quad \mathbf{R}'_S = \mathbf{R}_S + R_O^S.$$

The calculation results found for a certain discretization of the cross sections of the busbars and the shield must be compared with the solution of the problem for division of the busbars and the shield into a larger number of elementary conductors, and this process must be repeated until satisfactory agreement between the new solution and the results of the preceding calculation is attained.

The formation of the system of equations (7) is the central part of the algorithm under consideration, since all the characteristics of the transient process, including the temporal variation of the current density in the busbars and the shield, the load current, the magnetic fluxes, the force interactions between the busbars and the shield, etc., can be found as a result of its solution in an assigned time interval. The speed and reliability of the calculations depends significantly on the correctness of the first attempt to discretize the cross sections of the busbars and the shield, and thus the use of approximate estimates of the character of the penetration of the field or the current into the conductors is required.

ILLUSTRATION OF THE APPLICATION OF THE METHOD IN A MODEL PROBLEM

To illustrate the reliability and accuracy of the circuit-diagram models under consideration, let us apply them to the solution of one-dimensional problems with known answers in analytical form. One of them is shown in Fig. 2. The capacitor C with an initial voltage V_0 is discharged into a short-circuited system consisting of two copper busbars of length l. To be specific, we set $C=1$ F, $U_0=1$ V, $D_A=D_B=D=0.02$ m, $D_0=0$, $h=0.01$ m, $l=1$ m, and $\gamma=5.6\times10^7$ S/m. Before performing the main calculations it would be useful to estimate the temporal characteristics of the transient process, even if only on an approximate level. This is most easily done for the case of a pronounced skin effect under the assumption of an oscillatory, ''almost sinusoidally damped'' type of process. Let T_0 be the approximate ''total period'' of the oscillations of the discharge current, and let $\omega_0=2\pi/T_0$ be its angular frequency. Then the active resistance of one busbar r_0 can be calculated as the resistance to a constant current passing along a surface layer of the busbar of thickness[6]

$$\Delta_0 = \sqrt{\frac{2}{\omega_0 \mu_0 \gamma}} = \sqrt{\frac{T_0}{\pi \mu_0 \gamma}}.$$

Therefore,

$$r_0 = \frac{l}{\gamma h \Delta_0} = \frac{l}{h}\sqrt{\frac{\pi \mu_0}{\gamma T_0}},$$

and the internal inductance of each busbar (under the assumption of a pronounced skin effect) equals

$$L_0 = \frac{r_0}{\omega_0} = \frac{T_0}{2\pi} r_0,$$

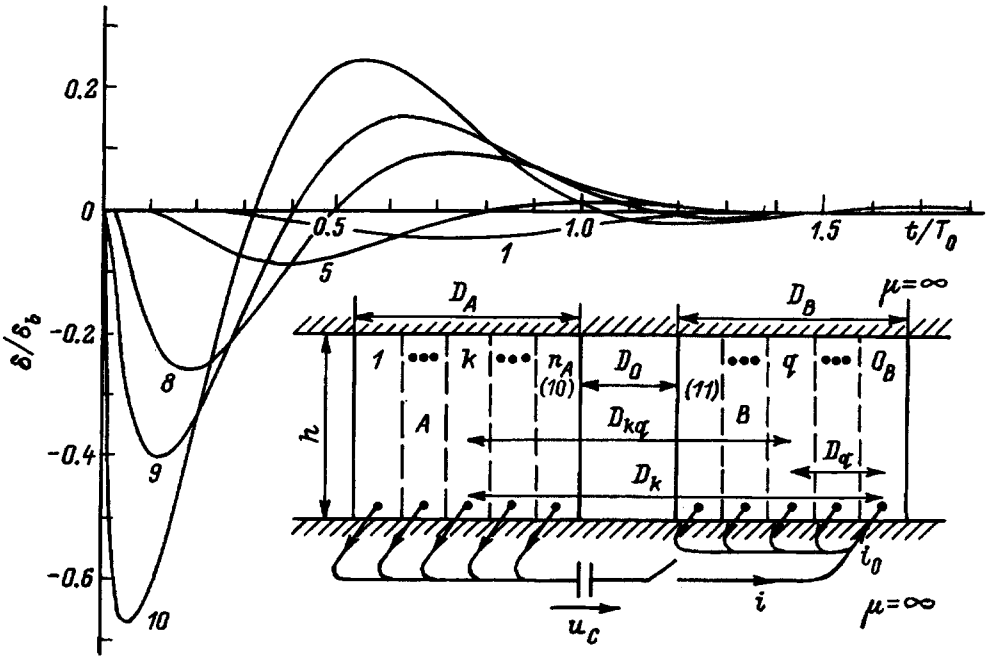

FIG. 2. Variation of the current density in busbar A of a conductor short-circuited at one end for $h=0.01$ m, $D_A=D_B=0.02$ m, $D_0=0$, and busbars divided into $n_A=n_B=10$ elementary conductors (the numbers on the curves correspond to the number of the elementary conductors, and the number of the base elementary conductor is $n=20$).

which allows us to represent a system consisting of two busbars and a capacitor in an approximation by an RLC circuit with a resistance $2r_0$ and an inductance $2L_0$, where

$$2r_0=\frac{2l}{h}\sqrt{\frac{\pi\mu_0}{\gamma T_0}}, \quad 2L_0=\frac{l}{h}\sqrt{\frac{\mu_0 T_0}{\gamma\pi}}\,.$$

The period of the damped oscillations in such a circuit equals

$$T_0=2\pi\sqrt{\frac{1}{C(2L_0)}-\frac{(2r_0)^2}{4(2L_0)^2}}$$

$$=2\pi\sqrt{1\left(C\frac{l}{h}\cdot\sqrt{\frac{\mu_0 T_0}{\gamma\pi}}\right)-\frac{1}{4}\left(\frac{2\pi}{T_0}\right)^2},$$

which gives the following approximate estimate of the total period of the oscillations in the system for a pronounced skin effect in the busbars

$$T_0=\pi^3\sqrt{\frac{25C^2l^2\mu_0}{h^2\gamma}}=0.0056\text{ s}. \tag{8}$$

The equivalent current penetration depth (the thickness of the skin layer) is found to be approximately equal to

$$\Delta_0=\sqrt{\frac{T_0}{\pi\mu_0\gamma}}=0.005\text{ m}.$$

Since $D_A=D_B=4\Delta_0$ in the present example, it can be assumed that we are considering the case of a very pronounced manifestation of the skin effect, for which an analytical solution is known[1,4] and can be used to evaluate the accuracy of the proposed numerical method.

The knowledge of Δ_0 enables us to correctly discretize the cross sections of the busbars in the initial stage of the calculations. The division of the busbars into elementary conductors (for the numerical solution) indicated by the dashed lines in Fig. 2 was made with consideration of the one-dimensional nature of the problem. To start we take a value of the width a_k of each elementary conductor approximately equal to half of the equivalent penetration depth, dividing the busbars into $n_A=n_B=10$ identical elementary conductors

$$a_k=\frac{D_A}{n_A}=\frac{D_B}{n_B}=0.002\text{ m}.$$

Thus, $a_k=0.4\Delta_0$. The cross-sectional areas S_k and the resistances R_k of all the elementary conductors, including the base elementary conductor, are identical:

$$S_k=a_k h=2\times10^{-5}\text{ m}^2,$$

$$R_k=R_0=\frac{l}{\gamma S_k}=8.9286\times10^{-4}\,\Omega.$$

The total number of elementary turns equals

$$n=n_A+n_B-1=19.$$

We determine the self- and mutual inductances of the elementary turns in the busbars from the formulas $(k,q=\overline{1,n})$

$$M_{kk}\equiv L_k=\frac{\mu_0}{h}\left(D_k-\frac{a_k}{3}\right), \tag{9}$$

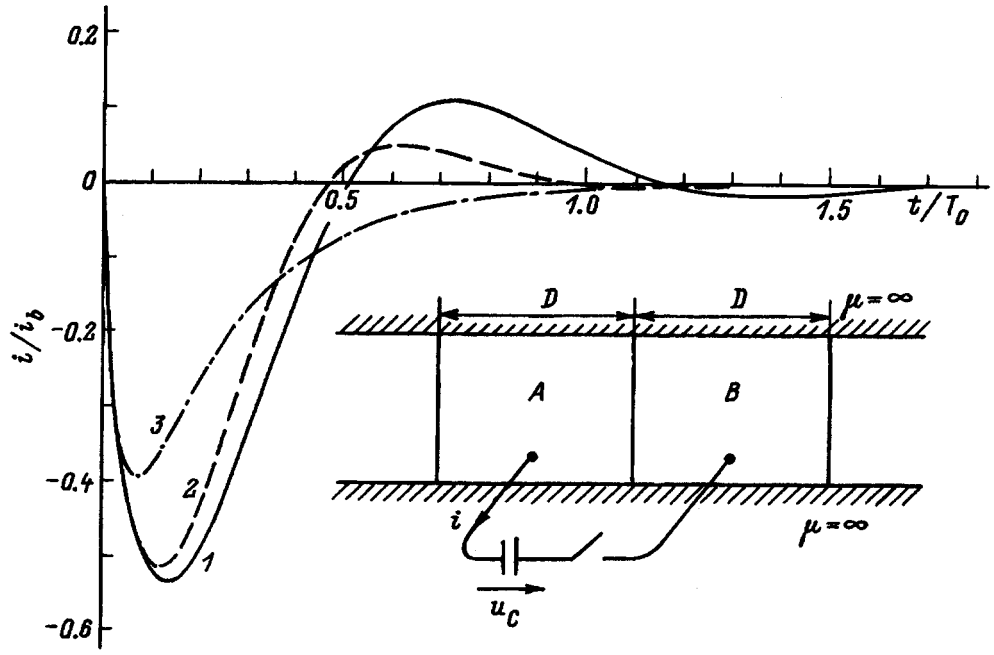

FIG. 3. Time dependence of the discharge current in the one-dimensional problem (Fig. 2). *1* — $D_A=D_B=D=4\Delta_0=0.02$ m, *2* — $D=\Delta_0=0.005$ m, *3* — $D=\Delta_0/2=0.0025$ m.

$$M_{kq}=\frac{\mu_0}{2h}\left(D_k+D_q-D_{kq}-\frac{a_k}{3}\right), \quad (k\neq q). \tag{10}$$

When $k,q=\overline{1,n_A}$, the elements M_{kk} and M_{kq} for the symmetric matrix $\mathbf{M_A}$ of dimension $n_A\times n_A=10\times 10$

$$\mathbf{M}_A=\frac{2\pi a_k}{3\times 10^5}\cdot\begin{array}{c}\begin{array}{cccccc}1 & 2 & & & 9 & 10\end{array}\\ \left[\begin{array}{cccccc}112 & 107 & \dots & & 65 & 59\\ & 106 & \dots & & 65 & 59\\ & & \vdots\,\vdots & & \vdots & \vdots\\ \text{Symmetric} & & & & 64 & 59\\ \text{values} & & & & 58 & \end{array}\right.\end{array}\begin{array}{c}1\\2\\ \\9\\10\end{array}.$$

The self- and mutual inductances of the elementary turns in busbar B are obtained from Eqs. (9) and (10) when $k,q=\overline{(n_A+1),n}$. In the case under consideration the matrix $\mathbf{M}_B$ is of dimension $(n_B-1)\times(n_B-1)=9\times 9$:

$$\mathbf{M}_B=\frac{2\pi a_k}{3\times 10^5}\cdot\begin{array}{c}\begin{array}{cccccc}11 & 12 & & & 18 & 19\end{array}\\ \left[\begin{array}{cccccc}52 & 47 & \dots & & 11 & 5\\ & 46 & \dots & & 11 & 5\\ & & \vdots\,\vdots & & \vdots & \vdots\\ \text{Symmetric} & & & & 10 & 5\\ \text{values} & & & & & 4\end{array}\right.\end{array}\begin{array}{c}11\\12\\ \\18\\19\end{array}.$$

The matrix of the mutual inductances between the elementary turns in busbar A and the elementary turns in busbar B, which has the dimension $n_A\times(n_B-1)=10\times 9$, equals

$$\mathbf{M}_{AB}=\frac{2\pi a_k}{3\times 10^5}\cdot\begin{array}{c}\begin{array}{cccccc}11 & 12 & & & 18 & 19\end{array}\\ \left[\begin{array}{cccccc}53 & 47 & \dots & & 11 & 5\\ 53 & 46 & \dots & & 11 & 5\\ \vdots & \vdots & \vdots\,\vdots & & \vdots & \vdots\\ 53 & 47 & \dots & & 11 & 5\\ 53 & 47 & \dots & & 11 & 5\end{array}\right.\end{array}\begin{array}{c}1\\2\\ \\9\\10\end{array}.$$

The circuit-diagram model is obtained from the diagram in Fig. 1b by eliminating the source $e(t)$, the load elements R and L, and the subdiagram corresponding to the shield. On the basis of system (7), the equations of state for this model can be represented in the form

$$\begin{bmatrix}\mathbf{M}_A & \mathbf{M}_{AB} & \mathbf{O}\\ \mathbf{M}_{BA} & \mathbf{M}_B & \mathbf{O}\\ \mathbf{O} & \mathbf{O} & C\end{bmatrix}\cdot\frac{d}{dt}\begin{bmatrix}\mathbf{i}_A\\ \mathbf{i}_B\\ u_C\end{bmatrix}=\begin{bmatrix} & & -\mathbf{1}\\ & -\mathbf{R}' & \\ & & \mathbf{O}\\ \mathbf{1}^T & \mathbf{O} & 0\end{bmatrix}\times\begin{bmatrix}\mathbf{i}_A\\ \mathbf{i}_B\\ u_C\end{bmatrix},$$

$$\mathbf{i}_A(+0)=\mathbf{O}, \quad \mathbf{i}_B(+0)=\mathbf{O}, \quad u_C(+0)=U_0. \tag{11}$$

The order of system (11) equals $n+1=19+1=20$. After solving it by some numerical method, it is not difficult, according to Eq. (1), to find and construct a plot of the time

TABLE I. Results of the calculation of the discharge current i (in amperes) of a capacitor into an ideal bifilar consisting of short-circuited busbars (Fig. 2) of thickness $D_A=D_B=D=4\Delta_0=0.02$ m for various values of the time t, $n=20$, and selection of the elementary conductor with the number n_0 as the basis elementary conductor (O_B) $(u_C(0)=1$ V)

		Numerical calculation			Exact value for $D=\infty$
t, s	t/T_0	$n_0=20$	$n_0=15$	$n_0=11$	(Ref. 1, p. 382)
0	0	0	0	0	0
0.001	0.179	−733.87	−733.88	−733.78	−725.38
0.002	0.358	−330.54	−330.55	−330.48	−325.40
0.003	0.537	41.49	41.50	41.45	39.27
0.004	0.716	158.49	158.51	158.40	153.54
0.005	0.896	107.80	107.81	107.74	105.05
0.006	1.074	26.89	26.89	26.89	26.62

dependence of the current density in the middle part of the kth elementary conductor, as well as of the current through the capacitor, from Eq. (6).

The results of these calculations were used to construct the dependence of the relative current densities for the centers of the elementary conductors in busbar A numbered *1*, *5*, *8*, *9*, and *10*. The value $\delta_b=\gamma U_0/2l$ was taken as the basis value. As expected, the current densities at the centers of the elementary conductors numbered *11*, *12*, *13*, and *16* in busbar B coincide with the current densities of the 10th, 9th, 8th, and 5th elementary conductors of busbar A. Curve *1* in Fig. 3 shows a plot of the current $i(t)=i(t)/i_b$ [it is assumed that $i_b=U_0/(2l/\gamma h\Delta_0)=\delta_b h\Delta_0$], the current $i(t)$ being calculated from Eq. (6). It is noteworthy that, despite the very rough preliminary discretization, the curve presented practically coincides with the current curves obtained when each busbar is divided into doubled and quadrupled numbers of elementary conductors.

It follows from the curves in Fig. 2 that the maximum value of the current density in the first elementary conductor is significantly, almost 15-fold, smaller than the current density in the 10th elementary conductor. Therefore, it can be assumed that the field scarcely reaches the outer faces of the busbars and that the case of a pronounced skin effect is realized, permitting evaluation of the accuracy of the solution obtained by comparing it with the data presented in Table I (from Ref. 1, p. 382). The current curve calculated from the formulas in Ref. 1 for the system under consideration practically coincides with curve *1* in Fig. 3.

To distinguish the real range of the first current oscillation (with the negative and positive half waves of the curve) from T_0, we denote it by T. It is seen from curve *1* in Fig. 3 that $T/T_0=1.17$; therefore, the error of approximate formula (8) in the present case amounts to $(T_0-T)/T\cdot 100\%=-15\%$. We note that, as expected, the results of the calculations scarcely depend on which of the elementary conductors is selected as the base elementary conductor (see Table I).

ANALYSIS OF THE DYNAMICS OF THE SKIN EFFECT

The literature does not offer an analytical solution for the case in which the thicknesses of the busbars in the system (Fig. 2) are commensurate with the approximate wave penetration depth ($D_A=D_B=D\leqslant\Delta_0$). The results of a numerical solution of this problem according to the proposed method are presented in the form of curves *2* and *3* in Fig. 3. When $D=\Delta_0$, it is sufficient to divide the busbars into five elementary conductors. It follows from curve *2* that $T=T_0$ in this case. When $D=\Delta_0/2$, the discharging of the capacitor already has an aperiodic character, and it is sufficient to represent each busbar by three elementary conductors.

In contrast to the one-dimensional problems considered above, Fig. 4 presents a system in which the discharge current creates a two-dimensional electromagnetic field. The capacitor discharges into a short-circuited system of busbars of rectangular profile with the same height $h=0.01$ m as in the systems in Figs. 2 and 3. The busbars are identical $(D_A=D_B=D)$, and the gap between the busbars is vanishingly small $(D_0=0)$. The main steps in the solution of this problem and the circuit-diagram model do not differ from the procedure and the diagrams presented above for the one-dimensional problems. If the busbars are divided into elementary conductors with cross sections of rectangular shape (with the sides a_k and b_k), as is shown in Fig. 4, the formulas for calculating the self- and mutual inductances of the elementary turns take the form

$$M_{kk}\equiv L_k=\frac{\mu_0}{2\pi}\ln\frac{D_{k0}^2}{g_k g_0},$$

$$M_{kq}=\frac{\mu_0}{2\pi}\ln\frac{D_{k0}D_{q0}}{D_{kq}g_0},$$

where D_{k0} is the distance between the centers of the cross sections of the kth elementary conductor and the base elementary conductor, and $g_k=0.2236(a_k+b_k)$ is the geometric-mean distance of the area of the rectangular cross section of the kth elementary conductor from itself.

The results of the calculations for three different busbar widths D are presented in Fig. 4. When the busbars were discretized, the sides of the rectangular elementary cross sections did not exceed $\Delta_0/2$. Curve *1′* in Fig. 4 mimics curve *1* in Fig. 3. A comparison of curves *1′* and *1* (Fig. 4), which were obtained for busbars with identical profiles $(D=4\Delta_0=0.02$ m, $h=0.01$ m, $D_0=0)$, shows that neglect of the two-dimensional character of the problem leads to significant errors.

Curve *3* in Fig. 4 was obtained for $h/D=0.01/0.0025=4$ (which corresponds to a tape-like busbar) and therefore practically coincides with curve *3* in Fig. 3 for the one-dimensional problem.

To provide an additional illustration of the possibilities of the calculation method described in this paper, Fig. 5 presents plots of the time dependence of the discharge current of the capacitor into a short-circuited conductor for several other (including asymmetric) busbar configurations.

In conclusion, we note that the accuracy of the solutions obtained depends on the degree of discretization of the cross sections of the conducting elements, the shapes and dimensions of the cross sections of the elementary conductors, and the accuracy of the relations used to calculate the self- and

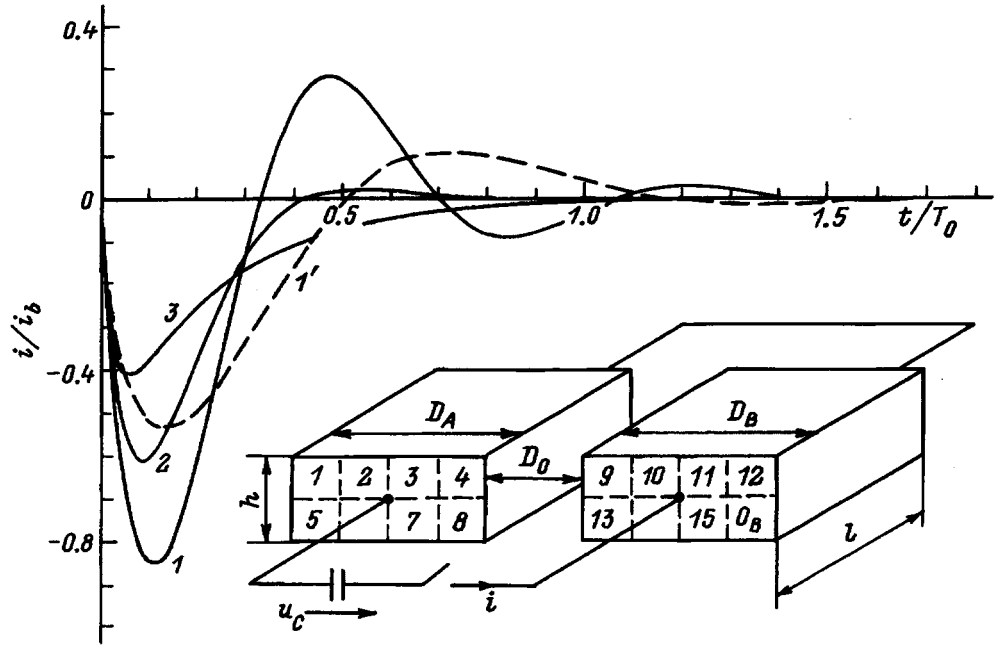

FIG. 4. Time dependence of the discharge current in the two-dimensional problem for $h=0.01$ m and $D_0=0$. *1* — $D_A=D_B=D=4\Delta_0=0.02$ m, *2* — $D=\Delta_0=0.005$ m, *3* — $D=\Delta_0/2=0.0025$ m.

mutual inductances of the elementary turns. These questions require separate investigations and have not been considered in the present work.

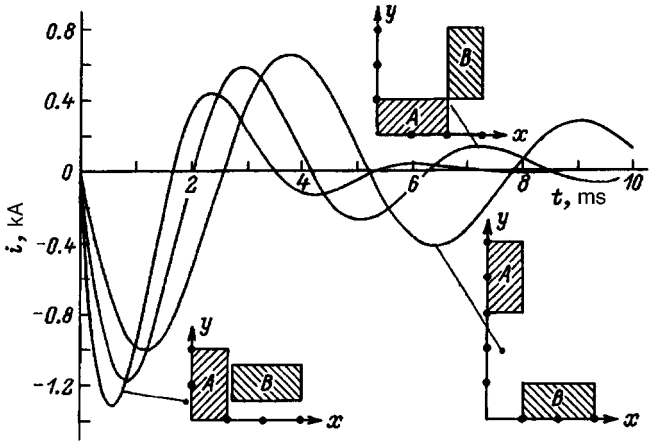

FIG. 5. Discharge current of a capacitor through short-circuited busbars *A* and *B* of identical cross section (0.01×0.02 m) in various configurations.

CONCLUSIONS

The approach considered is applicable to the investigation of transient processes with any degree of manifestation of the skin effect in two-dimensional problems. The reliability of the methodology is confirmed by the practically complete agreement between the results obtained on its basis and the solutions for one-dimensional problems having solutions in analytical form. Owing to the clear-cut physical basis of the method, it can easily be adapted to taking into account the proximity effect in the analysis of transient processes in systems of busbars running near not only real, but also ideal shields.

[1] G. A. Shneerson, *Fields and Transient Processes in Ultrastrong-Current Devices* [in Russian], Energoatomizdat, Moscow (1992).
[2] C. I. Mosani, IEEE Trans. Power Appar. Syst. **91**, 844 (1972).
[3] M. G. Vitkov, Zh. Tekh. Fiz. **35**, 410 (1965) [Sov. Phys. Tech. Phys. **10**, 324 (1965)].
[4] V. I. Yurchenko, Zh. Tekh. Fiz. **43**, 1866 (1973) [Sov. Phys. Tech. Phys. **18**, 1180 (1973)].
[5] V. M. Mikhaĭlov, *Pulsed Electromagnetic Fields* [in Russian], Vishcha Shkola, Kharkov (1979).
[6] W. R. Smythe, *Static and Dynamic Electricity*, 2nd ed., McGraw-Hill, New York (1950) [Russ. transl., IL, Moscow (1954)].

Translated by P. Shelnitz

Theory of one-dimensional and quasi-one-dimensional heat conduction

S. O. Gladkov

N. N. Semenov Institute of Chemical Physics, Russian Academy of Sciences, 117977 Moscow, Russia
(Submitted February 26, 1996)
Zh. Tekh. Fiz. **67**, 8–12 (July 1997)

It is shown that the heat conduction process in a one-dimensional flow of a fluid moving with a velocity **V** in a constant temperature field follows a law that is considerably more complicated than an ''ordinary'' exponential law. It is demonstrated that in the quasi-one-dimensional case the heat conduction process in an abstract space of dimension $1+\varepsilon$, where ε varies from zero to unity, is described by a modified Fourier equation. Its solution for an infinite space is found. © *1997 American Institute of Physics.* [S1063-7842(97)00207-9]

Problems associated with the investigation of the physical properties of structures of fractional dimension have become crucially important in the recent period. Most of the interest in them has been displayed since the foundations of fractal theory, percolation theory, and renormalization group analysis were laid.[1–5] One of the most widely encountered examples of a quasi-one-dimensional structure is a coastline. Its dimension is $1+\varepsilon$, where the number ε is considerably less than unity. One more example, but of a purely technical character, is a set of one-dimensional water pipes that are randomly arranged in a plane and connected to one another in an arbitrary manner. Such a quasi-one-dimensional structure is the subject of investigation in the present work. The crux of the problem is to ascertain the features of the heat conduction process in systems of dimension $1+\varepsilon$, in which a liquid (or gas) moves with a velocity **V** in a certain direction x. The boundary conditions are as follows. At the interface between the fluid and the external medium it is assumed that the external medium acts as a thermostat with a constant temperature T_0. Initially (at $t=0$), the temperature of the fluid is $T(x)$. When concrete initial and boundary conditions are assigned, the quasi-one-dimensional problem can be solved only if the solution of the one-dimensional problem is known. Therefore, we first study the heat conduction process for a one-dimensional flow of a liquid (or gas) with a temperature T_1 moving within another liquid (or gas) with a temperature T_2.

It is usually assumed[6] that the flow velocity can be taken into account by introducing it into the boundary conditions. Let us imagine the following picture. In the initial moment a fluid stream flows into another fluid, which is confined to a very narrow, but finite cylindrical reservoir of radius r. The channel length L significantly exceeds that radius: $L \gg r$. The system of coupled equations that takes into account the exchange of heat between the two media can be written in the following, highly general form (see, for example, Ref. 7):

$$\partial T_1/\partial t + V\partial T_1/\partial x = \chi_1 \partial^2 T_1/\partial x^2 + \alpha_{12}(T_2 - T_1), \qquad (1)$$

$$\partial T_2/\partial t = \chi_2 \partial^2 T_2/\partial x^2 + \alpha_{21}(T_1 - T_2), \qquad (2)$$

where T_1 is the temperature of the fluid, χ_1 is its thermal diffusivity, α_{12} is the coefficient of heat transfer from the fluid to the shell surrounding it, χ_2 is the thermal diffusivity of the shell, and α_{21} is the coefficient of heat transfer from the shell to the fluid.

It should be noted that consideration of the velocity **V** makes the last term in Eq. (1) necessary. In fact, since the characteristic ranges of variation of the temperature are $\delta x \gg l$, where l is the mean free path of the molecules, the condition that the second term on the left-hand side of Eq. (1) is smaller than the last term on the right-hand side can be written in the form of the following inequality: $V/\delta x \ll \alpha_{12}$. Hence it follows that $\delta x \gg V\tau_{12} = Vl/v_T$, and this condition always holds for all $V < v_T$. The latter calls for the consideration of heat transfer on the right-hand side of the heat conduction equation when the fluid undergoes hydrodynamic motion. We stress once again that if the fluid is motionless, the last term in Eq. (1) vanishes.

Thus, the system of equations presented faithfully describes the establishment of the temperature in a one-dimensional flow of a fluid moving along the x axis when the mutual influence of the two fluids on one another is taken into account. It should be noted that the problem formulated in this manner has not been solved in any publication known to us. For this reason, the solution of the system of equations (1) and (2) is of interest in itself not only from a purely cognitive standpoint, but also from a methodical standpoint.

We choose the initial conditions in the following form:

$$T_1(t=0, x) = T_0(1 - \tanh(x/d)), \qquad (3)$$

$$\lim T_1(t,x) = T_0, \quad t \to \infty, \qquad (4)$$

where d is a certain characteristic distance, to which the fluid penetrates at the initial moment in time, and T_0 is the constant temperature of the moving fluid at the initial moment.

The solution of the system of differential equations (1) and (2) can be sought in this case using the Laplace transformation. The conditions for doing so are that the time interval be assigned in the range from 0 to $+\infty$ and that the region for spatial variations of the temperature also be defined by a semi-infinite interval.

Finding an expression for T_2 from Eq. (1) and substituting it into Eq. (2), we find the following fairly cumbersome equation:

 1063-7842/97/070724-04$10.00

$$(1+\alpha_{21}/\alpha_{12})\partial T_1/\partial t+\alpha_{12}^{-1}\partial^2 T_1/\partial t^2+V\alpha_{12}^{-1}\partial^2 T_1/\partial t\partial x$$
$$-\chi_2 V\alpha_{12}^{-1}\partial^3 T_1/\partial x^3+\chi_1\chi_2\alpha_{12}^{-1}\partial^4 T_1/\partial x^4$$
$$+\alpha_{21}\alpha_{12}^{-1}(V\partial T_1/\partial x-\chi_1\partial^2 T_1/\partial x^2)=0. \quad (5)$$

Next, expanding the function $T_1(t,x)$ into a Laplace integral using the transformations

$$T_1(t,x)=\int_{\sigma-i\infty}^{\sigma+i\infty} e^{qx}T_{1q}(t)dq/2\pi i,$$

$$T_{1q}(t)=\int_0^{+\infty} e^{-qx}T_1(t,x)dx,$$

we find the equation which describes the temporal evolution of the temperature

$$d^2T_{1q}/dt^2+\varphi_{1q}dT_{1q}/dt+T_{1q}\varphi_2=0, \quad (6)$$

where

$$\varphi_{1q}=\alpha_{12}+\alpha_{21}+Vq-(\chi_1+\chi_2)q^2,$$
$$\varphi_{2q}=\chi_1\chi_2q^4-V\chi_2q^3-\alpha_{21}\chi_1q^2+Vq\alpha_{21}. \quad (7)$$

Let us consider the case in which the amount of heat transferred from the outer fluid to the moving fluid is small in comparison with the amount of heat transferred from the moving fluid to the outer fluid, i.e., $\alpha_{12}\ll\alpha_{21}$. In this approximation the roots of the characteristic equation corresponding to Eq. (6) are

$$k_{1,2}(q)=0.5[(\chi_1+\chi_2)q^2-\alpha_{21}-Vq]$$
$$\pm\{[(\alpha_{21}-(\chi_2-\chi_1)q^2)/2]^4$$
$$-V^2q^2(0.5\alpha_{21}-\chi_2q^2)^2\}^{1/4}e^{0.5iz(q)}, \quad (8)$$

where

$$z(q)=\arctan[4iqV(0.5\alpha_{21}-\chi_2q^2)/(\alpha_{21}-(\chi_2-\chi_1)q^2)].$$

As is easily understood, only the root k_1 (the radical with the + sign) has physical meaning. In this case the solution of the equation is

$$T_{1q}(t)=C_{1q}e^{k_1(q)t},$$

where C_{1q} is an integration constant.

According to the inverse Laplace transformation, we have

$$T_1(t,x)=\int_{\sigma-i\infty}^{\sigma+i\infty} e^{qx+k_1(q)t}dq/2\pi i,$$

and with consideration of the initial condition (3) we find

$$T_1(t,x)=\int_0^{\infty}T_1(0,y)dy\int_{\sigma-i\infty}^{\sigma+i\infty}e^{k_1(q)t-q(x-y)}dq/2\pi i, \quad (9)$$

where $T_1(0,x)$ is given by (3).

Unfortunately, the integral (9) has a fairly complicated form, and some limiting physical cases should be considered to calculate it. To start, let us consider the case in which the flow velocity V is small. More specifically, as follows from (8), it should satisfy the inequality

$$V\ll\alpha_{21}^{1/2}(\chi_2-\chi_1)^{3/2}(3\chi_2-\chi_1). \quad (10)$$

This condition corresponds, in particular, to a small value for the expression under the radical sign in $k_1(q)$. As a result, k_1 turns out to be

$$k_1(q)\cong 0.5qV+\chi_1q^2+(2qV(0.5\alpha_{21}-\chi_2q^2))/(\alpha_{21}$$
$$-(\chi_2-\chi_1)q^2). \quad (11)$$

Substituting this expression into the integral (9), after the replacement $q\Rightarrow -1q$ we obtain a formula, which describes the evolution of the temperature of a one-dimensional flow of the fluid undergoing laminar motion (small Reynolds numbers),

$$T_1(t,x)=\int_0^{\infty}T_1(0,y)dy\int_{-\infty}^{\infty}e^{i\varphi(q)t+iq(x-y)-\chi_1q^2t}dq, \quad (12)$$

where

$$\varphi(q)=qV(0.5+(\alpha_{21}+2\chi_2q^2))/(\alpha_{21}+(\chi_2-\chi_1)q^2). \quad (13)$$

If the flow is motionless ($V=0$), (12) automatically transforms into the known expression.[6]

To find the distribution of the temperature in the flow when $V\neq 0$, we should evaluate the exponent and isolate the saddle point. As is seen from (13), the function $\varphi(q)$ varies in a fairly narrow range from $1.5qV$ to $(0.5+2\chi_2/(\chi_2-\chi_1))$. Since the integrand achieves its largest value when $q\Rightarrow 0$ ($q\Rightarrow 1/L$, where L is the linear dimension of the system), $\varphi(q)$ can be replaced by $1.5qV$. In such a case

$$T_1(t,x)=T_0+\int_0^{\infty}T_1(0,y)dy$$
$$\times\int_{-\infty}^{+\infty}e^{iq(x-y)+1.5qVt-\chi_1q^2t}dq/2\pi$$
$$=T_0+\int_0^{\infty}T_1(0,y)e^{-(x-y+1.5Vt)^2/4\chi_1t}dy. \quad (14)$$

In the other limiting case, in which the velocity V is large and satisfies the condition

$$V\gg\alpha_{21}^{1/2}(\chi_2-\chi_1)^{3/2}/(3\chi_2-\chi_1), \quad (15)$$

the function k_1 (after the replacement $q\Rightarrow iq$) can be represented in the form

$$k_1q=0.5[iVq-(\chi_1+\chi_2)q^2-\alpha_{21}]$$
$$+\{Vq(0.5\alpha_{21}+\chi_2q^2)\}^{1/2}. \quad (16)$$

The maximum of this expression is found at

$$q_0=(V\alpha_{21})^{1/3}/2^{5/3}(\chi_1+\chi_2)^{2/3}. \quad (17)$$

Expanding the exponent near the point $q=q_0$, we find that at large velocities

$$T_1(t,x)=T_0+e^{-0.5\alpha_{21}t+iq_0x}$$
$$\times\int_0^{+\infty}T_1(0,y)e^{-iq_0y-(x-y)^2/4(\chi_1+\chi_2)t}dy. \quad (18)$$

It should be noted that Eq. (18) is valid provided $T_1(t,x) > T_0$. It can concluded from (18) that heat exchange is considerably more intense at large flow velocities, since the time for the establishment of a certain equilibrium temperature is determined not by heat conduction, but purely by heat transfer and corresponds in order of magnitude to $2/\alpha_{21}$.

The problem posed of ascertaining the character of the establishment of the temperature in a flow of a liquid (or gas) moving within another fluid raises the question of analyzing the heat conduction process in systems with a fractional dimension. Such structures include, for example, a system of randomly distributed pipes located in a single plane. Its dimension is $1+\varepsilon$. Clearly, when $\varepsilon=0$, the system simply becomes one-dimensional, and when $\varepsilon=1$, it becomes two-dimensional. The intermediate case of arbitrary ε is of interest in regard to purely methodical aspects of the case, which will be analyzed below.

It should be noted that the description of the character of heat conduction in a quasi-one-dimensional system is of fundamental importance from the standpoint of simulating complicated branched structures in order to adequately describe how thermal equilibrium is established in such structures and how the temperature is established. It is important to stress that ε is the only arbitrary parameter and that, once it is assigned, an attempt can be made to describe these macroscopically complicated systems.

Before proceeding to a solution of the heat conduction equation in the quasi-one-dimensional case, we must understand what the gradient operator represents in the present case. We assign a certain operator A acting in a Hilbert space of dimension $1+\varepsilon$ in the form of the following Fourier integral

$$Af(x)=\int_{-\infty}^{+\infty} ik^{1+\varepsilon}e^{ikx}dk/2\pi. \tag{19}$$

When $\varepsilon=0$, we obtain the ordinary result, which obviously indicates that A is simply an operator of differentiation with respect to x. It should be noted that the transition to the purely two-dimensional case should be made not by means of (19), but by means of the two-dimensional expansion, i.e.,

$$Af(x,y)=\lim_{\eta\Rightarrow 0}\int_{-\infty}^{+\infty} ik^{2-\eta}e^{i\mathbf{k}\cdot\mathbf{x}}d\mathbf{k}/(2\pi)^2, \tag{20}$$

where $\eta>0$.

After defining the action of the operator A in the form of (19), we can now introduce the concept of a quasi-one-dimensional heat flow. More specifically, assuming that

$$q=-\varkappa AT(x,t), \tag{21}$$

we find the Fourier equation sought in the quasi-one-dimensional case

$$\partial T/\partial t=-\chi\int_{-\infty}^{+\infty} k^{2(1+\varepsilon)}e^{ikx}T_k(t)dk/2\pi, \tag{22}$$

where χ is the thermal diffusivity in a space of dimension $1+\varepsilon$.

It should be noted that Eq. (22) ''works'' for an infinite space. Passage to a finite region of the space requires modification of this equation and a different definition for the operator A. This problem will not be considered in the present paper.

We next assign the properties of the operator A in the following manner: 1) the action of this operator on $f(x)$ has the fully concrete eigenvalues λ, i.e., $Af(x)=\lambda f(x)$; 2) the operator A is a bounded self-adjunct operator, i.e., $A=A^*$.[9]

According to Eq. (22), for the Fourier transform of the temperature we obtain

$$\partial T_k(t)/\partial t=-\chi k^{2(1+\varepsilon)}T_k(t). \tag{23}$$

Solving the equation obtained and performing the inverse Fourier transformation, we find

$$T(x,t)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T(\xi,0)e^{-\chi t k^{2(1+\varepsilon)}+ik(x-\xi)}dk\,d\xi/(2\pi)^2. \tag{24}$$

We calculate the integrals over k. Utilizing the saddle-point method for this purpose, we find that the saddle point lies at the value

$$k=k_0=[(iy)^{1/(1+2\varepsilon)}]/[2(1+\varepsilon)\chi t]^{1/(1+2\varepsilon)}, \tag{25}$$

where $y=x-\xi$.

Substituting this expression into the integrand in (24) and calculating the Gaussian integral obtained as a result, we find the following expression for the distribution of the temperature in a quasi-one-dimensional space:

$$T(x,t)=\int_{-\infty}^{+\infty} T(\xi,0)e^{Y(\varepsilon)}\cos(Z(\varepsilon))d\xi/2(\pi X(\varepsilon))^{1/2}, \tag{26}$$

where

$$\begin{aligned}
X(\varepsilon)&=\{u^{\varepsilon/(1+2\varepsilon)}\chi t(1+\varepsilon)(1+2\varepsilon)\\
&\quad\times\cos[\pi\varepsilon/(1+2\varepsilon)]\}/(1+\varepsilon)^{2\varepsilon/(1+2\varepsilon)},\\
Y(\varepsilon)&=(1+2\varepsilon)\cos[\pi(1+\varepsilon)/(1+2\varepsilon)]\\
&\quad\times u^{(1+\varepsilon)/(1+2\varepsilon)}/2(1+\varepsilon)[(1+\varepsilon)]^{1/1+2\varepsilon},\\
Z(\varepsilon)&=\tan(\pi\varepsilon/(1+2\varepsilon))-\varepsilon u^{(1+\varepsilon)/(1+2\varepsilon)}\\
&\quad\times(1+2\varepsilon)/2(1+\varepsilon)[(1+\varepsilon)]^{1/(1+2\varepsilon)},\\
u&=(x-\xi)^2/2\chi t.
\end{aligned} \tag{27}$$

In the limiting case, i.e., when ε tends to zero and the space becomes one-dimensional, we obtain a single expression for the thermal conductivity.[6] An analysis of (26) for several different values of ε is illustrated in Fig. 1.

It should be noted that when two-phase structures (for example, a two-component composite or a structure of the liquid+gas type) are treated, the thermal diffusivity can be represented in the following approximate form:[8]

$$\chi(p)=[p^2\varkappa_1+(1-p)^2\varkappa_2]/[pc_{1v}+(1-p)c_{2v}],$$

where the concentration $p=V_1/V_0$, V_1 is the volume of the impurity phase, $\varkappa_1$ and $\varkappa_2$ are the thermal conductivities of the first and second phases, respectively, c_1 and c_{2v} are their

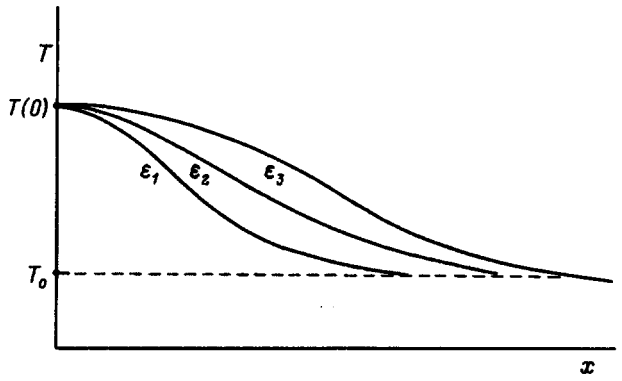

FIG. 1. Dependence of the temperature distribution on the coordinate x for several different values of ε ($\varepsilon_1<\varepsilon_2<\varepsilon_3$).

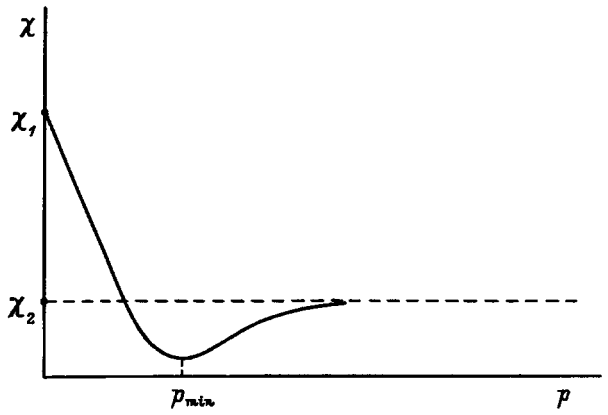

FIG. 2. Dependence of the thermal diffusivity of a structure on the concentration of the impurity phase p.

isochoric heat capacities per unit volume, and the denominator is obtained from the condition that the entropies of the two phases are additive.

More specifically, since the entropy of the entire structure is $S=pS_1+(1-p)S_2$, we can obtain the denominator sought by differentiating the entropy with respect to the temperature.

In addition, the relation presented for $\chi(p)$ is valid under the assumption of a weak interaction between the two phases, for which the term proportional to the product $p(1-p)$ is small. It was shown in Ref. 8 that this condition is realized quite frequently.

The minimum of the function $\chi(p)$ lies at the concentration

$$p_{\min}=c_{2v}/(c_{2v}-c_{1v})+|c_{1v}-c_{2v}|^{-1}\times\{(c_{2v}^2\varkappa_1+c_{1v}^2\varkappa_2)/(\varkappa_1+\varkappa_2)\}^{1/2}.$$

The corresponding value of the thermal diffusivity is

$$\chi_{\min}=2(c_{1v}-c_{2v})^{-2}\times[(c_{2v}^2\varkappa_1+c_{1v}^2\varkappa_2)(\varkappa_1+\varkappa_2)-c_{1v}\varkappa_2-c_{2v}\varkappa_1]^{1/2}.$$

A plot of $\chi(p)$ is shown in Fig. 2.

It is important to note that the thermal diffusivity minimum is achieved only in two-phase structures. In ''ordinary'' (single-phase) substances the thermal diffusivity depends only on the temperature and not on the concentration. This allows us to draw the unequivocal conclusion that the process leading to equalization of the temperature in two-phase structures at the concentration $p=p_{\min}$ takes considerably more time than in any single-phase substances (regardless of their dimension) and that cooling is, therefore, slowed. This can be very important from the practical standpoint.

[1] K. Wilson and G. Kogut, Phys. Rep. **12**, 75 (1974).
[2] A. Z. Patashinskiĭ and V. L. Pokrovskiĭ, *Fluctuation Theory of Phase Transitions*, Pergamon Press, Oxford (1979).
[3] B. I. Shklovskiĭ and A. L. Éfros, *Percolation Theory* [in Russian], Nauka, Moscow (1982).
[4] Yu. A. Izyumov and Yu. N. Skryabin, *Statistical Mechanics of Magnetically Ordered Systems*, Consultants Bureau, New York (1988).
[5] S. O. Gladkov, Phys. Lett. A **198**, 151 (1995).
[6] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, 2nd ed., Oxford University Press (1959) [Russ. transl., Nauka, Moscow (1964)].
[7] S. O. Gladkov, Solid State Commun. **94**, 787 (1995).
[8] S. O. Gladkov, Physica (Amsterdam) B **167**, 159 (1990).
[9] V. I. Smirnov, *A Course of Higher Mathematics, Vol. 5: Integration and Functional Analysis*, Addison-Wesley, Reading, Mass. (1964).

Translated by P. Shelnitz

Spectral and angular distribution of slow electrons emitted by hydrogen atoms in collisions with fast highly charged ions

A. B. Voĭtkiv

Institute of Electronics, Academy of Sciences of Uzbekistan, 700143 Tashkent, Uzbekistan
(Submitted November 13, 1995; resubmitted June 5, 1996)
Zh. Tekh. Fiz. **67**, 13–18 (July 1997)

The ionization of hydrogen atoms with the emission of slow electrons in collisions with fast highly charged ions is considered. Analytical expressions are obtained for the singly and doubly differential ionization cross sections and for a quantity which characterizes the angular asymmetry in the escape of slow electrons. A unique feature of the momentum balance for collisions that lead to the emission of slow electrons is discussed. © *1997 American Institute of Physics.* [S1063-7842(97)00307-3]

The investigation of the collisions of atoms with fast highly charged ions is of interest both for several branches of physics (atomic physics, solid-state physics, plasma physics, etc.) and for other closely related areas of science (for example, biophysics). The charges Z of these ions are frequently so great that, despite the large value of their velocity v ($v \gg v_0$, $v_0 = 1$ a.u. $\simeq 2\times 10^8$ cm/s), the relation $Z \gtrsim v$ holds (here and below, atomic units are used, unless stated otherwise). The total cross sections for the single, double, and multiple ionization of atoms in collisions with such ions have been the subject of numerous experimental and theoretical studies (see, for example, Refs. 1–4 and the literature cited therein). More detailed information on collisions between a fast highly charged ion and an atom can be obtained by investigating the various differential cross sections. The immense progress that has been achieved in recent years in developing experimental techniques has made it possible to perform so-called kinetamically complete experiments on the collisions of fast charged particles with atoms,[5,6] in which not only the total ionization cross sections, but also the angular and energetic distributions of the electrons leaving the atom, the momenta and energies of the recoil ions, etc. were determined. Such investigations were performed for the single ionization of helium atoms in collisions with fast highly charged ions ($Z = 24$, $v = 12$) characterized by a small amount of momentum imparted to the atom (such collisions will be called ''soft'' collisions below) in Ref. 6, where the characteristics of slow electrons with energies following escape from an atom that do not appreciably exceed its ionization potential, as well as the momentum balance in a system consisting of a highly charged ion, an electron, and a recoil ion, were studied [a classical-trajectory Monte Carlo (CTMC) calculation of such a collision process was also performed in that work]. The study of the characteristics of slow electrons is of great importance, since they comprise the bulk of the electrons emitted by atoms in collisions with fast highly charged ions in the range of values of Z and v considered.

In the present work we investigate the ionization of hydrogen atoms in ''soft'' collisions with fast highly charged ions, in which the atoms emit slow electrons when the parameters of the problem are as follows: $v \lesssim Z \ll v^2$, $v \gg v_0$, where v_0 is the characteristic orbital velocity of the electron in the ground state of the atom. We shall examine the energetic and angular distributions of the slow electrons ($v_e \lesssim v_0$, where v_e is the velocity of the emitted electron relative to the recoil ion), and we shall briefly discuss a unique feature of the momentum balance in such collisions on the basis of the analogy to photoionization.

Let a hydrogen atom with a nucleus lying at the origin of coordinates be initially in the ground state, and let a structureless, highly charged ion move along the classical linear trajectory $\mathbf{R}(t) = \mathbf{b} + \mathbf{v}t$, where $\mathbf{b}$ is the impact parameter vector. We divide the entire range of values of the impact parameter $0 \leqslant b < \infty$ into two subranges: 1) $b \lesssim Z/v$, 2) $b > Z/v$.

In collisions in the first subrange considerable energy, which significantly exceeds the ionization potential of the atom, is imparted on the average to the electron. In fact, when $b \gtrsim 1$, this energy can be estimated as $\varepsilon(b) \simeq 2Z^2/(b^2v^2)$,[7] and it is already fairly high at $b \simeq Z/v > 1$ and increases rapidly as the impact parameter decreases [for example, for the ions with $Z = 24$ and $v = 12$ used in Ref. 6, $\varepsilon(b=1) = 8$]. When $b < 1$, the mean energy imparted is even higher. Because of this (as well as the small width of the range $b < Z/v$), the contribution of the ''hard'' collisions to the emission of slow electrons is small, and these collisions will not be considered below (for information on the spectra of electrons emitted with a large momentum transfer, see, for example, Refs. 8–11 and the literature cited therein).

In the range of impact parameters $b > Z/v > 1$ calculations in various approximations[12–14] predict that the ionization probability decreases rapidly with increasing b and becomes far smaller than unity already at $b \simeq 1.5Z/v - 2Z/v$. To describe the transitions of an atom in such collisions we use the scattering matrix formalism, in which the transition amplitude of the atom has the form

$$A_{\mathbf{k}} = -i \int_{-\infty}^{+\infty} dt \langle \Psi_{\mathbf{k}}^{(-)}(\mathbf{r},t) | W(\mathbf{r},t) | \phi_0(\mathbf{r},t) \rangle, \qquad (1)$$

where $\phi_0(\mathbf{r},t) = \phi_0(\mathbf{r})\exp(-it^2/2)$, $\phi_0(\mathbf{r}) = \pi^{-1/2}\exp(-r)$ is the wave function of the ground state of the hydrogen atom, $\psi_{\mathbf{k}}^{(-)}(\mathbf{r},t)$ is the wave function of the electron in the final state in the simultaneous presence of the fields of the nucleus

and the highly charged ion, $W(\mathbf{r},t)=Z/|\mathbf{R}(t)|-Z/|\mathbf{R}(t)-\mathbf{r}|$ is the interaction of the atom with the field of the highly charged ion, and $\mathbf{r}$ is the radius vector of the electron.

Let us estimate the relative influence of both centers on the electron in the final state $\Psi_{\mathbf{k}}^{(-)}(\mathbf{r},t)$ from the ratio between the classical forces exerted on it by the fast highly charged ion (F_i) and the nucleus of the atom (F_a). The field of the fast highly charged ion in the region where the atom is located has a maximum as a function of time with a center at the point $t=0$ and an effective width $T\simeq b/v$ ($b>1$, see Ref. 15 and 7). In the range of impact parameters $b>v$ the field of the highly charged ion at the atom is not only already very weak (when $v^2\gg Z$), but also adiabatically slowly varying, and the ionization probability of the atom is then exponentially small (see, for example, Ref. 12). In the collisions with $Z/v<b<v$, which make the main contribution in the presence of slow electrons, $T<1$, i.e., at such values of the impact parameter the field of the highly charged ion has a fairly sharp maximum at the times $|t|<T$, at which ionization occurs for the most part. Then the distance between the proton and the electron leaving it is estimated as $\overline{v}\,t$ ($t>0$), where $\overline{v}$ is the mean velocity at which the electron passes through the region of space $r\sim 1$ and which, for slow electrons, is equal in order of magnitude to $v_0=1$. The distance between the highly charged ion and the electron at $t>0$ can be assumed to be proportional to the difference between their velocities: $|\mathbf{v}-\overline{\mathbf{v}}|t\simeq vt$. Thus, for the ratio between the forces we have $F_i/F_a\sim Zv_0^2/v^2=Z/v^2$,[1)] whence it follows that for $Z/v^2\ll 1$ the behavior of the slow electron in the final state is ''controlled'' mainly by the field of the nucleus of the atom. Therefore, the influence of the Coulomb interaction of a fast highly charged ion with such an electron can be taken into account approximately. As will be seen below, the main effect of this influence is asymmetry in the angular distribution of the slow electrons, and the majority of them are dragged by the Coulomb attraction of the recoiling highly charged ion in the direction of its motion.

There is another effect, which leads to asymmetry in the angular distribution of the slow electrons. Distributing the interaction potential of the atom with the field of the highly charged ion ($Z/|\mathbf{R}(t)|-Z/|\mathbf{R}(t)-\mathbf{r}|$) as a function of the coordinates of the electron and the time among monochromatic plane waves, we can easily see that the field of the highly charged ion has a longitudinal (parallel to the velocity of the ion) momentum $q_p\sim 1/v$ and that its absolute value and direction do not depend (when $Z\neq 0$) on the magnitude and sign of the charge of the impinging particle (it is important to note that if only the dipole term in the expansion of this interaction is taken into account, the calculated value of q_p vanishes). The absorption of this momentum by an atomic electron leads (see below) to additional asymmetry in the angular distribution of the slow electrons.

We represent the interaction $W(\mathbf{r},t)$ in the range of values of the impact parameter $b>Z/v>1$ in the form

$$W(\mathbf{r},t)=W_1(\mathbf{r},t)+W_2(\mathbf{r},t), \tag{2}$$

where

$$W_1(\mathbf{r},t)=-\frac{Z(vtz+by)}{R^3(t)}=-\frac{Z\mathbf{R}(t)\mathbf{r}}{r^3(t)}=-\mathbf{E}(t)\mathbf{r},$$

$$W_2(\mathbf{r},t)=\frac{Zr^2}{2R^3(t)}-\frac{3Z}{2R^5(t)}(vtz+by)^2. \tag{3}$$

In Eq. (3) the z axis is directed along the velocity vector of the highly charged ion, and the y axis is directed along the impact parameter vector. In accordance with the foregoing, the interaction $W(\mathbf{r},t)$ has been expanded in (2) to the quadrupole terms, inclusively.

The wave function $\Psi_{\mathbf{k}}^{(-)}(\mathbf{r},t)$ of the electron in the final state in the field of the two centers satisfies the Schrödinger equation

$$i\frac{\partial}{\partial t}\Psi_{\mathbf{k}}^{(+)}=\left(-\frac{\Delta}{2}-\frac{1}{r}-W_1-W_2\right)\Psi_{\mathbf{k}}^{(+)}. \tag{4}$$

In accordance with the arguments presented above regarding the relative roles of the two centers in the final state, we take into account the interaction of the slow electron with the nucleus of the atom precisely and the interaction with the field of the fast highly charged ion in the zeroth sudden approximation[16]

$$\Psi_{\mathbf{k}}^{(+)}(\mathbf{r},t)=\phi_{\mathbf{k}}^{(+)}(\mathbf{r})\exp\left(-ik^2t/2-i\int_{-\infty}^{t}dt'\,W(t')\right). \tag{5}$$

Here $\phi_{\mathbf{k}}^{(+)}(\mathbf{r})$ is the Coulomb wave function for the scattering of an electron on a proton, which, at $r\to\infty$, is the result of the superposition of the incident ''plane'' wave on the diverging spherical wave and is normalized according to the condition $\langle\phi_{\mathbf{k}}^{(+)}|\phi_{\mathbf{k}}^{(+)}\rangle=(2\pi)^{-3}\delta(\mathbf{k}-\mathbf{k}')$, where $\mathbf{k}$ is the wave vector of the motion of the electron relative to the nucleus of the atom. When (5) is substituted into (1) we have

$$A_{\mathbf{k}}=-i\left\langle\phi_{\mathbf{k}}^{(-)}(\mathbf{r})\left|\int_{-\infty}^{+\infty}dt\,W(\mathbf{r},t)\exp\left(i\omega t-i\int_{-\infty}^{t}dt'\,W(t')\right)\right|\phi_0(\mathbf{r})\right\rangle, \tag{6}$$

where $\omega=(1+k^2)/2$ is frequency of the transition, and $\phi_{\mathbf{k}}^{(-)}(\mathbf{r})=(\phi_{-\mathbf{k}}^{(+)}(\mathbf{r}))^*$.

For $b>Z/v>1$ the term $\int_{-\infty}^{t}dt'\,W(t')$ at any t is small compared with unity. Expanding the corresponding exponential function in (6) in a series and leaving the principal terms in (6), we obtain

$$A_{\mathbf{k}}=A_{\mathbf{k}}^d+A_{\mathbf{k}}^q+A_{\mathbf{k}}^i. \tag{7}$$

Here

$$A_{\mathbf{k}}^d=-i\left\langle\phi_{\mathbf{k}}^{(-)}(\mathbf{r})\left|\int_{-\infty}^{+\infty}dt\,W_1(\mathbf{r},t)\exp(i\omega t)\right|\phi_0(\mathbf{r})\right\rangle,$$

$$A_{\mathbf{k}}^q=-i\left\langle\phi_{\mathbf{k}}^{(-)}(\mathbf{r})\left|\int_{-\infty}^{+\infty}dt\,W_2(\mathbf{r},t)\exp(i\omega t)\right|\phi_0(\mathbf{r})\right\rangle \tag{8}$$

are the amplitudes of the dipole and quadrupole transitions, respectively, and

$$A_{\mathbf{k}}^{i}=-i\left\langle \phi_{\mathbf{k}}^{(-)}(\mathbf{r})\right| \int_{-\infty}^{+\infty} dt\ \exp(i\omega t) \times(\mathbf{q}(t)\mathbf{r})W_1(t)\left|\phi_0(\mathbf{r})\right\rangle \tag{9}$$

takes into account the principle term for the interaction of the slow electron in the final state with the field of the fast highly charged ion, where $\mathbf{q}(t)=\int_{-\infty}^{t}dt\mathbf{E}(t)$.

For the transition probability of an electron to a state with a definite value of $\mathbf{k}$ in a collision with the impact parameter b we have

$$w(\mathbf{k},\mathbf{b})=|A_{\mathbf{k}}|^2\simeq|A_{\mathbf{k}}^{d}|^2+2\ \mathrm{Re}(A_{\mathbf{k}}^{d})\mathrm{Re}(A_{\mathbf{k}}^{q}+A_{\mathbf{k}}^{i}) +2\ \mathrm{Im}(A_{\mathbf{k}}^{d})\mathrm{Im}(A_{\mathbf{k}}^{q}+A_{\mathbf{k}}^{i}). \tag{10}$$

Using the explicit forms of the interactions W_1 and W_2 and of the Coulomb wave functions and averaging (by virtue of the geometry of the problem) the probability (10) over the electron escape angle ϕ ($0\leqslant\phi<2\pi$, ϕ is the azimuthal angle in the plane of the impact parameter), we can obtain

$$\begin{aligned} w(k,\Theta,b)&=\frac{1}{2\pi}\int_0^{2\pi}d\phi w(\mathbf{k},\mathbf{b})\\ &=\frac{2Z^2\omega^2}{v^4}\alpha(k)\Bigg(2K_0^2(\xi)\cos^2\Theta+K_1^2(\xi)\sin^2\Theta\\ &+\frac{8k}{v}\cos\Theta(K_1^2(\xi)\sin^2\Theta+K_0^2(\xi)(3\cos^2\Theta\\ &-1))+\frac{4Z}{v^2}\frac{\cos\Theta}{\omega}\Bigg(K_0(\xi)K_1(\xi)\\ &-\frac{\pi k}{2}\exp(-\xi)(2K_1(\xi)\sin^2\Theta\\ &+K_0(\xi)(3\cos^2\Theta-1))\Bigg)\Bigg), \end{aligned} \tag{11}$$

where the electron escape angle Θ ($0\leqslant\Theta\leqslant\pi$) is measured from the direction of the velocity vector of the highly charged ion, $\xi=\omega b/v$, K_0 and K_1 are modified Bessel functions,[17] and

$$\alpha(k)=\frac{2^7k^{-1}}{(1+k^2)^5}\frac{\exp\left(-\frac{4}{k}\arctan k\right)}{(1-\exp(-2\pi/k))}. \tag{12}$$

For the probability of ionization with the escape of a slow electron in collisions with a fixed impact parameter we have

$$w(b)=\int_0^{k_{\max}}dkk^2\int d\Omega w(k,\Theta,b), \tag{13}$$

where $k_{\max}\simeq v_0=1$ and $d\Omega=2\pi\sin\Theta d\Theta$.

The doubly differential (with respect to the escape angle and with respect to the energy $E=k^2/2$) ionization cross section is defined by the expression

$$\frac{d^2\sigma}{dEd\Omega}=2\pi\int_{b_{\min}}^{\infty}dbb(2E)^{1/2}w(k(E),\Theta,b), \tag{14}$$

where $b_{\min}=\lambda Z/v\ll v$ (λ is a constant of order unity) is the lower bound of the range of impact parameters in which the ionization probability is appreciably less than unity.

Performing the integration in (14) over the impact parameter, we find

$$\begin{aligned} \frac{d^2\sigma}{dEd\Omega}&=2^8\frac{Z^2}{v^2}\frac{1}{(1+2E)^5}\frac{\exp(-(4/(2E)^{1/2})\arctan(2E)^{1/2})}{(1-\exp(-2\pi/(2E)^{1/2}))}\\ &\times(\sin^2\Theta\ \ln\beta+\cos^2\Theta-0.5\sin^2\Theta\\ &+2^{3.5}(E^{1/2}/v)\cos\Theta(\sin^2\Theta\ \ln\beta+\cos 2\Theta)\\ &+2(Z/v^2)\cos\Theta(\ln^2\beta-2^{1.5}\pi E^{1/2}(\sin^2\Theta\ \ln\beta\\ &+\cos 2\Theta-0.5\sin^2\Theta))), \end{aligned} \tag{15}$$

where $\beta=1.123v/(b_{\min}(E+0.5))=(1.123/\lambda)(v^2/Z)/(E+0.5)$.

Since the parameter β, which is determined to within a constant coefficient of order unity, contains the large cofactor v^2/Z and appears under the logarithm sign in Eq. (15), we simply assume that $\lambda=1$ in it. We note that the accuracy of the approach employed increases as the geometric dimensions of the range of impact parameters $Z/v<b<v$ increase.

The energy distribution of the electrons is specified by the differential cross section

$$\begin{aligned} \frac{d\sigma}{dE}&=\int d\Omega\frac{d^2\sigma}{dEd\Omega}=\frac{2^{11}\pi}{3}\frac{Z^2}{v^2}\frac{1}{(1+2E)^5}\\ &\times\frac{\exp(-(4/(2E)^{1/2})\arctan(2E)^{1/2})}{(1-\exp(-2\pi/(2E)^{1/2}))}\ \ln\left(\frac{2.25v^2}{Z(1+2E)}\right). \end{aligned} \tag{16}$$

As follows from (16), the electron escape probability decreases rapidly with increasing energy: the bulk ($\sim$90%) of the escaping electrons have energies that do not surpass the ionization potential of the atom $I_0=0.5$. We note that in the approximation under consideration contributions to (16) are made only by the dipole transitions between the states ϕ_0 and $\phi_{\mathbf{k}}^{(-)}$ due to the interaction $W_2(\mathbf{r},t)$ (which leads to electric monopole and quadrupole transitions) and that the transitions to final states distorted by the field of the highly charged ion, which are responsible for the asymmetry in the angular distribution of the slow electrons, do not make a contribution to (16) and, accordingly, do not influence the total number of slow electrons emitted in the approximation under consideration.

To obtain the angular distribution of these electrons, we must integrate (15) over the energies E in the range $0\leqslant E<I_0$. However, since the electron escape probability decreases rapidly with increasing E, the upper bound of the energy range can formally be set equal to infinity, and then we find

$$\frac{d\sigma}{d\Omega}=\int_0^{\infty}dE\frac{d^2\sigma}{dEd\Omega}=3(0.283)\frac{Z^2}{v^2}\Bigg(\sin^2\Theta\ \ln\beta_1$$

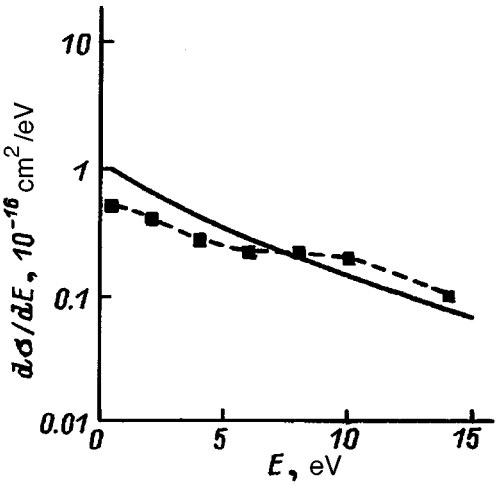

FIG. 1. Differential ionization cross section of hydrogen atoms with respect to the energy in collisions with highly charged ions for $Z=6$ and $v=5$. Solid line — calculation based on (16), ■ — data from Ref. 13.

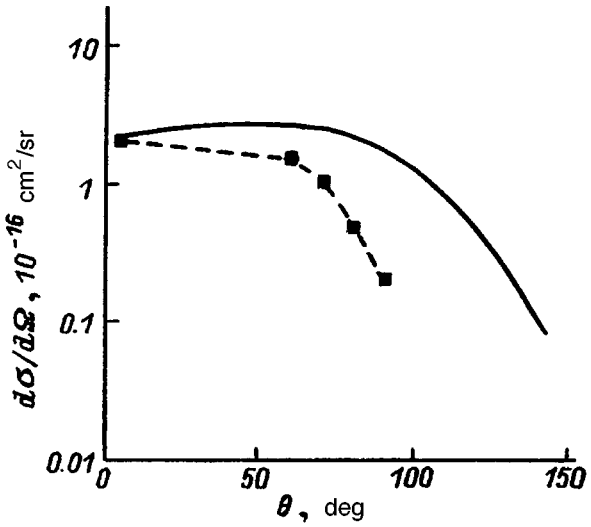

FIG. 2. Differential ionization cross section of hydrogen atoms with respect to the angle in collisions with highly charged ions for $Z=6$ and $v=5$. Solid line — calculation based on (17), ■ — data from 13.

$$+\cos^2\Theta-0.5\sin^2\Theta$$
$$+\frac{8(0.61)}{v}\cos\Theta(\sin^2\Theta\ \ln\beta_2+\cos 2\Theta)$$
$$+\frac{2Z}{v^2}\cos\Theta(\ln^2\beta_1+\langle\ln^2\omega\rangle-\ln^2\omega_1$$
$$-2\pi 0.61(\sin^2\Theta\ \ln\beta_2+\cos 2\Theta$$
$$-0.5\sin^2\Theta))\Bigg), \tag{17}$$

where

$$\beta_1=1.12v^2/(Z\omega_1),\quad \beta_2=1.12v^2/(Z\omega_2),$$
$$\omega_1=\exp\left(\int_0^\infty dk\alpha(k)\ln(\omega)/\int_0^\infty dk\alpha(k)\right)=0.71,$$
$$\omega_2=\exp\left(\int_0^\infty dkk\alpha(k)\ln(\omega)/\int_0^\infty dkk\alpha(k)\right)=0.81,$$
$$\langle\ln^2\omega\rangle=\int_0^\infty dk\alpha(k)\ln^2(\omega)/\int_0^\infty dk\alpha(k)=0.234,$$
$$\ln^2\omega_1=0.1. \tag{18}$$

Figure 1 presents a comparison of the differential cross section $d\sigma/dE$ found from Eq. (16) with the CTMC data in Ref. 13 for $Z=6$ and $v=5$. Our results for this cross section appreciably exceed the data from the calculation in Ref. 13 in the range of emitted electron energies $E\lesssim 5$ eV. Figure 2 compares the differential cross section $d\sigma/d\Omega$ calculated from Eq. (17) with the data from Ref. 13 for the same values of Z and v. The disparity between our results for $d\sigma/d\Omega$ and the CTMC calculation, which is not very appreciable at small values of the electron escape angle Θ, becomes significant as Θ increases. The disparity between our results for the differential cross sections $d\sigma/dE$ and $d\sigma/d\Omega$ and the data from the CTMC calculation can be attributed to the following known causes. According to classical mechanics, collisions with sufficiently large values of the impact parameter, in which the mean energy imparted to the atom is smaller than its ionization potential, do not make an appreciable contribution to the ionization process (the contribution of this range of values of b is ''classically suppressed''[18]). On the other hand, according to quantum mechanics, just this range of values of b makes the main contribution to electron emission when $v\gg v_0$, and in this case ionization takes place with the escape of slow electrons. This difference, in particular, leads to different asymptotes of the cross sections for single ionization at the collision velocities $v\gg v_0$ (the subscripts ''cl'' and ''qu'' refer to classical and quantum mechanics, respectively)

$$\sigma''_{\rm cl}v^{-2}(\text{Ref. 19}),\quad \sigma''_{\rm qu}(\text{const}_1+\ln v)\quad v^{-2}(v\gg v_0,Z)$$
$$\times(\text{Ref. 20}),$$
$$\sigma''_{\rm qu}(\text{const}_2+\ln(v^2/Z))v^{-2}\quad (Z\sim v\gg v_0)(\text{Ref. 1}).$$

In addition, as we know (see, for example, Ref. 18 and the literature cited therein), formidable difficulties arise when an attempt is made to use classical mechanics to describe the emission of slow electrons in directions corresponding to large values of Θ. Thus, the disparities between our calculations and the data from 13 can be attributed to the fact that the methods for calculating ionization based on classical mechanics are inapplicable for describing ''soft'' collisions.[21]

The asymmetry in the escape of electrons can be characterized by the parameter

$$\eta=\left(\int_0^{\pi/2}d\Theta\ \sin\Theta\frac{d\sigma}{d\Omega}-\int_{\pi/2}^{\pi}d\Theta\ \sin\Theta\frac{d\sigma}{d\Omega}\right)$$
$$\times\left(\int_0^{\pi}d\Theta\ \sin\Theta\frac{d\sigma}{d\Omega}\right)^{-1}. \tag{19}$$

From (18) and (19) we find

$$\eta=\frac{1.83}{v}+\frac{Z}{v^2}\left(1.5\ \ln\left(\frac{1.6v^2}{Z}\right)-2.87\right.$$
$$\left.+2.15\ \ln^{-1}\left(\frac{1.6v^2}{Z}\right)\right). \qquad (20)$$

It follows from Eq. (20) that the majority of the slow electrons are emitted by the atom (when $Z>0$) in the direction of motion of the fast highly charged ion. The first term in (20), which does not depend on the magnitude or sign of the charge of the highly charged ion, and the second term, which depends on both (when $Z<0$, $|Z|$ should be taken under the logarithm sign), are consequences, respectively, of the absorption of the longitudinal momentum $q_p\sim 1/v$ by the atom and the dragging of the electron leaving the atom by the electric field of the recoiling highly charged ion, which were discussed above. We note that the simple additivity of these two effects in (20) is a consequence of expansions (7) and (10), in which the terms that lead to the angular asymmetry (but are not the dominant terms for the total emission) are taken into account in the first nonvanishing approximation.

For the contribution to the ionization cross section of hydrogen atoms from collisions in which slow electrons escape, from (16) [or (17)] we obtain

$$\Delta\sigma=8\pi 0.283\ \frac{Z^2}{v^2}\ \ln\left(\frac{1.6v^2}{Z}\right). \qquad (21)$$

This expression practically coincides with the expression for the (total) ionization cross section found in Ref. 12 and is distinguished only by the numerical multiplier under the logarithm sign from the total ionization cross section calculated in Ref. 7

$$\sigma_i=8\pi 0.283\ \frac{Z^2}{v^2}\ \ln\left(\frac{5v^2}{Z}\right), \qquad (22)$$

which faithfully describes the existing experimental data for $v_0\ll v\lesssim Z\ll v^2$. It is seen from Eqs. (21) and (22) that $\Delta\sigma\simeq\sigma_i$ for $v_0\ll v\lesssim Z\ll v^2/3$, i.e., the collisions which result in the emission of slow electrons essentially determine the value of the ionization cross section in this case.

We note that all the cross sections (15)–(17), (21), and (22) as functions of the charge and velocity of the incident particle satisfy the scaling relation $\sigma/Z=f(v^2/Z)$, which is characteristic of the ionization cross section of hydrogen in collisions in the ranges of values of the parameters of the problem $Z/v>1$ and $v\gg v_0$.[1]

To simplify the treatment, we assumed above that the highly charged ion is structureless. It is obvious, however, that a fast highly charged ion bearing electrons can also be regarded as a point charge when $b>Z/v>1\gg r_Z$, where $r_Z\sim 1/Z$ is the radius of the ion.

In conclusion, let us briefly dwell on the features of the momentum balance for a collision of a fast highly charged ion with an atom when $b>Z/v$. In such collisions the mean momentum $Q\simeq Z\mathbf{b}/(b^2v)$ imparted by the field of the impinging particle to the atomic electron (see, for example, Ref. 17) is small in comparison with the characteristic momentum of the electron in the ground state of the atom $Q_0\simeq 1$. At the same time, the field of the fast highly charged ion contains the characteristic frequencies $\Omega\sim v/b$, which are small in comparison with the frequencies of atomic transitions even when $b\simeq v$. Therefore, the ionization of an atom in collisions with $b>Z/v$ is very similar[22–25,12] to the ionization of an atom by the field of a light wave when the atom absorbs a quantum, whose energy is sufficient for ionization, while its momentum is negligibly small. In the case of photoionization by a field of not excessively high frequency, the momentum of the escaping electron is balanced by the momentum of the atomic residue. The same situation clearly arises in the case of collisional ionization in the range $b>Z/v>1$, which leads to the escape of slow electrons, and was experimentally detected in Ref. 6 in an investigation of ''soft'' collisions with helium atoms.

[1)]For electrons with a velocity $v_0<v_e<v$ we have $\overline{v}\simeq v_e$ and $F_i/F_a\sim Zv_e^2/v^2\sim ZE/v^2$. Hence it follows that the influences of the two centers become comparable at electron energies $E\sim v^2/Z$ and that the influence of the highly charged ions predominates at $E\gg v^2/Z$.

[1] L. P. Presnyakov, V. P. Shevel'ko, and R. K. Janev [in Russian], Énergoatomizdat, Moscow (1986).
[2] C. L. Cocke and R. E. Olson, Phys. Rep. **205**, 205 (1991).
[3] J. H. McGuire, Adv. At. Mol. Opt. Phys. **29**, 217 (1992).
[4] H. Berg, Doctoral Thesis, Universität Frankfurt, 1993.
[5] J. Ullrich, R. Doerner, V. Mergel *et al.*, Preprint No. GSI-94-63.
[6] R. Moshammer, J. Ullrich, M. Unverzagt *et al.*, Phys. Rev. Lett. **73**, 3371 (1994).
[7] A. B. Voĭtkiv and A. V. Koval', Zh. Tekh. Fiz. **64** (3), 188 (1994) [Tech. Phys. **39**, 335 (1994)].
[8] P. D. Fainstein, V. H. Ponce, and R. D. Rivarola, J. Phys. B **24**, 3091 (1991).
[9] N. Stolterfoht, D. Schneider, J. Tanis *et al.*, Europhys. Lett. **4**, 899 (1987).
[10] N. Stolterfoht, H. Platten, G. Schiwirtz *et al.*, Phys. Rev. A **52**, 3796 (1995).
[11] J. H. Macek, in *Ionization of Solids by Heavy Particles*, R. A. Baragiola (ed.), Plenum Press, New York (1993).
[12] E. L. Duman, L. I. Men'shikov, and B. M. Smirnov, Zh. Éksp. Teor. Fiz. **49**, 516 (1979) [Sov. Phys. JETP **22**, 260 (1979)].
[13] C. O. Reinhold, C. A. Falcon, and J. E. Miraglia, J. Phys. B **20**, 3737 (1987).
[14] V. S. Nikolaev, V. A. Sidorovich, and V. N. Novozhilova, Zh. Éksp. Teor. Fiz. **101**, 1198 (1992) [Sov. Phys. JETP **74**, 640 (1992)].
[15] N. H. D. Bohr, *The Penetration of Atomic Particles through Matter*, E. Munksgaard, Copenhagen (1948) [Russ. transl., IL, Moscow (1950)].
[16] A. M. Dykhne and G. L. Yudin, Usp. Fiz. Nauk **125**, 377 (1978) [Sov. Phys. Usp. **21**, 549 (1978)].
[17] *The Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun (eds.), Dover, New York (1976) [Russ. transl., Nauka, Moscow (1969)].
[18] C. Reinhold and J. Burgdorfer, J. Phys. B **26**, 3101 (1993).
[19] J. J. Thomson, Philos. Mag. **23**, 449 (1912).
[20] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions*, 3rd ed., Oxford University Press, London (1965) [Russ. transl., Mir, Moscow (1969)].
[21] N. Stolterfoht, ''One- and Two-Center Electron Emission in Energetic Ion-Atom Collisions,'' Invited Talk at the Symposium for Two-Center Effects in Ion-Atom Collisions, Lincoln, 1994.
[22] E. Fermi, Z. Phys. **29**, 315 (1924).
[23] C. Weizsacker, Z. Phys. **88**, 612 (1934).
[24] E. Williams, Phys. Rev. **45**, 729 (1934).
[25] H. Bethe, Ann. Phys. (Leipzig) (5), 325 (1930).

Translated by P. Shelnitz

Plasma parameters of a nonself-sustained microwave discharge created by a programmed pulse

A. F. Aleksandrov, A. S. Zarin, A. A. Kuzovnikov, V. M. Shibkov, and L. V. Shibkova

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia
(Submitted August 14, 1996; resubmitted September 30, 1996)
Zh. Tekh. Fiz. **67**, 19–23 (July 1997)

The kinetics of the electrons in the plasma of a nonself-sustained discharge formed at the focus of a microwave beam when the gas is exposed to electromagnetic radiation with an energy flux density that varies with time in a programmed-pulse mode are investigated. It is shown that the temperature of the electrons in the plasma of the localized microwave discharge is of the order of 1 eV and varies weakly during the pump pulse and as the air pressure is varied, while the electron density is an order of magnitude or more lower than the critical density and depends on the level of the pump generator power. It is shown that the degree of ionization of the plasma can be regulated by altering the programmed-pulse mode. [S1063-7842(97)00407-8]

A freely localized microwave discharge appearing at the focus of a beam of electromagnetic energy[1–19] is a complicated nonlinear phenomenon, which includes nonstationary breakdown of the gas, propagation of the ionization fronts interacting with the radiation, maintenance of the nonequilibrium plasma formed at the focus of the beam by the incident energy flux, and excitation and heating of the molecules accompanied by deformation of the density of the neutral gas. It is known that if the electric field strength at the focus of a microwave beam exceeds the threshold value, breakdown of the gas occurs in that region of free space. After breakdown, the plasma formed at the beam focus begins to rapidly absorb the energy supplied to the discharge, and the microwave discharge in a beam with a sufficiently small convergence angle has a tendency to move toward the radiation source by some mechanism[20] (a breakdown wave, diffusion of resonant radiation, diffusion of charged particles, a regime of slow thermal-conductive heating, and several other processes). At the same time, the zone of effective energy release rapidly moves from the focus toward the energy flux, precluding strict fixation of the position of the discharge in space, and the energy supplied to the discharge is distributed over a large body of gas.

Many practical applications call for the localization of a microwave discharge at a fixed position in space. There are various techniques for localizing a microwave discharge in a focused beam. The formation of the discharge must take place with a minimum delay relative to the leading edge of the pulse of microwave radiation, and the leading edge of the breakdown wave must not deviate far from the focus during the entire time of action of the microwave energy on the plasma formed. As experiments performed by several investigators in different systems have shown, the use of a rectangular microwave pulse of long duration and large amplitude does not lead to the efficient introduction of energy at an assigned position in free space and, consequently, does not lead, in particular, to efficient heating of the gas.

Let us briefly examine several techniques for localizing a microwave discharge in a focused beam. For example, if short (of the order of microseconds) microwave pulses with a low repetition rate are used, the discharge can be confined to a given region in space.[15] However, the mean energy imparted to the discharge in this case is small. It is known that the repeated pulsed breakdown of air differs from the initial breakdown process, since the elimination of the charged particles after completion of the first pulse is not complete, and electrons, as well as positive and negative ions, remain in the region where the discharge existed when the next pulse begins. Considerable variation of the principal components of the gas associated with the appearance of long-lived electronically and vibrationally excited molecules and variation of the chemical composition of the neutral gas is also possible. The specific energy imparted to the plasma can be increased somewhat by increasing the pulse repetition rate up to a certain limit. However, as the pulse repetition rate is increased, there is an increase in the linear dimensions of the region where the discharge exists, and, for example, the efficiency of the heating of the gas does not increase.

To fix the location of an artificially ionized region in the Earth's atmosphere, several investigators[7,21,22] proposed creating such a region at the site of the intersection of two or more microwave beams. The energy of each beam would then be insufficient for breakdown of the gas, but the formation of a self-sustained microwave discharge would be possible at the site of intersection of the beams. In the case of the pulsed sustaining of a discharge in intersecting beams, the existence of a thin plasma layer capable of reflecting radio waves with frequencies up to 1 GHz is possible. The ionized region then acquires a definite structure, depending on the geometry of the beams and the pressure of the gas.

Another method for stopping a microwave discharge, viz., the programmed pulsed method, was proposed in our laboratory to localize a microwave discharge in a focused beam of electromagnetic waves at a fixed position in free space.[2,16,17] The essential point of this method is that breakdown of the gas is effected by a short power pulse, during which the discharge front does not manage to depart from the focal region, and the plasma is sustained by a second pulse of small amplitude (the pump pulse), which is not capable of causing breakdown of the gas by itself, but if break-

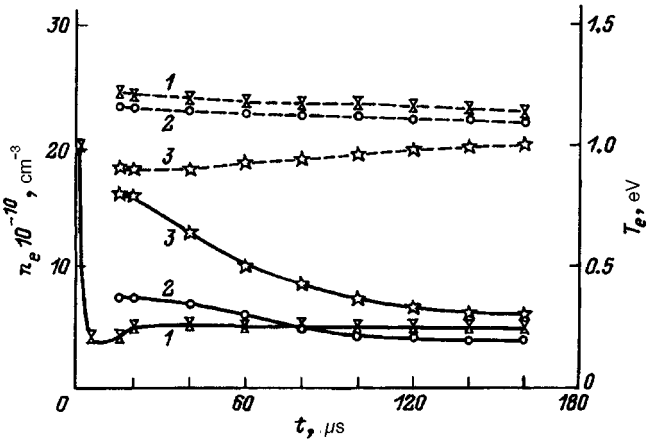

FIG. 1. Time dependence of the electron density (solid curves) and temperature (dashed curves) in the plasma in a programmed-pulse mode, p, Torr: *1* — 3, *2* — 5, *3* — 13.

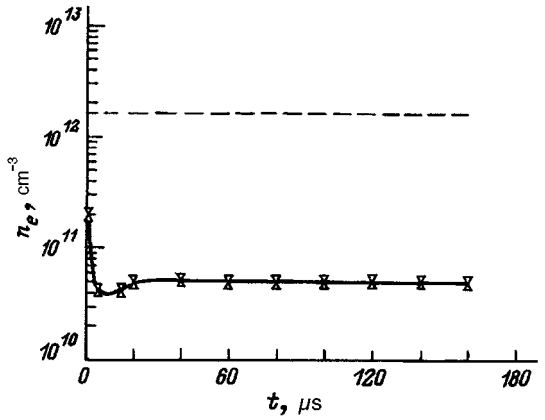

FIG. 2. Time dependence of the electron density in the plasma at a pressure $p=3$ Torr (dashed line — n_{ec2}).

down has already been effected, the microwave power of the pump pulse is sufficient for sustaining the discharge in a fixed position in free space for a long time.

The experiments performed in our laboratory showed that the use of programmed pulses makes it possible, in principle, to solve the problem of localizing a discharge created by a focused microwave beam (stopping it at an assigned position in free space). The experimental setup,[23] on which the plasma parameters of a localized air microwave discharge were studied by optical methods, included two pulsed magnetron generators operating in the centimeter wavelength range at $\lambda_1=10$ cm and $\lambda_2=2.4$ cm. The first (breakdown) generator produced a series of (from one to a hundred) short ($\tau_1=3$ μs) powerful ($W_1 \leqslant 1$ MW) microwave pulses, which appeared with a repetition rate of 400 Hz and effected breakdown of the gas under investigation. The second generator (the pump generator) produced a pulse with a duration $\tau_2=160$ μs and a power $W_2<200$ kW, which immediately followed the last of the series of pulses from the first generator.

The electron density and temperature were investigated at a gas pressure $p<15$ Torr using the method of simultaneously determining n_e and T_e in a nonisothermal nitrogen plasma from measurements of the absolute radiated intensities of bands of the second positive and first negative nitrogen systems.[18,24] At higher pressures the electron temperature was not determined, but the electron density was measured by microwave diagnostics.

The results for various air pressures are presented in Fig. 1. The measured electron density in the tenth pulse (the last pulse under the present experimental conditions) of the series from the breakdown generator amounted to 2×10^{11} cm^{-3} at $p=3$ Torr, i.e., it was close to the critical density $n_{ec1}=10^{11}$ cm^{-3} for the first generator. The electron density measured when breakdown was effected only by a pulse from the second generator (the train of pulses from the first generator was absent) with $W_2=200$ kW amounted to 2×10^{12} cm^{-3} at the beginning of the pulse and to $\sim6\times10^{11}$ cm^{-3} 5 μs after the time of breakdown, i.e., to a value close to the critical density $n_{ec2}=1.6\times10^{12}$ cm^{-3} for the pump generator. In the case of the localized microwave discharge created in a programmed mode (Figs. 1 and 2) the electron density in the plasma during the action of the pump pulse was an order of magnitude lower than the critical value for the second generator. It is also seen from the results obtained that the effective temperature of the electrons in the plasma of the localized air microwave discharge is $T_e\simeq1$ eV and varies only slightly with time and as the pressure is varied, while the electron density decreases with increasing exposure time and reaches a stationary level at the end of the pump pulse.

The low value of the electron density in the second pulse is attributed to the fact that the power of the pulse from the pump generator in the experiment was so small that if breakdown of the gas was not caused by the train of pulses from the first generator, the pulse from the pump generator would not cause breakdown of the gas, i.e., the value of E/p for the second generator was knowingly small (less than the breakdown value). Hence the frequency ν_i of ionization by the field of the second generator was small. Under the conditions of the experiment we obtained $T_e\simeq0.9-1.2$ eV. Such an electron temperature corresponds to ionization frequencies $\nu_i/p=10^1-10^3\text{s}^{-1}\text{Torr}^{-1}$.

When the air pressure is greater than 1 Torr, the main processes leading to the loss of electrons from the discharge are attachment, recombination, and diffusion. Electrons which attach to oxygen molecules either depart from the discharge (negative ions can leave the discharge zone by diffusion or recombine with positive ions) or return to the discharge as a result of detachment. According to estimates, under our conditions the negative-ion density n_- is of the order of the electron density n_e, as is confirmed by the data in Ref. 13, where it was shown experimentally that the negative-ion density increases with increasing air pressure and reaches $n_-\simeq n_e$ at $p\simeq10$ Torr. In this case it follows from the balance equations for the charged particles[25] that the main process maintaining the electron density during the second pulse is electron detachment from negative ions, rather than direct ionization, i.e., the small term $\nu_i n_e$ just compensates the small difference $\nu_a n_e-\nu_d n_-$ between two large quantities (ν_a and ν_d are the frequencies of electron

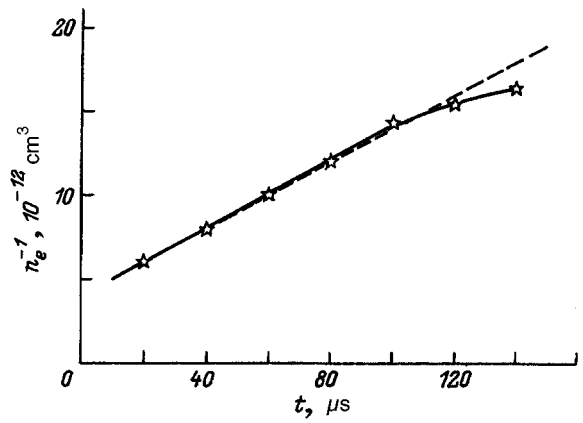

FIG. 3. Time dependence of the value of n_e^{-1} in the plasma at $p=13$ Torr in a programmed microwave pulse mode.

attachment and detachment). Hence it follows that the electron density in the second pulse under the conditions of our experiment cannot be greater than the density created in the plasma by the first generator (which is less than the critical density for the second generator).

To account for the recombination decay of the plasma in a glow discharge in air (an N_2-O_2 mixture), it was theorized in Ref. 26 that the loss of charged particles is determined by the resultant effect of electron attachment and detachment. It was shown in Ref. 26, in which the decay of the plasma of a pulsed air discharge in a constant field was considered, that the decrease in the electron density with time must take place in two stages. In the initial stage of deionization of the plasma, equilibrium is rapidly established between the electron density and the negative-ion density, so that $\nu_a n_e \simeq \nu_d n_-$, and in the second stage slow decay of the plasma as a result of electron–ion and ion–ion recombination takes place with maintenance of that equilibrium. Then, for slow decay of the plasma after the field is removed we have[26]

$$\frac{dn_e}{dt} = -\alpha_{\text{eff}} n_e^2, \tag{1}$$

where $\alpha_{\text{eff}} = \alpha + \beta \nu_a / \nu_d$ is the effective recombination coefficient, which takes into account that apart from the electron–ion recombination channel α there is another channel for the loss of charged particles as a result of ion–ion recombination β.

It is seen from the results of our experiment (Fig. 2) that at an air pressure of 3 Torr the electron density drops rapidly (within a time of the order of several microseconds at the beginning of the pump pulse from a value of 2×10^{11} cm^{-3}, which is characteristic of the plasma created by the radiation from the first generator, to a value of $\sim 5\times10^{10}$ cm^{-3} and subsequently remains constant during the pump pulse, since in this case the electron attachment process is compensated by electron detachment, i.e., $\nu_a n_e = \nu_d n_-$, and ionization compensates diffusion and recombination. As the gas pressure is raised, recombination begins to play an increasingly greater role, and at $p=13$ Torr the decay of n_e (Fig. 3) is described by the effective recombination constant $\alpha_{\text{eff}} \simeq 10^{-7}$ cm^{3}s^{-1}, which is in good agreement, for example, with the data in Ref. 8. It is also seen from Fig. 1 that the electron density achieves a stationary level at the end of the pump generator pulse and does not vary thereafter. This is explained in the following manner. After the electron density has dropped to a value of $\sim 6.5\times10^{10}$ cm^{-3} in our case, the subsequent decrease in n_e due to recombination is completely compensated by bulk ionization, i.e., the balance equation for electrons can be written in the form (since $\nu_a n_e = \nu_d n_-$)

$$\frac{dn_e}{dt} = \nu_{i,\text{eff}} n_e - \alpha_{\text{eff}} n_e^2 = 0, \tag{2}$$

where $\nu_{i,\text{eff}} = \nu_i/(1+\eta)$ is the effective ionization frequency, which takes into account that some of the electrons attaching to oxygen molecules were, in effect, not created at all: $\eta = \nu_a/\nu_d$.

Hence, the frequency of ionization by the field of the second generator under the conditions of our experiment is

$$\nu_{i,\text{eff}} = \alpha_{\text{eff}} n_e. \tag{3}$$

According to the data from the experiment (Fig. 1), when the air pressure equals 13 Torr, we obtain $\nu_i = 7\times10^3$ s^{-1}. Using the formulas in Refs. 27 and 28, in which the functional relations between the ionization frequency and the reduced field E/n were presented, we found the amplitude of the electric field strength in the plasma of a localized air microwave discharge: $E_0 \simeq 240$ V/cm according to the formulas in Ref. 27 and $E_0 \simeq 300$ V/cm according to the formulas in Ref. 28. These values are consistent with the maintenance of the pump generator power at a level that is insufficient for independent breakdown in our experiment.

It follows from (3) that, since the ionization frequency depends on E/n, the electron density n_e in a freely localized microwave discharge created in a programmed-pulse mode is a function of the electric field strength in the plasma

$$n_e = \frac{\nu_{i,\text{eff}}}{\alpha_{\text{eff}}} \sim f(E/n) \sim f(W), \tag{4}$$

i.e., the electron density in the second pulse depends on the level of the pump generator power, as is confirmed by the data from our experiment, which were obtained for different pump pulse powers ($n_e = 6\times10^{11}$ cm^{-3} for $W_2 = 200$ kW and $n_e = 6\times10^{10}$ cm^{-3} for $W_2 = 50$ kW).

The creation of a programmed-pulse mode with the same microwave radiation frequency in the first and second pulses and a study of the threshold characteristics of the repeated microwave discharge in a wave beam led to disclosure of features in the formation of the ionization regions, their dynamics, and their relative positions.[17] The amount of energy imparted to the discharge during the second pulse influences the dynamics of the discharge processes so strongly that it leads to alteration of the final form of the microwave discharge, in particular, in the appearance of a specific type of discharge, viz., a localized microwave discharge.

The possibility of controlling the electron density in the plasma of a nonself-sustained discharge, such as a localized discharge created in a fixed region of free space by a focused microwave beam, is also confirmed by the experiments (Fig.

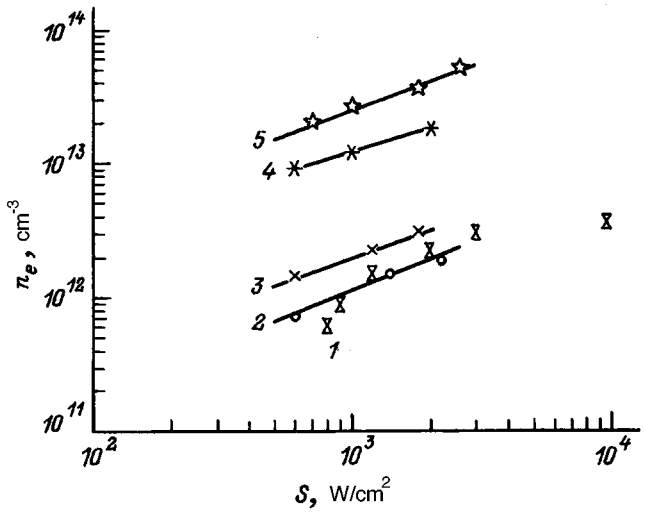

FIG. 4. Dependence of the stationary level of the electron density in the plasma on the flux density of the microwave power supplied during the pump pulse. p, Torr: *1* — 10, *2* — 30, *3* — 150, *4* — 300, *5* — 350.

4) that we performed on the system described in Ref. 29. The parameters of the system allowed us to vary the supplied energy flux density over the course of 800 μs according to a preassigned program. In particular, in our experiments we used a program in which the energy flux density S_1 in the focal region of the beam was at a level sufficient for independent breakdown of the gas during a time $\tau_1 = 1-20$ μs, and the microwave pulse was maintained at a level S_2 that is subthreshold for the unperturbed gas during the remainder of the time. By varying S_2 it was possible to control the geometric dimensions of the plasma of the microwave discharge, the rate at which it propagates toward the focusing antenna, and the gas heating kinetics, particularly the degree of ionization of the plasma (Fig. 4).

Thus, it has been demonstrated experimentally in two systems that the use of a programmed-pulse mode makes it possible to localize a discharge created by a focused microwave beam in an assigned region of free space and that the properties of the nonself-sustained microwave discharge formed at the focus of the beam when electromagnetic radiation with a time-varying energy flux density acts on the gas can be regulated over a broad range without a loss of localization by adjusting the pulse parameters.

[1] *High-Frequency Discharges in Wave Fields* [in Russian], A. G. Litvak (ed.), Inst. Prikl. Fiz. Akad. Nauk, Gor'kiĭ (1988).
[2] *Strong Microwave in Plasmas*, A. G. Litvak (ed.), Inst. Appl. Phys., Acad. Sci. USSR, Nizhny Novgorod (1991).
[3] *Strong Microwave in Plasmas*, A. G. Litvak (ed.), Inst. Appl. Phys., Russ. Acad. Sci., Nizhny Novgorod (1994).
[4] G. M. Batanov, S. I. Gritsinin, I. A. Kossyĭ *et al.*, Tr. FIAN SSSR **160**, 174 (1985).
[5] S. I. Gritsinin, I. A. Kossyĭ, N. M. Tarasova, and V. M. Shibkov, Teplofiz. Vys. Temp. **25**, 625 (1987).
[6] G. A. Askar'yan, G. M. Batanov, A. É. Barkhudarov *et al.*, Fiz. Plazmy **18**, 1198 (1992) [Sov. J. Plasma Phys. **18**, 625 (1992)].
[7] A. L. Vikharev, V. V. Gil'denburg, O. A. Ivanov, and A. N. Stepanov, Fiz. Plazmy **10**, 165 (1984) [Sov. J. Plasma Phys. **10**, 96 (1984)].
[8] A. L. Vikharev, O. A. Ivanov, and A. N. Stepanov, Fiz. Plazmy **10**, 792 (1984) [Sov. J. Plasma Phys. **10**, 460 (1984)].
[9] V. V. Gil'denburg and A. V. Kim, Zh. Éksp. Teor. Fiz. **74**, 141 (1978) [Sov. Phys. JETP **47**, 72 (1978)].
[10] V. V. Gil'denburg and A. V. Kim, Fiz. Plazmy **6**, 904 (1980) [Sov. J. Plasma Phys. **6**, 496 (1980)].
[11] V. G. Borovkin and Yu. F. Kolesnichenko, Pis'ma Zh. Tekh. Fiz. **16**, 55 (1990) [Sov. Tech. Phys. Lett. **16**, 107 (1990)].
[12] V. V. Zlobin, A. A. Kuzovnikov, and V. M. Shibkov, Vestn. Mosk. Univ. Fiz. Astron. **29**(1), 89 (1988).
[13] 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)].
[14] A. M. Devyatov, A. A. Kuzovnikov, V. V. Lodinev, and V. M. Shibkov, Vestn. Mosk. Univ. Fiz. Astron. **32**(2), 29 (1991).
[15] A. F. Aleksandrov, A. A. Kuzovnikov, and V. M. Shibkov, Inzh.-fiz. Zh. **62**, 726 (1992).
[16] A. F. Aleksandrov, A. A. Kuzovnikov, V. M. Shibkov *et al.*, [in Russian] Moscow (1994), No. 4, pp. 20–29.
[17] N. A. Armand, A. S. Zarin, V. E. Mitsuk *et al.*, in *Radiophysics*, MRTI Akad. Nauk SSSR, Moscow (1991), pp. 135–142.
[18] V. N. Shibkov, in *Proceedings of the 20th ICPIG, Contributed Papers*, Pisa, Italy, 1991, Paper 5, pp. 1137–1140.
[19] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **64**, 74 (1994) [Tech. Phys. **39**, 40 (1994)].
[20] Yu. P. Raizer, *Laser-Induced Discharge Phenomena*, Consultants Bureau, New York (1977).
[21] A. V. Gurevich, Usp. Fiz. Nauk **132**, 685 (1980) [Sov. Phys. Usp. **23**, 862 (1980)].
[22] N. D. Borisov, A. V. Gurevich, and G. M. Milikh, *Artificially Ionized Region in the Atmosphere* [in Russian], IZMIRAN, Moscow (1985).
[23] A. A. Kuzovnikov, V. M. Shibkov, and L. V. Shibkova, Zh. Tekh. Fiz. **67**(6), 10 (1997) [Tech. Phys. **42**, 595 (1997)].
[24] M. G. Berdichevskiĭ and V. V. Marusin, Zh. Prikl. Spektrosk. **18**, 1055 (1973).
[25] V. M. Shibkov, K. Sh. Isaev, V. V. Lodinev, and L. V. Shibkova, in *Proceedings of the 11th ESCAMPIG*, St. Petersburg, 1992, pp. 244–245.
[26] A. P. Napartovich, V. G. Naumov, and V. M. Shashkov, Fiz. Plazmy **1**, 821 (1975) [Sov. J. Plasma Phys. **1**, 449 (1975)].
[27] J. T. Mayhan, J. Appl. Phys. **42**, 5362 (1971).
[28] A. V. Gurevich, Geomagn. Aeron. **19**, 633 (1979).
[29] G. I. Batskikh and Yu. I. Khvorostyanoĭ, Radiotekh. Elektron. **37**, 311 (1992).

Translated by P. Shelnitz

Use of perturbing probes for plasma-jet diagnostics

V. I. Batkin and O. Ya. Savchenko

Institute of Chemical Kinetics and Combustion, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia

(Submitted December 4, 1995)
Zh. Tekh. Fiz. **67**, 24–27 (July 1997)

The experience gained in using an array of cylindrical Langmuir probes as a perturbing body for plasma-jet diagnostics is described. The plasma potential on a part of the probe is determined from the ion-scattering minimum at that part. The potential values thus obtained are compared with the results of an analysis of the probe characteristics. The conditions for applicability of the method are discussed. A density marker is created by applying a voltage pulse with a duration of 30 ns to the probes, and time-of-flight probing of the plasma jet is performed. © *1997 American Institute of Physics.* [S1063-7842(97)00507-2]

A plasma jet created by a gas-discharge source is characterized by high directivity of the motion of the ions, and in some cases[1,2] the energy of the directed motion of the ions is 10^3 times greater than the energy characterizing the spread of their transverse velocities. These conditions make it possible to record the transverse perturbation which the ions experience on their route along a plasma jet by observing the motion of ions extracted from the plasma. If a cylindrical Langmuir probe is placed in the plasma jet, the measurement of its current–voltage characteristic can be supplemented by observing the influence of the probe on the ions. The plasma potential can be determined from the ion-perturbation minimum independently of the current–voltage characteristic $I(U)$.[3] In contrast to the method for determining the potential from d^2I/dU^2 (Ref. 4), this method for measuring it utilizes the first derivative of the experimental dependence and is less sensitive to the temporal instability of the plasma, the influence of the magnetic field, and the anisotropy of the plasma. The design of the heatable cylindrical probe is simplified, because there is no need for shielding.[5] In the present report determinations of the potential from the perturbation minimum are compared with the results of measurements using the classical probe method. Static and pulsed probe measurements were performed on the plasma jet of an arc plasma generator.[6,7]

The setup used to perform the measurements is shown in Fig. 1. A hydrogen-plasma arc generator hurls a plasma jet from the hole in anode *1* onto grid diodes *2* and *3*, which form a beam of protons with an energy of 6 keV. There is an apparatus,[8] which makes it possible to investigate the phase volume of the beam using slit collimator *7* and multiple-wire pickup *8*. Cylindrical probes *4*, which are actually 10 gold-plated tungsten filaments with a diameter of 30 μm at a distance of 500 μm from one another, run through the central region of the plasma jet and are at a distance of 70 mm from diode grid *2*. It is possible to move the probes across the plasma jet, and the plasma generator can be moved in the longitudinal direction. The plasma jet is focused by the magnetic field of solenoid *5* together with the electronic current on anode cover *6* (Ref. 9). The design of the ion-source expander and the solenoid were described in Ref. 9.

When the array of probes *4* is located in the central part of the plasma jet and casts a shadow onto diode grid *2*, the potential of the probes influences the angular divergence of the part of the ion beam located in the region of the shadow. At a certain potential the angular spread of the ions has a minimum value and coincides with the spread when the probes are removed. An example of the angular distributions of the ions is given in Fig. 2. Points *1* correspond to the ion-scattering minimum at a probe potential equal to 31 V. Distributions *2* and *3* are obtained, if the potential is shifted 14.5 V in the downward (-45.5 V) and upward (-16.5 V) directions from that value. Here the distance from the anode to the probes is 28 mm, and the current in the arc discharge equals 100 A. The potential corresponding to the ion-scattering minimum is naturally identified with the plasma potential. The plasma potential at the sites of the filaments can be determined by isolating portions of the shadow of the array by the collimator. This technique for measuring the local potential made it possible to make the probes heatable without fabricating miniature shielded leads.[5,6] The conditions for applicability of this method and the mechanism of ion scattering are discussed in the appendix.

Since the perturbation-minimum technique is not associated with a specific manifestation of the perturbation, different perturbation indicators can be used. In Fig. 3 the indicators are the angular spread of the ions (curve *1*), the mean transverse velocity of the ions (curve *2*), and the current in the grid diode (curve *3*). Curves *1–3* are compared with the logarithm of the current–voltage characteristic of the probes (curve *4*). The vertical line at 31 V illustrates the identical nature of results of the determination of the potential from the perturbation minimum and from the current–voltage characteristic of the probes. The measurements were performed in the absence of a magnetic field using a sufficiently homogeneous plasma, so that the current–voltage characteristic of the probes could be associated with the local properties of the plasma. Figures 2 and 3 depict the results of some measurements.

The perturbing probes were used to determine the axial distribution of the potential in a plasma jet with and without focusing of the latter by a magnetic field. The data are presented in Fig. 4. The distance Z is measured from the anode. Points *1* were obtained in the absence of a magnetic field and

 1063-7842/97/070737-04$10.00

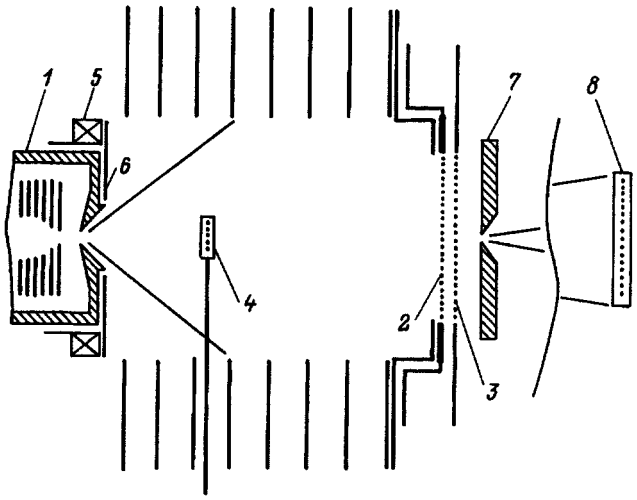

FIG. 1. Experimental setup.

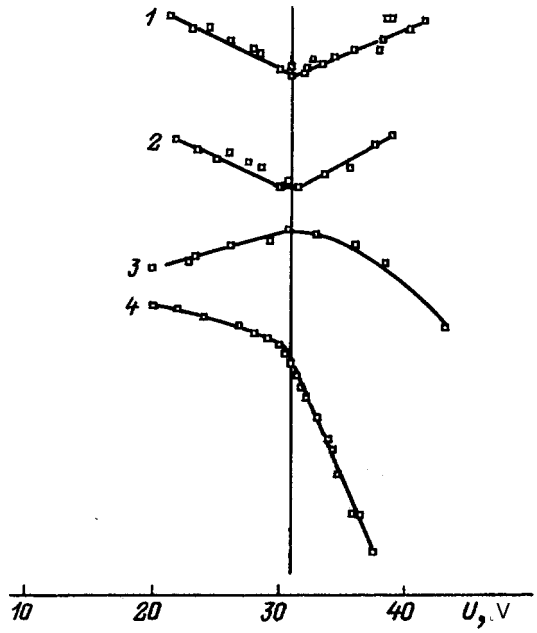

FIG. 3. Indicators of the perturbation minimum.

an electronic current in the expander walls. Points *2* correspond to the case in which a longitudinal magnetic field equal to 41 Oe was created at a distance of 10 mm from the anode. In this case no focusing of the plasma jet was observed, the increase in the density of the plasma was caused by the accumulation of secondary slow ions. When an electron current equal to 0.7 A was created in electrode *6* (Fig. 1) in addition to the magnetic field, focusing of the plasma jet occurred in the near-anode region and was manifested in the longitudinal potential profile (points *3*). Where possible (cases *1* and *2*), the determinations of the potential from the perturbation minimum were supplemented by determining it from the current–voltage characteristic of the probes (points *4* and *5*, respectively). The data presented in Fig. 4 illustrate the good agreement between the two methods for determining the potential with considerable variation of the plasma density.

The use of perturbing probes permits the determination not only of the potential of the plasma, but also of its kinematic characteristics. For this purpose a pulsed negative voltage must be supplied to the probes, and modulation of the plasma density is caused by the resultant variation of the capture of ions by the probes. The current response of the ions extracted from the plasma by the diode can be used to determine the mean velocity of the center of the density marker and the rate at which it spreads out in the gap l between the probes and the diode. The former gives the mass velocity of the ions V_i, and the latter gives the velocity of ion sound c, which characterizes the electron temperature $c^2 = T_e/m_i$ (Ref. 10). The typical oscillogram of the ion current in Fig. 5 illustrates the wave character of the spreading of the density marker. The temporal characteristics of the current pulse do not depend on the amplitude of the analyzing pulse, and thus the spreading process is linear. The duration of the voltage pulse at the 0.5 level is 30 ns. If the delay between the beginning of the leading edge of the current pulse and the beginning of the leading edge of the voltage pulse is T_1 and the time of the end of the trailing edge is T_2, we have

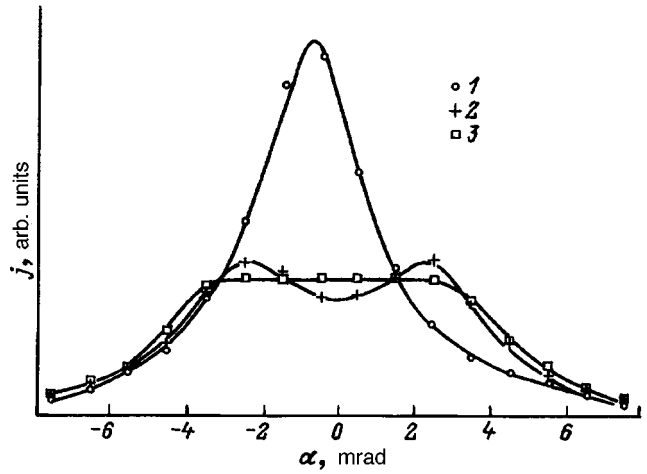

FIG. 2. Angular distribution of the ions.

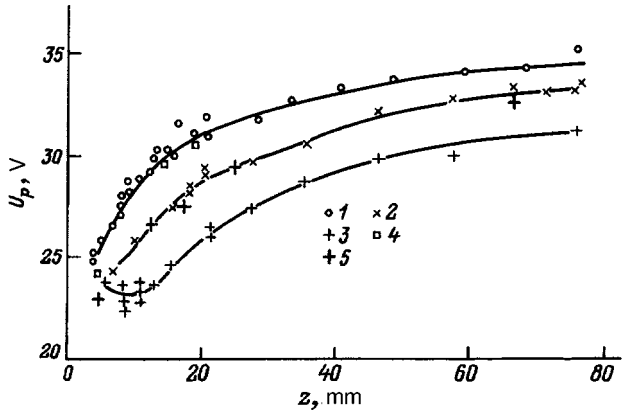

FIG. 4. Axial distribution of the potential in a plasma jet.

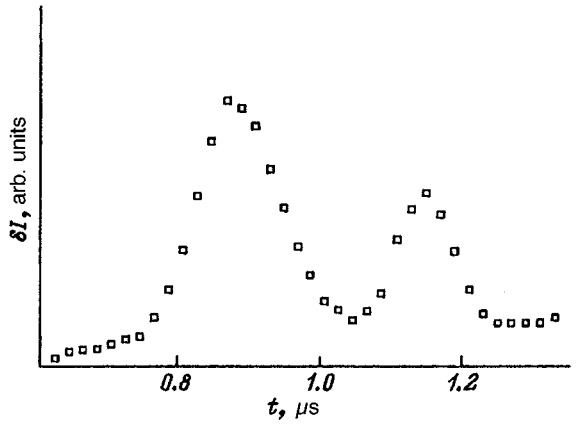

FIG. 5. Current response of ions extracted by the diode to a voltage pulse on the probes. The time is measured from the beginning of the voltage pulse, and $I_e=0$. The oscillogram is inverted.

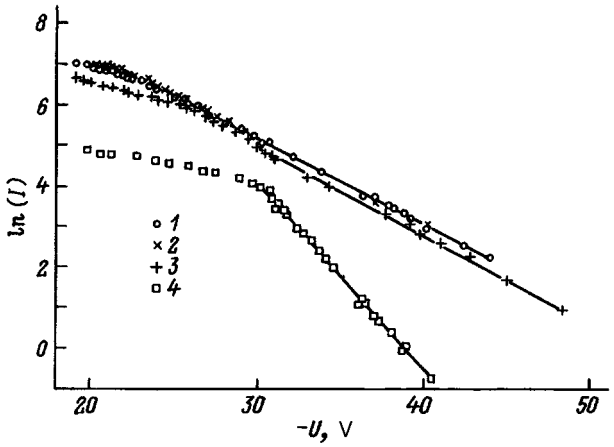

FIG. 6. Electronic branch of the probe characteristics. Z, mm: *1* — 8, *2* — 4.5, *3* — 14, *4* — 28; *1*, *3* — without a magnetic field; *2*, *4* — with a magnetic field.

$$V_i=(l/T_1+l/T_2)/2; \quad c=(l/T_1-l/T_2)/2. \tag{1}$$

The presence of slow ions in the plasma jet causes two additional ion-sound waves having a velocity $\approx\pm c$ instead of $\approx V_i\pm c$ to be excited. For $V_i>2c$ the use of formulas (1) in the presence of slow ions with the same density as the fast ions does not lead to appreciable distortion of V_i and underestimates T_e by $<20\%$. The velocity of the ions when $Z>1.5$ cm corresponds to a kinetic energy equal to 28 eV. The mean electron temperature T_e calculated from c depends on the distance Z between the probes and the anode. When Z is increased from 4 to 25 mm, it drops from 3 to 2.4 eV, and when $Z>25$ mm, it drops to 2 eV. The duration of the pulse edges δI is significantly greater than the duration of the voltage pulse on the probes. This difference is naturally attributed to Landau damping of the short-wavelength components of the spectrum of δI. When $T_e=2$ eV and $\varepsilon_i=30$ eV, the damping of the sound waves with a ratio of the frequency to the damping coefficient equal to 4–5 takes place at a longitudinal temperature of the ions ≈ 0.5 eV, which is close to the results of direct measurements of the spread of the longitudinal velocities of ions.[2]

The estimates of T_e obtained from measurements of the wave velocity can be checked against the probe characteristics. The electronic components of the probe characteristics are depicted in Fig. 6. The ion current I_i, which was subtracted from the probe current, was approximated by the formula

$$I_i=\alpha\sqrt{|U|}+\beta\sqrt{U_p-U},$$

where U and U_p are, respectively, the probe potential and the plasma potential.

The first term is associated with ions moving away from the anode orifice, and the second term is associated with the ions formed within the plasma jet as a result of ionization and charge exchange. The ratio α/β is close to 3/2. Curves *1* and *2* were obtained at a distance of 4.5 mm from the anode and are distinguished by the absence and presence of a magnetic field, respectively. Curve *3* was plotted for $Z=14$ mm in a magnetic field, and curve *4* was plotted for $Z=28$ mm without such a field. Characteristics *1* and *2* and, to some extent, characteristics *3* correspond to the Boltzmann distribution of electrons with $T_e=4.5$ eV. As the distance from the near-anode region increases, a low-temperature electronic component appears in the form of a step on probe characteristic *3* and then becomes dominant at large distances. The temperature of this component is 2 eV. The ''hot'' electrons are supplied by the anode orifice, and the ''cold'' electrons are secondary particles, which accumulate in the potential trap created by the plasma jet. The predominance of the primary electrons having a small angular momentum supports the dynamic model of the focusing of a plasma jet in the near-anode region,[6] which attributes the appearance of the electric field that focuses the ions to the need to compensate the Lorentz force acting on the electrons in a longitudinal magnetic field. The data on the energy distribution of the electrons obtained from the probe characteristics are consistent with the results of the estimates of T_e from the velocity of the ion sound. The Boltzmann distribution, the values of T_e measured here, and the picture of the geometric dispersal of the ions provide an estimate of the longitudinal drop in the plasma potential that is close to the data in Fig. 4. At the same time, the ambipolar mechanism for accelerating the ions does not account for their total kinetic energy, which reaches 30–50 eV.[1,2]

CONCLUSIONS

The method of determining the plasma potential from the minimum of the scattering of ions by an array of cylindrical probes can be recommended for use in cases in which the plasma flow has a sufficient degree of directivity and its density is not excessively great. An array of probes can also be employed in time-of-flight plasma probing, which provides quick information on the velocity of the ions and the electron temperature.

APPENDIX SCATTERING OF IONS BY AN ARRAY OF CYLINDRICAL PROBES

If the plasma is not excessively rarefied, a potential well of width r is filled by electrons during the characteristic time

$t > (\lambda/V_e)(D/r)^2$, where λ is the electron mean free path and D is the Debye radius. Under these conditions, for the vicinity of the probes up to a distance $\approx D$, as well as a probe radius $R \ll D$, the field of the probes is shielded at a characteristic distance $\approx D$. When the distance between the probes $H \gg D$, the mutual influence of the probes can be neglected. In this case the electric field around each of them is assigned in a linear approximation by a modified Bessel function: $E \sim K_1(x/D)$.

A calculation of the transverse velocity $v_\perp$, which an ion acquires as it passes at a distance x from a probe under the condition $v_\perp/v_\parallel \ll 1$, gives

$$v_\perp/c = UF(x/D).$$

Here $U = e\varphi/[f(R/D)\sqrt{\varepsilon_i T_e}]$; φ is the potential of the probe relative to the plasma; $c = \sqrt{T_e/m_i}$ is the velocity of the ion sound; ε_i and T_e are the kinetic energy of the ions and the electron temperature in energy units;

$$F(x) = x\sqrt{2}\int_0^\infty dy K_1(\sqrt{x^2+y^2})/\sqrt{x^2+y^2}$$

$$\approx 2.2\ \exp(-x);$$

$$f(x) = \int_x^\infty K_1(t)dt; \quad \text{when } R \ll D, \quad f \approx 2.$$

Two principal factors displace the potential of the ion scattering minimum: the ion shadow and the absorption of electrons by the probe. When $R \ll D$ and the transverse temperature of the ions T_i is small, the former effect predominates. If the shadow width h is determined from the trajectory of an ion near a probe, then $h = 2r\cos(\pi Z/2L_1)$, and the shadow length $L_1 = \pi D\sqrt{\varepsilon_i/T_e}$. An estimate of the shadow length from the angular spread of the ions gives $L_2 \approx R\sqrt{T_i/\varepsilon_i}$. Combining these estimates, we obtain $L \approx 1/\sqrt{L_1^{-2} + L_2^{-2}}$. When $R \ll D$ and $T_i \ll \varepsilon_i$, we have $h \ll D \ll L$, so that the shadow attracts ions like a flat plate charged by electrons. With consideration of the shadow

$$v_\perp/c = UF(x/D) + \sqrt{2}(R/D)\exp(-x/D).$$

When $U < 0.4$, the probes excite ion-sound waves in the transverse direction, which undergo rapid Landau damping and increase the transverse temperature of the ions after the probes:

$$\Delta T_i \approx 4.8 T_e (D/H)(U - U_0)^2,$$

where $U_0 \approx 0.64(R/D)[1 + 4(D/R)^2(T_i/T_e)]^{0.5}$.

When $U \gg 0.4$, many-flow dispersal of the ions occurs, and it can be stated that the ions scattered by the probes undergo drift, which is weakly perturbed by the bulk field.

The data in Fig. 2 were obtained for $T_e \approx 2$ eV, $T_i \approx 0.1$ eV, $D \approx 0.1$ mm, and $\varepsilon_i \approx 30$ eV when $R/D \approx 0.15$ and the shadow effect is insignificant. The two-peaked distributions *2* and *3* correspond to $U \approx \pm 1$. When $T_i = 0$, a single probe disperses the ions with an angular distribution $\sim 1/\alpha$. The finite spacing H of the probes in the array causes the peak of the distribution to be cut off at the scattering angle of the ions corresponding to a distance $\approx H/2$ from a probe. Under these conditions $H/2 = 2.5D$. The array of probes has finite dimensions with an angle of vision corresponding to $v_\perp/c = 0.2$. Ions with small scattering angles come from the center of the array with a ''cut-off'' velocity distribution, and the ions with large angles come from the edge of the array with a $1/\alpha$ distribution. This effect is enhanced by the ambipolar diffusion of the plasma toward the shadow of the array.

[1] V. I. Batkin, V. N. Getmanov, O. Ya. Savchenko *et al.*, Prikl. Mekh. Tekh. Fiz. (6), 30 (1982).
[2] V. I. Batkin, Zh. Tekh. Fiz. **66** (3), 37 (1996) [Tech. Phys. **41**, 250 (1996)].
[3] V. I. Batkin, *All-Union Seminar on Charged-Particle Linear Accelerators. Abstracts* [in Russian], Khar'kov, 1989, p. 79.
[4] B. V. Alekseev and V. A. Kotel'nikov, *Probe Method of Plasma Diagnostics* [in Russian], Énergoatomizdat, Moscow (1988).
[5] E. V. Shun'ko, Preprint No. 136, Institute of Nuclear Physics, Academy of Sciences of the USSR, Siberian Branch, Novosibirsk (1983).
[6] G. I. Dimov, Yu. G. Kononenko, O. Ya. Savchenko *et al.*, Zh. Tekh. Fiz. **38**, 997 (1968) [Sov. Phys. Tech. Phys. **13**, 754 (1969)].
[7] G. I. Dimov, G. V. Roslyakov, and O. Ya. Savchenko, Preprint No. 145, Institute of Nuclear Physics, Academy of Sciences of the USSR, Siberian Branch, Novosibirsk (1967).
[8] V. I. Batkin and V. N. Getmanov, Prib. Tekh. Éksp. (5), 72 (1983).
[9] V. I. Batkin, Prikl. Mekh. Tekh. Fiz. (5), 11 (1991).
[10] L. A. Artsimovich and R. Z. Sagdeev, *Plasma Physics* [in Russian], Atomizdat, Moscow (1978).

Translated by P. Shelnitz

Instability of Abrikosov vortices in a type-II superconductor–ferrite structure with a longitudinal electric field

Yu. I. Bespyatykh and V. D. Kharitonov

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 141120 Fryazino, Russia

W. Wasilewski

Higher Technical School, Radom, Poland
(Submitted February 7, 1996)
Zh. Tekh. Fiz. **67**, 28–34 (July 1997)

The influence of the interaction of the Abrikosov vortices with the magnetization on the longitudinal vortex instability in a layered type-II superconductor–ferrite structure is analyzed. It is shown that in the vicinity of the orientational phase transition in the magnet, where the transverse susceptibility of the magnet is high, the longitudinal critical current in the structure can be almost 1.5 times smaller than the corresponding value in the isolated superconductor. Because of the influence of the nonlocality of the interaction between the vortices, such an effect can be observed only in structures with superconductors that have weak or moderate pinning. A structure is considered in which the thickness of the superconductor is significantly greater than the London magnetic-field penetration depth and the wavelength of the critical mode. © *1997 American Institute of Physics.* [S1063-7842(97)00607-7]

1. Most ferrites are semiconductors at room temperature and can be regarded as insulators at liquid-nitrogen and lower temperatures. Because of the small penetration depth of the conduction electrons of a superconductor into a ferrite, the exchange interaction between such layers apparently does not play a major role. In type-II superconductor–ferrite structures with an electromagnetic interaction between the Cooper pairs and the electron spins in the ferrite, the expulsion of magnetic flux from the superconductor leads to displacement of the transition point of the magnet from a homogeneous magnetic state to an inhomogeneous state toward lower magnetizing fields,[1,2] and in the case of sufficiently thin magnetic films it can render the domain structure energetically unfavorable.[2] The interaction of the magnetization with the vortices can also alter the type of phase transition to the inhomogeneous state and the orientation of the domain boundaries with respect to the magnetizing field[3] (from the qualitative standpoint the role of the interaction of the magnetization with the vortices is similar to the role of magnetostriction in an isolated ferromagnet[4]). Finally, the magnetic field of the domains in a bulk magnet can create weak links in a thin superconducting film adjoining it.[5]

Studies of the dynamic properties of superconductor–ferrite structures have focused predominantly on the analysis of the excitation and propagation of spin waves in the microwave range. The dispersion and damping of surface and bulk magnetostatic waves in such structures,[6–8] as well as the emf of the entrainment of electrons in the superconductor by magnetostatic waves,[9] have been measured and estimated. The efficiency of exciting spin waves by scattering an electromagnetic field on a lattice of Abrikosov vortices has been analyzed theoretically,[10,11] and the possibility of amplifying magnetostatic waves by interacting them with vortices drifting under the action of a transport current,[12] as well as by creating a negative differential conductivity in the superconductor,[13] has been evaluated.

As far as we know, the influence of the magnitude of the transport current on the instability of the ground state of type-II superconductor–ferrite structures toward low-frequency excitations has not been investigated heretofore. The spectrum and damping of the high-frequency excitations in such structures in the presence of a transport current have also been inadequately studied, although experimental investigations have been carried out in that area. In the present work we investigated the influence of the electromagnetic interaction of the vortices with the magnetization in a type-II superconductor–ferrite structure on the transition of the superconductor from the nonresistive state to the resistive state under the action of a transport current.

The investigation of the transition of a type-II superconductor to the resistive state under the influence of a transport current is of great importance for practical applications of such materials, and the literature devoted to it is quite extensive (see, for example, Refs. 14 and 15 and the literature cited therein). As a whole, the behavior of a system of vortices in the case where the transport current reaches or exceeds the critical current for the transition of a type-II superconductor from the nonresistive state to the resistive state is complicated and has not been solved hitherto. The case of longitudinal vortex instability, in which the transport current is parallel to the direction of the vortices, has been studied most thoroughly by theoretical methods.

The longitudinal critical current in type-II superconductors having the shape of long thin cylinders was estimated with allowance for the external stray field in Ref. 16. The critical longitudinal current in thick superconducting layers was calculated in Ref. 17 taking into account the influence of the stray field and of the bulk and surface pinning. The Labusch model[15] was used to describe the pinning. As was shown in Refs. 16 and 17, the stray field energy is positive and always leads to an increase in the critical current.

The main purpose of the present work is to ascertain to

what extent the longitudinal critical current in a type-II superconductor–ferrite structure can be varied by varying the amplitude of the magnetizing field and the magnetic parameters of the ferrite. In addition, it would be interesting to compare the role of the magnetic field in the ferrite with the role of the electric field in an insulator with a large dielectric constant in a layered semiconductor–insulator structure when the Gunn electric domain instability is suppressed.[18] In fact, the static magnetic susceptibility of ferrites in the domain phase[19] and in the homogeneous phase can be considerable near the field corresponding to the transition from the homogeneous state to the inhomogeneous state,[20] and, at first glance, the drawing of magnetic flux into the magnet and, accordingly, extension of the vortex instability can be expected. However, as will be shown below, in the case under consideration here the magnet has an influence on the longitudinal vortex instability that is the exact opposite of the influence of an insulator with a large dielectric constant on the Gunn domain instability in semiconductors. As the permeability of the ferromagnet increases, both the magnetic field energy of the vortices outside the superconductor and the threshold for longitudinal vortex instability decrease. The value of the critical longitudinal current depends in a complicated manner on the vortex pinning in the superconductor, as well as on the magnetizing field, the magnetization, and the magnetic anisotropy of the ferrite. The nonlocality of the interaction of the vortices with one another has an appreciable influence on the parameters of the critical mode at any values of the pinning constants in the superconductor.

2. Let us assume that the superconductor–ferromagnet structure consists of a superconducting half-space $y>0$ and a ferromagnetic layer $-L<y<0$. The magnet has ''easy-axis'' magnetic anisotropy. The direction of the anisotropy axis $\mathbf{n_a}\|\mathbf{n_y}$. The structure is immersed in a tangential magnetizing field $\mathbf{H}_e=H_e\mathbf{n}_z$, $H_e>H_a$ ($H_a=\beta M_0$ is the anisotropy field, $\beta>0$ is the anisotropy constant, and M_0 is the saturation magnetization of the ferromagnet), and $H_{c1}\ll H_e\ll H_{c2}$ (H_{c1} and H_{c2} are the lower and upper critical fields of the superconductor). The last condition makes it possible to use the London approximation to describe the superconductor. A small transport current **I** flows near the surface of the superconductor and is parallel to the magnetizing field ($\mathbf{I}\|\mathbf{n}_z$). The superconductor is in the mixed state, the distance between neighboring vortices is much smaller than the London penetration depth of the magnetic field in the superconductor λ, and the magnetization **M** in the ferromagnet is uniform ($\mathbf{M}\|\mathbf{n}_z$).

The energy of the system U including the interaction of the vortex and magnetic subsystems can be represented in the form

$$U=U_S+U_M+U_{\text{int}}, \tag{1}$$

where U_S is the energy of the vortices in the isolated superconductor, U_M is the magnetization energy with allowance for the Meissner effect and the additional external field in the magnet induced by the transport current, and U_{int} is the energy of the interaction of the vortices with the magnetization.

The energy of a vortex lattice in the absence of a transport current was calculated in Ref. 21, and the energy of the interaction of vortices with a transport current was obtained in Ref. 22. The energy of interaction of magnetization with the vortices and the magnetic energy with the Meissner effect taken into account were found in Ref. 7. Here, however, we must take into account the renormalization of the magnetizing field in the ferromagnet due to the field induced by the transport current. As was established in Ref. 17, a long-wavelength mode is stable in superconductors with weak or moderate pinning; therefore, the continuum model of a system of vortices can be used to calculate the critical transport current. Here we shall use the same approximation, but, in contrast to Ref. 17, we shall neglect the nonlocal character of the interaction between the vortices. To find the longitudinal critical current I_c at which the system of vortices becomes unstable, it is sufficient to examine small static perturbations in the structure. In this case we need only the part of the energy of the system which is quadratic with respect to the small displacements **u** of the vortices and the small deviations of the magnetization **m** from the ground-state values.

The energy U_S of the vortices in an isolated superconductor with a longitudinal transport current equals[17]

$$U_S=U_v+U_{\text{source}}+U_{\text{stray}}+U_{\text{pin}}+U_{\text{bind}}+U_I, \tag{2}$$

where U_v is the energy of the interaction of vortices with other vortices and with image vortices

$$\begin{aligned}U_v=\frac{1}{4\lambda^2}C_{11}\int_0^\infty dy\int\frac{d\mathbf{k}}{4\pi^2}\Bigg\{\int_0^\infty dy'\,\frac{1}{\tau}\{[k^2u^x_{\mathbf{k}}u^{x'}_{-\mathbf{k}}\\ +(k_z^2-\tau^2)u^y_{\mathbf{k}}u^{y'}_{-\mathbf{k}}]e^{-\tau|y-y'|}-[k^2u^x_{\mathbf{k}}u^{x'}_{-\mathbf{k}}\\ +(\tau^2-k_z^2)u^y_{\mathbf{k}}u^{y'}_{-\mathbf{k}}]e^{-\tau(y+y')}\}+2u^y_{\mathbf{k}}u^y_{-\mathbf{k}}(1-e^{-y/\lambda})\\ +2\varepsilon\lambda^2\left[k_x^2u^y_{\mathbf{k}}u^y_{-\mathbf{k}}+\frac{du^x_{\mathbf{k}}}{dy}\frac{du^x_{-\mathbf{k}}}{dy}\right]\Bigg\},\end{aligned} \tag{3}$$

U_{source} is the energy of the interaction of the vortices with the external field $\mathbf{H}_e$

$$U_{\text{source}}=\frac{C_{11}H_e}{2\lambda^2B}\int_0^\infty dy\int\frac{d\mathbf{k}}{4\pi^2}u^y_{\mathbf{k}}u^y_{-\mathbf{k}}e^{-y/\lambda}, \tag{4}$$

U_{stray} is the stray field energy

$$\begin{aligned}U_{\text{stray}}=\frac{1}{2\lambda^2}C_{11}\int_0^\infty dy\,dy'\\ \times\int\frac{d\mathbf{k}}{4\pi^2}k_z^2\left(\frac{1}{k}-\frac{1}{\tau}\right)u^y_{\mathbf{k}}u^{y'}_{-\mathbf{k}}e^{-\tau(y+y')},\end{aligned} \tag{5}$$

U_{pin} is the bulk pinning energy

$$U_{\text{pin}}=\frac{1}{2}C_{11}k_p^2\int_0^\infty dy\int\frac{d\mathbf{k}}{4\pi^2}\mathbf{u}_{\mathbf{k}}u_{-\mathbf{k}}, \tag{6}$$

U_{bind} is the surface pinning energy

$$\begin{aligned}U_{\text{bind}}=\frac{1}{2}C_{11}\int\frac{d\mathbf{k}}{4\pi^2}[k_1u^x_{\mathbf{k}}(0)u^x_{-\mathbf{k}}(0)\\ +k_2u^y_{\mathbf{k}}(0)+u^y_{-\mathbf{k}}(0)],\end{aligned} \tag{7}$$

and U_I is the energy of the interaction of the vortices with the transport current

$$U_I = -iC_{11}t\int_0^\infty dy \int \frac{d\mathbf{k}}{4\pi^2} k_z \hat{j} u_{\mathbf{k}}^x u_{-\mathbf{k}}^y. \tag{8}$$

Here B is the magnetic induction in the superconductor; $\varepsilon = C_{11}/C_{66}$, where C_{11} and C_{66} are the bulk and shear moduli of the vortex lattice in the local limit; $\mathbf{k}=(k_x,0,k_z)$; $k_p^2 = \alpha_L/C_{11}$; α_L is the Labusch constant; k_1 and k_2 are the phenomenological surface pinning constants; $\tau = (\lambda^{-2}+k^2)^{1/2}$; $t = H_I/H_e$; H_I is the field induced by the transport current on the surface of the superconductor

$$H_I = \frac{4\pi}{c} I,$$

where I is the transport current through a unit of width of the superconductor and c is the velocity of light in a vacuum; $\mathbf{u}_{\mathbf{k}} \equiv \mathbf{u}_{\mathbf{k}}(y)$; and $\mathbf{u}'_{\mathbf{k}} \equiv \mathbf{u}_{\mathbf{k}}(y')$, where $\mathbf{u}_{\mathbf{k}}$ is the Fourier transform of the displacement of the vortices. In the range of magnetizing fields under consideration the induction B in the superconductor equals $B \cong H_e - (\Phi_0/8\pi\lambda^2) \times [-3.872 + \ln(\Phi_0/\lambda^2 H_e)] \cong H_e$ (Ref. 23). The stiffness ratio $\varepsilon \cong H_{c2}/8B\varkappa^2$ in the London limit $1/2\varkappa^2 \leqslant B/H_{c2} \leqslant 0.3$ for superconductors with a large value for the Ginzburg-Landau parameter $\varkappa \gg 1$ is small.[22] The spatial distribution of the transport current in a superconductor was calculated in Refs. 24 and 25. It was shown that the dependence $\hat{j}(y)$ has the form

$$\hat{j}(y) = \gamma e^{-\gamma y}, \tag{9}$$

where $\gamma \cong \lambda^{-1}(1 - B/H_{c2}) \cong \lambda^{-1}$.

Taking into account the arguments in Ref. 22, we shall henceforth regard γ as an assigned parameter and use the value $\gamma = \lambda^{-1}$ for estimates. The magnetic energy of the ferromagnet U_M equals

$$U_M = U_Z + U_d + U_a + U_{ex}, \tag{10}$$

where U_Z is Zeeman energy of the ferromagnet in the external field $\mathbf{H}_0 = \mathbf{H}_e + \mathbf{H}_I$

$$U_Z \cong 2\pi\Omega_0 \int \frac{d\mathbf{k}}{4\pi^2} \int_{-L}^0 dy (m_{\mathbf{k}}^\xi m_{-\mathbf{k}}^\xi + m_{\mathbf{k}}^y m_{-\mathbf{k}}^y), \tag{11}$$

U_d is the dipolar energy

$$\begin{aligned} U_d = 2\pi \int \frac{d\mathbf{k}}{4\pi^2} \int_{-L}^0 dy \Bigg\{ & m_{\mathbf{k}}^y m_{-\mathbf{k}}^y + \frac{1}{2k} \\ & \times \int_{-L}^0 dy' \Bigg\{ [k_\xi^{y'} m_{\mathbf{k}}^{\xi'} m_{-\mathbf{k}}^\xi - k^2 m_{\mathbf{k}}^{y'} m_{-\mathbf{k}}^y \\ & + 2ik_\xi k m_{\mathbf{k}}^{\xi'} m_{-\mathbf{k}}^y \operatorname{sgn}(y-y')] e^{-k|y-y'|} \\ & + \frac{(\tau-k)}{(\tau+k)} (k_\xi^2 m_{\mathbf{k}}^{\xi'} m_{-\mathbf{k}}^\xi + k^2 m_{\mathbf{k}}^{y'} m_{-\mathbf{k}}^y \\ & - 2ik_\xi k m_{\mathbf{k}}^{\xi'} m_{-\mathbf{k}}^y) e^{-k(y+y')} \Bigg\} \Bigg\}, \end{aligned} \tag{12}$$

U_a is the magnetic anisotropy energy

$$U_a = 2\pi Q \int \frac{d\mathbf{k}}{4\pi^2} \int_{-L}^0 dy m_{\mathbf{k}}^y m_{-\mathbf{k}}^y, \tag{13}$$

and U_{ex} is the exchange energy

$$\begin{aligned} U_{\text{ex}} = 2\pi D \int \frac{d\mathbf{k}}{4\pi^2} \int_{-L}^0 dy \Bigg[& k^2 (m_{\mathbf{k}}^\xi m_{-\mathbf{k}}^\xi + m_{\mathbf{k}}^y m_{-\mathbf{k}}^y) \\ & + \frac{dm_{\mathbf{k}}^\xi}{dy} \frac{dm_{-\mathbf{k}}^\xi}{dy} + \frac{dm_{\mathbf{k}}^y}{dy} \frac{m_{-\mathbf{k}}^y}{dy} \Bigg]. \end{aligned} \tag{14}$$

Here $\Omega_0 = H_0/4\pi M_0$, Q is the Q factor, $D = \alpha/4\pi$ is the inhomogeneous exchange constant of the ferromagnet, ξ is the axis in the Cartesian coordinate system $\{\xi, y, \zeta\}$ associated with the total external field $(\mathbf{n}_\zeta \| \mathbf{H}_0)$, $k_\xi = k_x \cos\theta - k_z \sin\theta$, $\theta = \arctan(H_I/H_e)$, $\mathbf{m}_{\mathbf{k}} \equiv \mathbf{m}_{\mathbf{k}}(y)$, $\mathbf{m}'_{\mathbf{k}} \equiv \mathbf{m}_{\mathbf{k}}(y')$, and $\mathbf{m}_{\mathbf{k}}$ is the Fourier transform of the deviations of the magnetization from the ground state. The energy of interaction of the vortices with the magnetization has the form

$$U_{\text{int}} = \frac{B}{\lambda^2} \int \frac{d\mathbf{k}}{4\pi^2} \frac{k_z}{\tau+k} \int_{-L}^0 dy dy' \left(\frac{k_\xi}{k} m_{-\mathbf{k}}^{\xi'} - i m_{-\mathbf{k}}^{y'} \right) u_{\mathbf{k}}^y e^{ky' - \tau y}. \tag{15}$$

Using the equation of state for the magnetization

$$\delta U / \delta \mathbf{m}_{\mathbf{k}} = 0 \tag{16}$$

and the relation

$$\int d\mathbf{k} \int_{-L}^0 dy \mathbf{m}_{\mathbf{k}} \delta U / \delta \mathbf{m}_{\mathbf{k}} = 2U_M + U_{\text{int}}, \tag{17}$$

we find the expression for the sum of the magnetic energy U_M and the energy of interaction U_{int} of the vortices with the magnetization:

$$U_M + U_{\text{int}} = U_{\text{int}}/2, \tag{18}$$

which simplifies the calculation of the total energy of the system. Since the magnetic energy in fields exceeding the field corresponding to the transition to the domain phase is positive ($U_M > 0$), the conditions

$$U_{\text{int}} < 0, \quad U_M + U_{\text{int}} < 0 \tag{19}$$

follow from (18). Inequalities (19) allow us to conclude that the interaction of the vortices with the magnetization reduces to a decrease in the influence of the stray field. Such a conclusion is natural, since the appearance of additional degrees of freedom as a result of such displacements of the vortices should lower the total energy of the system. Solving the equation of state for the magnetization (16), we can find the magnetization as a function of the displacement of the vortices and eliminate the explicit dependence of the energy of the system on it. As a result, the energy of the system U will be described by Eqs. (2)–(9), if the stray field energy is preliminarily renormalized. Let the thickness of the ferro-

magnetic layer be sufficiently large ($L \gg D^{1/2}$). Then, expressing the magnetization in terms of the displacement of the vortices and substituting it into (15), we obtain

$$U_{\text{int}} = -\frac{C_{11}}{\lambda^2}\int \frac{d\mathbf{k}}{4\pi^2}\int_0^\infty dy\,dy'\,k_z^2 \times\left(\frac{1}{k}-\frac{1}{\tau}\right)F(\mathbf{k})u_{\mathbf{k}}^{y}u_{-\mathbf{k}}^{y'}e^{-\tau(y+y')}, \tag{20}$$

where

$$F(\mathbf{k}) = \frac{[(k_\xi^2+k^2\tilde{\Omega}_0)^2-k^2q^2\tilde{\Omega}_0^2]\sinh qL}{(k_\xi^2-k^2\tilde{\Omega}_0)[(k_\xi^2+k^2\tilde{\Omega}_0)\sinh qL + kq\tilde{\Omega}_0\cosh qL]},$$

$$q = \frac{\tilde{\Omega}_0-Q}{\tilde{\Omega}_0(\tilde{\Omega}_0+1-Q)}(k_\xi^2+k^2\tilde{\Omega}_0), \quad \tilde{\Omega}_0 = \Omega_0 + Dk^2. \tag{21}$$

The quantity $F(\mathbf{k})$ also contains a dependence on parameters of the ferromagnet. The renormalization of the stray field energy reduces to the additional multiplier $P(k)=[1-F(\mathbf{k})]$ in the integrand in Eq. (5). Generally speaking, the influence of the magnetization on the critical current increases with increasing thickness of the ferromagnet L; therefore, we confine ourselves to estimates for a structure with a semi-infinite ferromagnet ($L\to\infty$). Then, from (21) we find

$$P(\mathbf{k}) = \frac{k}{\sqrt{k_\xi^2+k^2\tilde{\Omega}_0}}\sqrt{\frac{\tilde{\Omega}_0(\tilde{\Omega}_0-Q)}{(\tilde{Q}_0+1-Q)}}. \tag{22}$$

The inhomogeneous exchange constant for ferrites is small ($D\sim 10^{-11}-10^{-12}$ cm^2), so that under conditions for which the continuum approximation is applicable the inequality $Dk^2 \ll 1$ definitely holds in the range of external fields $H_e \sim 10^3-10^4$ Oe. Therefore, we shall henceforth set $\tilde{\Omega}_0=\Omega_0$.

It is easy to see that $0<P(\mathbf{k})<1$. When $\Omega_0\to\infty$, i.e., when the magnetization is completely held in place by the magnetizing field and the coupling with it can be neglected, $P(\mathbf{k})\to 1$; when $\Omega_0\to Q$, i.e., when the field outside the superconductor can be disregarded, $P(\mathbf{k})\to 0$. The latter condition is realized in the vicinity of the field $H_c \cong \beta M_0$ corresponding to the transition of the magnet from the homogeneous phase to the domain phase. Although the permeability component μ_{yy} of the magnet is large ($\mu_{yy}\gg 1$) near the critical field, suppression of the longitudinal vortex instability does not occur. Conversely, the influence of the field outside the superconductor decreases, and the critical field thus decreases. The reason for this becomes clearer, if we analyze the stray field in a structure consisting of a semi-infinite type-II superconductor and a magnet whose permeability tensor $\hat{\mu}$ has the nonzero components $\mu_{xx}=\mu_1$, $\mu_{yy}=\mu_2$, and $\mu_{zz}=1$. For the Fourier harmonics of the stray magnetic field in the superconductor ($y>0$) and in the magnet ($y<0$), in analogy to Ref. 21 we obtain the expression

$$\mathbf{H}_{\text{stray},\mathbf{k}} = \begin{cases} \left(i\mathbf{n}_x\dfrac{k_x\tau}{k^2}+i\mathbf{n}_z\dfrac{k_z\tau}{k^2}-\mathbf{n}_y\right)\dfrac{k^2}{k^2+\mu_2 p\tau}(B_v^y+B_v^{y'})_{\mathbf{k}}e^{-\tau y}, & y>0, \\ \left(i\mathbf{n}_x\dfrac{k_x}{k}+i\mathbf{n}_z\dfrac{k_z}{k}-\mathbf{n}_y\dfrac{p}{k}\right)\dfrac{k\tau}{k^2+\mu_2 p\tau}(B_v^y+B_v^{y'})_{\mathbf{k}}e^{ky}, & y<0, \end{cases}$$

where $\mathbf{B}_v+\mathbf{B}_v'$ is the sum of the fields of the vortices and their images on the surface of the superconductor, and $p^2=(k_z^2+\mu_1 k_x^2)/\mu_2$.

The stray field appears because of the jump in the normal component of the magnetic induction on the superconductor surface (surface ''magnetic charges''). It follows from the boundary condition on the normal component of the induction that, if $\mu_2\gg 1$, the stray field $H_{\text{stray}}\sim\mu_2^{-1/2}$, and the stray field energy $U_{\text{stray}}\sim 1/\mu_2$. The other contributions to the energy (2) of the vortices depend weakly on μ_2, and they determine the instability threshold in the vortex system. As follows from Eq. (23), when $\mu_2\to\infty$, the sources of the stray field are completely compensated by the surface ''magnetic charges'' created by the jump in the magnetization component m_y on the surface of the ferrite.

In the case of Gunn domain instability, a mode with a predominantly tangential component of the electric field in the semiconductor is unstable. As a consequence of the continuity of the tangential component of the electric field on the semiconductor–insulator interface, the electric induction is drawn into an insulator with a large dielectric constant $\varepsilon\gg 1$. The electric field energy of the unstable mode in the insulator is $\sim\varepsilon e^2$ (e is the amplitude of the electric field of the perturbation in the semiconductor and the insulator) and significantly exceeds the electric field energy in the semiconductor, $\sim e^2$. Thus the roles of the stray fields for the types of instability compared are different.

3. To calculate the critical current and the structure of the critical mode, we shall utilize a variational procedure, as in Ref. 17. An exact solution of this problem is possible even when the nonlocality is taken into account, but it does not have any special advantages over the solution obtained by the variational procedure, since the solution of the equation of state for the vortices has the form of an infinite series in powers of t^2. Let the displacement field in the superconductor be described by the trial functions

$$u_{\mathbf{k}'}^{x} = 2\pi^2 a e^{-\eta y}[\delta(\mathbf{k}'-k)+\delta(\mathbf{k}'+\mathbf{k})],$$

$$u_{\mathbf{k}'}^{y} = 2\pi i b e^{-\sigma y}[\delta(\mathbf{k}'-\mathbf{k})-\delta(\mathbf{k}'+\mathbf{k})], \tag{24}$$

where a, b, η, σ, and $\mathbf{k}$ are parameters that can be determined from the conditions for the minimum of the energy of the system U and the critical current.

Substituting (24) into (1)–(20), we bring the energy of the system U into the form

$$8U/C_{11}S = Aa^2+Bb^2-2tCab, \tag{25}$$

where S is the surface area of the superconductor, and the coefficients A, B, and C equal

$$A=\frac{1}{\eta}\left[k_p^2+\varepsilon(k_x^2+\eta^2)+\frac{k^2}{\lambda^2(\tau+\eta)^2}+2k_1\eta\right],$$

$$B=\frac{1}{\sigma}\left\{k_p^2+\varepsilon(k_x^2+\sigma^2)+\frac{1}{\lambda^2(\tau+\sigma)^2}\left[\frac{(\tau+2\sigma)}{\tau}k_z^2+\sigma^2+2\frac{\sigma k_z^2}{k}\frac{(\tau-k)}{\tau}P\right]+2k_2\sigma\right\},$$

$$C=\frac{2l\gamma}{(\gamma+\eta+\sigma)}. \tag{26}$$

Minimizing U with respect to a and b, we find the normalized critical current t and the polarization of the critical mode at the superconductor surface $r=a/b$ as functions of $\mathbf{k}$, σ, and η

$$t^2=\frac{AB}{C^2},\quad r=t\frac{C}{A}. \tag{27}$$

We go over to the normalized variables $\tilde{\tau}=\tau\lambda$, $\tilde{\gamma}=\gamma\lambda$, $\tilde{\eta}=\eta\lambda$, $\tilde{k}=k\lambda/\sqrt{\varepsilon}$, $\tilde{k}_p=k_p\lambda/\sqrt{\varepsilon}$, $\tilde{k}_1=k_1\lambda/\varepsilon$, and $\tilde{k}_2=k_2\lambda/\sqrt{\varepsilon}$. We shall henceforth use only these normalized quantities; therefore, the tilde over them will be omitted. As a result, we obtain

$$t^2=\sqrt{\varepsilon}\frac{(\gamma+\eta+\sigma\sqrt{\varepsilon})^2}{4\gamma^2k_z^2\eta\sigma}\left\{k_p^2+\varepsilon k_x^2+\eta^2+\frac{k^2}{\lambda^2(\tau+\eta)^2}+2k_1\eta\right\}\left\{k_p^2+\varepsilon(k_x^2+\sigma^2)+\frac{1}{\lambda^2(\tau+\sigma\sqrt{\varepsilon})^2}\left[\frac{(\tau+2\sigma\sqrt{\varepsilon})}{\tau}k_z^2+\sigma^2+2\frac{\sigma k_z^2}{k}\frac{(\tau-k\sqrt{\varepsilon})}{\tau}P\right]+2k_2\sigma\right\}, \tag{28}$$

$$\frac{A}{C}=\sqrt{\varepsilon}\frac{(\gamma+\eta+\sigma\sqrt{\varepsilon})}{4\gamma k_z\eta}\left\{k_p^2+\varepsilon k_x^2+\eta^2+\frac{k^2}{\lambda^2(\tau+\eta)^2}+2k_1\eta\right\}. \tag{29}$$

The wave vector of the critical mode increases with increasing pinning; therefore, the energy of the magnetic field outside the superconductor, which is proportional to $(\tau-k\sqrt{\varepsilon})/\tau$, decreases with increasing pinning. Comparing the relative magnitude of the terms in the square brackets in (28), which is proportional to k_z^2, we note that the relative influence of the field outside the superconductor decreases with increasing pinning. For this reason, the estimates of the critical current and the structure of the critical mode for the region of weak and moderate pinning are of greatest interest.

Minimization of the expression (28) for the normalized critical current t with respect to $\mathbf{k}$, η, and σ leads to a system of algebraic equations for determining the remaining independent parameters. Since this system can be solved exactly in its general form only by numerical methods, we shall henceforth confine ourselves to a simplified analysis, which is valid in the case of a vortex lattice with a small shear stiffness ($\varepsilon\ll1$). In addition, we shall neglect the dependence of Ω_0 on t, assuming that the critical current I_c is small ($t_c=4\pi I_c/cH_e\ll1$).

The longitudinal vortex instability thresholds at the surface and in the bulk of a thick superconductor were compared in Ref. 17 for a distribution of the transport current in a layer considerably thicker than the London penetration depth ($\gamma\ll1$) without consideration of the influence of the nonlocality. It was assumed that the results obtained can also be used for estimates in the case of $\gamma\cong1$. However, according to Ref. 17, for superconductors with moderate or strong pinning ($k_p^2>1$ or $k_2^2>1$) one has $\eta\gtrsim1$, and thus the local approximation is inapplicable in the general case. Taking the nonlocality into account greatly complicates the calculation of the instability threshold; however, when $\gamma\cong1$, it can be simplified for the most interesting cases of weak and moderate pinning ($k_1\ll\varepsilon^{-1}$, $k_p^2\ll\varepsilon^{-1}$, $k_2^2\ll\varepsilon^{-1}$) by setting $\varepsilon=0$ on the right-hand sides of Eqs. (28) and (29). This approximation enables one to find analytical expressions for the parameters $\mathbf{k}$, η, and σ and for t_2 and r as functions of the pinning constant in many special cases.

We present the asymptotic estimates of the parameters of the critical mode for two limiting cases. If the magnetization and the Q factor of the ferrite are small ($\Omega_0\gg1$, $\Omega_0\gg Q$, and $P\cong1$), the magnetization is held in place by the field, and the structure behaves as an isolated superconductor. If the external field is close to the field for passage of the magnet from the homogeneous to the domain phase ($\Omega_0\cong Q$, and $P\cong0$), the influence of the magnetic field of the critical mode outside the superconductor on the instability can be neglected. In both cases the x component of the wave vector of the critical mode $k_x=0$.

For a superconductor with weak pinning ($k_1\ll1$, $k_p\ll1$, and $k_2\ll1$) and $P=1$ the critical current and the polarization in the local limit equal

$$t_c^2\varepsilon^{1/2}=2,\quad r\varepsilon^{1/4}=2^{1/2}, \tag{30}$$

and the remaining parameters are

$$k=\eta=\sigma=\begin{cases}(k_1/2)^{1/2} & \text{for, } k_1\neq0,\quad k_p=k_2=0,\\ (3k_p^2/4)^{1/3} & \text{for, } k_p\neq0,\quad k_1=k_2=0,\\ (k_2/4)^{1/2} & \text{for, } k_2\neq0,\quad k_1=k_p=0.\end{cases} \tag{31}$$

For a superconductor with weak pinning and $P\cong0$ the expressions for the critical current and the polarization without consideration of the nonlocality have the form

$$t_c^2\varepsilon^{1/2}=1,\quad r\varepsilon^{1/4}=1, \tag{32}$$

and the inverse spatial scales equal

$$k=\eta=\sigma=\begin{cases}(k_1/2)^{1/2} & \text{for, } k_1\neq0,\quad k_p=k_2=0,\\ (k_p^2)^{1/3} & \text{for, } k_p\neq0,\quad k_1=k_2=0,\\ (k_2/2)^{1/2} & \text{for, } k_2\neq0,\quad k_1=k_p=0.\end{cases} \tag{33}$$

The expressions for the critical current and the polarization of the critical mode with the nonlocality taken into account coincide with (30) and (32), and the parameters k, η, and σ can be obtained from (31) and (33) by making the replacements $k_1 \to 2k_1$, $k_p \to 2k_p$, and $k_2 \to 2k_2$. Thus, taking the nonlocality into account leads only to quantitative changes in the spatial scales of the critical mode. The ratio between the critical currents for $P=1$ and $P=0$, as in the local limit, equals $\sqrt{2}$.

If the superconductor has moderate pinning, in the local limit (a) and with allowance for the nonlocality (b) for $P=1$ we have: in the range $1 \ll k_1 \ll \varepsilon^{-1}$, $k_p = k_2 = 0$

$$k=\sigma=(2k_1/3)^{1/2}, \quad \eta=1/3,$$
$$t_c^2\varepsilon^{-1/2}=8.71(k_1)^{1/2}, \quad r\varepsilon^{1/4}=(2/3k_1)^{1/4}, \tag{34a}$$
$$k=\sigma=(8k_1)^{1/2}, \quad \eta=1,$$
$$t_c^2\varepsilon^{-1/2}=(32k_1)^{1/2}, \quad r\varepsilon^{1/4}=(8/k_1)^{1/4}, \tag{34b}$$

in the range $1 \ll k_p \ll \varepsilon^{-1/2}$, $k_1 = k_2 = 0$

$$k=1.25k_p, \quad \sigma=1.6k_p, \quad \eta=1,$$
$$t_c^2\varepsilon^{-1/2}=9.34k_p, \quad r\varepsilon^{1/4}=3.6k_p^{1/2}, \tag{35a}$$
$$k=\sigma=(k_p)^{5/3}, \quad \eta=(k_p)^{2/3},$$
$$t_c^2\varepsilon^{-1/2}=2k_p, \quad r\varepsilon^{1/4}=(8k_p)^{1/6}, \tag{35b}$$

and in the range $1 \ll k_2 \ll \varepsilon^{-1/2}$, $k_1 = k_p = 0$

$$k=\sigma=k_2^{1/3}, \quad \eta=1,$$
$$t_c^2\varepsilon^{-1/2}=2k_2, \quad r\varepsilon^{1/4}=(8k_2)^{1/6}, \tag{36a}$$
$$k=\sigma=k_2/2, \quad \eta=(k_2^2/12)^{1/4},$$
$$t_c^2\varepsilon^{-1/2}=2.28(k_2^2)^{1/4}, \quad r\varepsilon^{1/4}=(12k_2^2)^{1/8}. \tag{36b}$$

Accordingly, for a superconductor with moderate pinning in a field close to the transition field in the magnet ($P \cong 0$), we have: in the range of pinning constants $1 \ll k_1 \ll \varepsilon^{-1}$, $k_p = k_2 = 0$

$$k=\sigma=(2k_1/3)^{1/2}, \quad \eta=1/3,$$
$$t_c^2\varepsilon^{-1/2}=4.35(k_1)^{1/2}, \quad r\varepsilon^{1/4}=(6k_1)^{-1/4}, \tag{37a}$$
$$k=\sigma=(8k_1)^{1/2}, \quad \eta=1,$$
$$t_c^2\varepsilon^{-1/2}=(8k_1)^{1/2}, \quad r\varepsilon^{1/4}=(2/k_1)^{1/4}, \tag{37b}$$

in the range $1 \ll k_p \ll \varepsilon^{-1/2}$, $k_1 = k_2 = 0$

$$k=\sigma=2^{1/2}k_p, \quad \eta=1,$$
$$t_c^2\varepsilon^{-1/2}=3^{3/2}k_p, \quad r\varepsilon^{1/4}=(4/3k_p^2)^{1/4}, \tag{38a}$$
$$k=\sigma=(k_p)^{5/3}, \quad \eta=(k_p)^{2/3},$$
$$t_c^2\varepsilon^{-1/2}=k_p, \quad r\varepsilon^{1/4}=(k_p)^{1/6}; \tag{38b}$$

and in the range $1 \ll k_2 \ll \varepsilon^{-1/2}$ $k_1 = k_p = 0$

$$k=\sigma=(2k_2)^{1/3}, \quad \eta=1,$$
$$t_c^2\varepsilon^{-1/2}=2k_2, \quad r\varepsilon^{1/4}=(2k_2)^{1/6}, \tag{39a}$$
$$k=\sigma=k_2, \quad \eta=(k_2^2/3)^{1/4},$$
$$t_c^2\varepsilon^{-1/2}=1.75(k_2^2)^{1/4}, \quad r\varepsilon^{1/4}=(3k_2^2)^{1/8}. \tag{39b}$$

Comparing Eqs. (34a)–(39a) with (34b)–(39b), we note that at moderate pinning the spatial scales of the critical mode with the nonlocality taken into account can differ from their values in the local limit not only quantitatively, but also qualitatively. The longitudinal vortex instability threshold is found to be lower in the nonlocal theory than in the local theory. This is due to the decrease in the stiffness modulus $C_{11}(\mathbf{k})$ with increasing k in the nonlocal theory. The ratios between the vortex instability thresholds in an isolated superconductor ($P=1$) and in a superconductor–ferromagnet structure near the transition point to the domain phase ($P=0$) for the nonlocal and local models do not differ strongly, if k_1 or k_p is large [compare Eqs. (34)–(37) and (35)–(38)]. However, if k_2 is large, in the local limit $k \sim \sigma \ll k_2$ [see Eqs. (36a) and (39a)], and the influence of the field outside the superconductor is small [this can be seen directly from the expression (28) for the critical current]. When the nonlocality is taken into account, $k \sim \sigma \sim k_2$ [see (36b)–(39b)], and the thresholds differ by about 15%. We also note that the longitudinal vortex instability threshold for $1 \ll k_2 \ll 1/\varepsilon^{1/2}$ and $k_1 = k_p = 0$ is proportional to k_2 in the local limit and is proportional to $k_2^{1/2}$ when the nonlocality is taken into account, i.e., the threshold increases with increasing k_2 significantly more slowly in the nonlocal theory than in the local theory. It is seen that in the region of moderate pinning $t_c \ll 1$, and thus the variation of the external field in the ferromagnet due to the transport current can be neglected ($\Omega_0 \cong \Omega_e$). When the inequality $t_c \ll 1$ holds, the simplifications that were made in Eqs. (28) and (29) are automatically valid.

If the magnetizing field has a perpendicular component $Q < H_e < \infty$, the x component of the wave vector of the critical mode k_x is nonzero in the general case. However, numerical calculations are then needed to determine the instability threshold and the parameters of the critical mode.

[1] S. Yu. Bereza, Yu. I. Gorobets, and A. A. Simonov, Fiz. Tverd. Tela (Leningrad) **34**, 1903 (1992) [Sov. Phys. Solid State **34**, 1015 (1992)].
[2] Yu. I. Bespyatykh, W. Wasilewski, M. Gajdek *et al.*, Fiz. Tverd. Tela (St. Petersburg) **36**, 586 (1994) [Phys. Solid State **36**, 323 (1994)].
[3] Yu. I. Bespyatykh, W. Wasilewski, M. Gajdek *et al.*, Fiz. Tverd. Tela (St. Petersburg) **37**, 2611 (1995) [Phys. Solid State **37**, 1434 (1995)].
[4] Yu. I. Bespyatykh, I. E. Dikshteĭn, and V. V. Tarasenko, Fiz. Tverd. Tela (Leningrad) **23**, 3013 (1981) [Sov. Phys. Solid State **23**, 1757 (1981)].
[5] E. B. Sonin, Pis'ma Zh. Tekh. Fiz. **14**, 1640 (1988) [Sov. Tech. Phys. Lett. **14**, 714 (1988)].
[6] B. M. Lebed' and S. V. Yakovlev, Pis'ma Zh. Tekh. Fiz. **15**(19), 27 (1989) [Sov. Tech. Phys. Lett. **15**, 757 (1989)].
[7] Yu. I. Bespyatykh, W. Wasilewski, M. Gajdek *et al.*, Fiz. Tverd. Tela (St. Petersburg) **35**, 2983 (1993) [Phys. Solid State **35**, 1466 (1993)].
[8] Yu. I. Bespyatykh, W. Wasilewski, M. Gajdek, and V. D. Kharitonov, Fiz. Tverd. Tela (St. Petersburg) **37**, 3049 (1995) [Phys. Solid State **37**, 1680 (1995)].
[9] V. S. Babushkin and N. A. Morozova, Pis'ma Zh. Tekh. Fiz. **17**(19), 1 (1991) [Sov. Tech. Phys. Lett. **17**, 681 (1991)].
[10] Yu. I. Bespyatykh, W. Wasilewski, M. Gajdek *et al.*, Fiz. Tverd. Tela

(Leningrad) **33**, 1545 (1991) [Sov. Phys. Solid State **33**, 870 (1991)].
[11] Yu. I. Bespyatykh, A. D. Simonov, and V. D. Kharitonov, Fiz. Met. Metalloved. (4), 87 (1992).
[12] A. F. Popkov, Pis'ma Zh. Tekh. Fiz. **15**(5), 9 (1989) [Sov. Tech. Phys. Lett. **15**, 166 (1989)].
[13] N. I. Polzikova and A. O. Raevskiĭ, Pis'ma Zh. Tekh. Fiz. **16**(22), 59 (1990) [Sov. Tech. Phys. Lett. **16**, 871 (1990)].
[14] A. Campbell and J. Evetts, *Critical Currents in Superconductors*, Taylor and Francis, London (1972) [Russ. transl., Mir, Moscow (1975)].
[15] R. P. Khyubener, *Structure of Magnetic Fluxes in Superconductors* [in Russian], Mashinostroenie, Moscow (1984).
[16] J. R. Clem, Phys. Rev. Lett. **38**, 1425 (1977).
[17] E. H. Brandt, Low Temp. Phys. **44**, 59 (1981).
[18] V. L. Bonch-Bruevich, I. P. Zvyagin, and A. G. Mironov, *Domain Electrical Instabilities in Semiconductors*, Consultants Bureau, New York (1975).
[19] P. W. Shumate, IEEE Trans. Magn. **7**, 586 (1971).
[20] V. V. Tarasenko, E. V. Chenskiĭ, and I. E. Dikshteĭn, Zh. Éksp. Teor. Fiz. **70**, 2178 (1976) [Sov. Phys. JETP **43**, 1136 (1976)].
[21] E. H. Brandt, Low Temp. Phys. **42**, 557 (1981).
[22] E. H. Brandt, Low Temp. Phys. **44**, 33 (1981).
[23] V. V. Shmidt and G. S. Mkrtchyan, Usp. Fiz. Nauk **112**, 459 (1974) [Sov. Phys. Usp. **17**, 170 (1974)].
[24] V. G. Kogan, Phys. Rev. B **21**, 2799 (1980).
[25] A. Sikora and B. Makiej, Phys. Status Solidi A **88**, K197 (1985).

Translated by P. Shelnitz

Amorphization and shear microbands near grain boundaries in mechanically alloyed metallic materials

I. A. Ovid'ko and A. V. Osipov

Institute of Mechanical-Engineering Problems, Russian Academy of Sciences, 199178 St. Petersburg, Russia
(Submitted March 7, 1996)
Zh. Tekh. Fiz. **67**, 35–38 (July 1997)

A theoretical model which effectively describes a novel micromechanism of solid-state amorphization in mechanically alloyed metallic materials is proposed. Within the model the amorphous phase nucleates as a result of the intersection of a grain boundary by a plastic-shear microband. It is shown that the growth of a nucleus of the amorphous $\alpha-\beta$ alloy near a site on a grain boundary intersected by a shear microband is effectively controlled by the rate of diffusion mixing of atoms of the mechanically alloyed metals α and β. © *1997 American Institute of Physics.* [S1063-7842(97)00707-1]

INTRODUCTION

The phenomenon of solid-state amorphization in mechanically alloyed metallic materials has been a subject of intensive theoretical and experimental research (see, for example, Refs. 1–4). There is special interest in the study of the micromechanisms of solid-state amorphization, whose disclosure is extremely important for determining the optimum parameters of any technology for synthesizing amorphous metallic alloys by mechanical alloying.

Mechanical alloying provides a powder of an amorphous metallic alloy ($\alpha-\beta$) as a result of the treatment of a mixture of powders of the crystalline metals (α and β) in a ball mill for many hours. The micromechanisms of solid-state amorphization in mechanically alloyed materials (in which intensive diffusion and plastic deformation take place) are associated with diffusion and/or plastic deformation. These amorphization micromechanisms include the special, diffusion-induced migration of grain boundaries,[4,5] the splitting of disclinations at three-way grain-boundary joints,[6–8] and the formation of high-density ensembles of point defects.[9,10] This paper offers a theoretical model description of a novel micromechanism of solid-state amorphization in mechanically alloyed materials, viz., the diffusion-limited growth of an amorphous phase near grain boundaries that are intersected by plastic-shear microbands.

SHEAR MICROBANDS, DIFFUSION, AND NUCLEATION OF AN AMORPHOUS PHASE NEAR GRAIN BOUNDARIES IN MECHANICALLY ALLOYED MATERIALS

The novel micromechanism of solid-state amorphization proposed in this paper is associated with both plastic deformation and diffusion processes in mechanically alloyed materials. Let us first discuss the role of plastic deformation in the nucleation of an amorphous phase near grain boundaries in mechanically alloyed materials.

Treatment in a ball mill induces high degrees of spatially nonuniform plastic deformation in metallic crystalline powders. Such deformation is often localized in crystals in narrow plastic-shear bands consisting of high-density groups of moving dislocations.[11,12] As dislocations migrate in a polycrystalline material, they intersect grain boundaries (Fig. 1a), which are effective stops for moving defects (see, for example, Refs. 11 and 13). Dislocation groups are usually stopped by grain boundaries. However, under the effect of high tangential stresses, dislocations can ''overcome'' barriers in the form of grain boundaries and thereby become able to continue to move into the adjacent crystalline grain.

One of the possible ways for lattice dislocations to ''overcome'' grain boundaries is as follows. When a microscopic shear band and a grain boundary intersect, the dislocations found at the edge of the shear microband enter the boundary (Fig. 1b). The core of each such dislocation is delocalized in the boundary (Fig. 1a). More precisely, a dislocation with a Burgers lattice vector splits in the grain boundary into several new grain-boundary dislocations with small Burgers vectors belonging to the displacement-shift complete lattice of the boundary (Fig. 1a) (such splitting of implanted dislocations is common in grain boundaries[13]). Then, under the action of the high tangential stresses (and the diffusional driving force; see below) the new dislocations with small Burgers vectors pass into the neighboring grain (Fig. 1d). Since the Burgers vectors of the new dislocations are not lattice vectors in the crystalline grain, a local region with an unordered structure (an ''extended stacking fault'') forms in the grain behind the leading edge of motion of such dislocations and is naturally interpreted as a nucleus of the amorphous phase (Fig. 1d).

Thus, the type of intersection of a grain boundary by a microscopic plastic-shear band just described leads to the appearance of a nucleus of the amorphous phase near the grain boundary. The nucleus of the amorphous phase is essentially an amorphous microscopic plastic-shear band, which propagates from the grain boundary toward the interior of the grain (Fig. 1d, hatched area).

Let us consider the character of the driving forces which act when an amorphous phase is formed according to the above scheme in greater detail. The first driving force F_m has a mechanical nature; it is the result of the action of the tangential stresses on the dislocations. The second driving force F_d has a diffusional character. In fact, intense diffusion mixing of the atoms of the mechanically alloyed metals α and β takes place during treatment in ball mills (see, for example, Refs. 1–3). The metals α and β participating in solid-

 1063-7842/97/070748-04$10.00

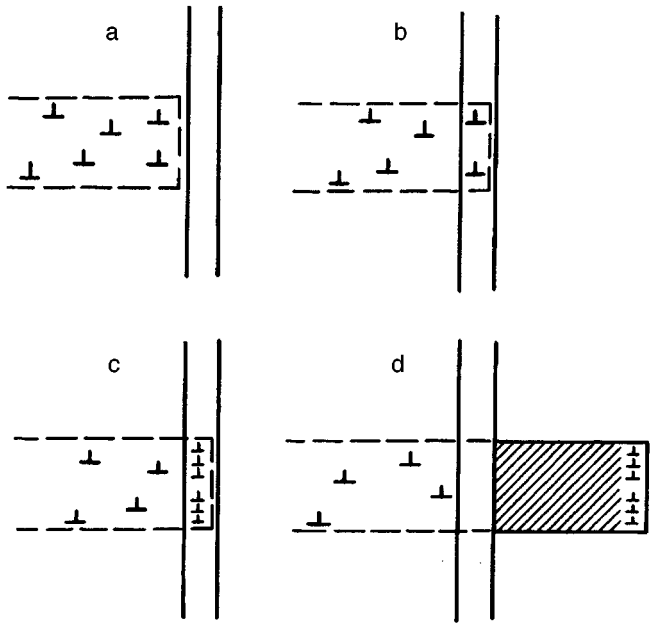

FIG. 1. Intersection of a shear microband with a grain boundary accompanied by amorphization.

state amorphization (the formation of an amorphous $\alpha-\beta$ alloy) exhibit the following experimentally discovered feature during mechanical alloying.[1,3] The heat of diffusion mixing of atoms of α and β is negative in the case of a ''β crystal to amorphous $\alpha-\beta$ phase'' transformation. In other words, such a phase transformation is a thermodynamically advantageous process, whose rate is controlled by the diffusion rate of atoms of α in β. This causes the appearance of the diffusional driving force F_d for the nucleation and propagation of a shear microband (Fig. 1d) having the structure of the amorphous $\alpha-\beta$ alloy.

More precisely, diffusion participates in the nucleation and propagation of an amorphous shear microband in β in the following manner. Since the grain-boundary diffusion rate is many times higher than the bulk diffusion rate,[13] atoms of α (filled circles) at first diffuse along grain boundaries of β (from the contact surface between particles of α and β into a particle of β) (Fig. 2a). Then, at a site on a grain boundary intersected by a shear microband, the atoms of α diffuse into the grain parallel to the motion of the leading edge of the shear microband (Fig 2b). Here (1) the microscopic shear band has the structure of the amorphous $\alpha-\beta$ alloy, and (2) the simultaneous effective action of the driving forces F_m and F_d in the nucleation and further evolution of the amorphous $\alpha-\beta$ phase near the grain boundary is ensured.

Generally speaking, nucleation of the amorphous $\alpha-\beta$ phase can occur near the boundary of a grain of β under the action of F_d at any site on the boundary. However, the additional action of the mechanical driving force F_m makes nucleation of the amorphous phase at the sites on the grain boundary intersected by shear microbands preferable. In fact, at such sites on the grain boundary the driving force of the nucleation of the amorphous $\alpha-\beta$ phase is F_m+F_d, while at other sites the driving force is $F_d<F_m+F_d$.

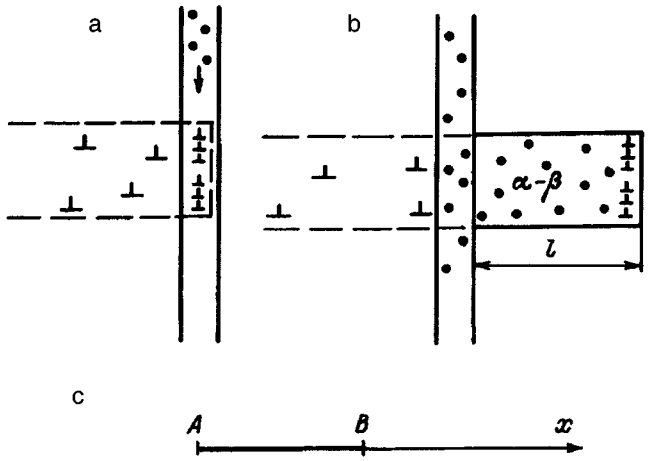

FIG. 2. Amorphization near a grain boundary involving diffusion (l is the time-varying length of the amorphous shear microband).

The condition of simultaneous and thus effective action of the driving forces F_m and F_d of solid-state amorphization defines the growth rate of the amorphous $\alpha-\beta$ phase (or, stated differently, the rate of propagation of an amorphous shear microband) as the velocity of the diffusion mixing front of the atoms of α and β. This feature will be taken into account in the next section in describing the evolution of an amorphous shear microband.

EVOLUTION OF AMORPHOUS MICROSCOPIC SHEAR BANDS IN MECHANICALLY ALLOYED MATERIALS

Let us investigate the characteristics of the evolution of an amorphous shear microband propagating from a grain boundary into the interior of a grain with the velocity of the diffusion mixing front of the atoms of α and β. For simplicity, we assume that an amorphous shear microband propagates in one direction, which corresponds to the direction of propagation of the original (''crystalline'') microscopic shear band before it intersected the grain boundary. This natural assumption permits the effective use of the one-dimensional model of an amorphous shear microband, within which such a shear microband is described as a one-dimensional segment AB between point A (which corresponds to a grain boundary) and point B (which corresponds to the leading edge of the amorphous microscopic shear band) (Fig. 2c). Now the diffusion of the atoms of α from grain boundary A into the crystalline grain and the simultaneous (diffusion-controlled) motion of the leading edge B of the amorphous shear microband are described by the following system of equations:

$$\frac{\partial n}{\partial t}=D\frac{\partial^2 n}{\partial x^2}, \tag{1}$$

$$n(0,\,t)=n_0, \tag{2}$$

$$n(l,(t),t)=n_c, \tag{3}$$

$$\frac{dl}{dt}=-\frac{D}{n_0}\frac{\partial n}{\partial x}\bigg|_{x=l}. \tag{4}$$

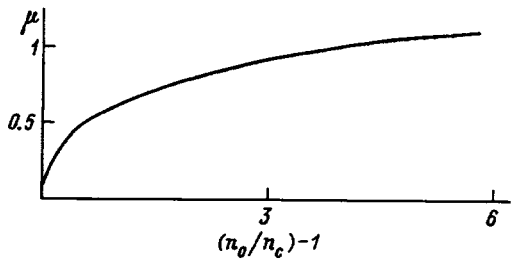

FIG. 3. Dependence of μ on (n_0/n_c-1).

Here $n(x,t)$ is the concentration of atoms of α within the amorphous shear microband; x is the coordinate along the direction of motion of the amorphous shear microband (Fig. 2b); t is the time; D is the diffusion coefficient of the atoms of α in the amorphous $\alpha-\beta$ phase (in the amorphous shear microband); n_c is the concentration of atoms of α at the leading edge B of the amorphous microscopic shear band (or, stated differently, the concentration corresponding to equilibrium between the amorphous and crystalline phases); n_0 is the concentration of atoms at grain boundary A; and l is the length of the amorphous shear microband (Fig. 2b). Equation (1) describes the diffusion of atoms of α from grain boundary A in the direction of propagation of the amorphous shear microband. Equations (2) and (3) are the boundary conditions for the concentration of atoms of α at grain boundary A and leading edge B, respectively. Equation (4) describes the diffusion-controlled motion of the leading edge B of the amorphous shear microband (or, stated differently, the motion of the boundary B between the amorphous and crystalline phases).

The solution of system of equations (1)–(4) has the following form:

$$n(x,t)=n_0-(n_0-n_c)f[x/2(Dt)^{1/2}], \tag{5}$$

$$l(x,t)=2\mu(Dt)^{1/2}, \tag{6}$$

where μ is an unknown dimensionless constant.

We use the following procedure to find μ. First, from Eqs. (2)–(3) and (5) we obtain the system of equations

$$\frac{d^2f}{dz^2}+2z\frac{df}{dz}=0, \tag{7}$$

$$f(0)=0, \quad f(\mu)=1, \tag{8}$$

where $z=x/2(Dt)^{1/2}=\mu x/l(t)$ is the argument of $f(z)$.

The solution of system of equations (7) and (8) is

$$f(z)=\Phi(z)/\Phi(\mu), \quad \text{where} \quad \Phi(z)=2\pi^{1/2}\int_0^z \exp(-y^2)dy$$

is a standard error function. With consideration of this result, from Eq. (4) we obtain the following nonlinear equation for μ:

$$\pi^{1/2}\mu\exp(\mu^2)\Phi(\mu)=n_0/n_c-1. \tag{9}$$

The exact numerical solution of Eq. (9) is presented in Fig. 3.

For the range of typical values $n/n-1<6$ we have $\mu\ll 1$. Then

$$\Phi(\mu)\approx 2\pi^{-1/2}\mu(1-\mu^2/3), \tag{10a}$$

$$\exp(\mu^2)\approx 1+\mu^2. \tag{10b}$$

Substituting (10) into (9), we obtain the following approximate analytical expression for μ:

$$\mu=\frac{\sqrt{3}}{2}\left[\sqrt{\frac{4}{3}\left(\frac{n_0}{n_c}-1\right)+1}-1\right]^{1/2}. \tag{11}$$

From Eqs. (6) and (11) we obtain

$$l(t)=[(\sqrt{3}\sqrt{4n_0/n_c-1}-3)Dt]^{1/2}. \tag{12}$$

We use Eq. (12) to estimate the mean rate of motion l/t of the leading edge of the amorphous shear microband (or, stated differently, the growth rate of the amorphous phase near the grain boundary intersecting the original microscopic plastic-shear band). At first, from Eq. (12) we obtain $l=6-130$ Å for the typical values $D\approx 10^{-21}-10^{-19}\text{m}^2\cdot\text{s}^{-1}$, $n_0/n_c\approx 2-3$, and $t\approx 600$ s. Thus $l/t\approx 0.01-0.2$ Å/s.

DISCUSSION OF RESULTS. CONCLUSIONS

There are convincing data indicating that the micromechanisms of solid-state amorphization in crystalline materials during thermal treatment and mechanical alloying are similar.[1] At the same, the range of concentrations of the components α and β of an amorphous alloy synthesized by mechanical alloying is broader than in the case of thermal treatment. The similarity and differences between the amorphization processes taking place during thermal treatment and mechanical alloying are naturally explained in the following manner.[4,5] There are diffusional micromechanisms of solid-state amorphization, which are basic and identical for the cases of thermal treatment and mechanical alloying. They account for the similarity between the solid-state amorphization processes taking place during thermal treatment and mechanical alloying. At the same time, along with the diffusional micromechanisms of amorphization, in mechanical alloying contributions are also made by ''additional'' micromechanisms, which have a mechanical character (i.e., are associated only with plastic deformation) or have a ''mixed'' (mechanical and diffusional) character. They account for the difference between the solid-state amorphization processes taking place during thermal treatment and mechanical alloying.

It has been shown in this work that one of the ''additional'' micromechanisms of amorphization operating during mechanical alloying is the diffusion-controlled propagation of an amorphous shear microband. This micromechanism, which is simultaneously associated with diffusion and plastic deformation, is characterized quantitatively by the mean rate of motion l/t of the leading edge of the amorphous shear microband. To estimate the contribution of the micromechanism under consideration to solid-state amorphization during mechanical alloying, it would be natural to compare the characteristic rate l/t and the experimentally measured growth rate of the amorphous phase during mechanical al-

loying. However, as far as we know, there are no data in the scientific literature from direct experiments aimed at measuring the growth rate of the amorphous phase during mechanical alloying. Under these circumstances we can only compare the model rate $l/t \approx 0.01-0.2 \text{Å}\cdot\text{s}^{-1}$ (see the preceding section) only with the experimentally measured[14] growth rate $l/t \approx 0.01-0.7 \text{Å}\cdot\text{s}^{-1}$ of the amorphous phase in metallic materials (''sandwiches'') during thermal treatment.

Within the ideas regarding the similarity and differences between the amorphization processes taking place during thermal treatment and mechanical alloying presented above (in the present section) a comparison of l/t and $\tilde{l}/t$ provides evidence that the diffusion-controlled motion of amorphous shear microbands (Fig. 2b) can make a significant contribution as an amorphization micromechanism to solid-state amorphization in metallic materials during mechanical alloying.

This work was performed with support from the Russian Foundation for Fundamental Research (Grant No. 95-02-03807) and the International Science Foundation (Grants Nos. R4F000 and R4F300).

[1] W. L. Johnson, Mater. Sci. Eng. **97**, 1 (1988).
[2] L. Schultz, Mater. Sci. Eng. **97**, 15 (1988).
[3] K. Samwer, Phys. Rep. **161**, 1 (1988).
[4] I. A. Ovid'ko, *Defects in Condensed Media: Glasses, Crystals, Quasicrystals, Magnets, and Superfluids* [in Russian] Nauka, Leningrad (1991).
[5] I. A. Ovid'ko, J. Phys. D **2,** 2190 (1991).
[6] I. A. Ovid'ko and A. V. Osipov, Fiz. Tverd. Tela (Leningrad) **34**, 288 (1992) [Sov. Phys. Solid State **34**, 152 (1992)].
[7] A. V. Osipov and I. A. Ovid'ko, Appl. Phys. A **54**, 517 (1992).
[8] M. Yu. Gutkin and I. A. Ovid'ko, Philos. Mag. A **70**, 561 (1994).
[9] M. Yu. Gutkin, I. A. Ovid'ko, and Yu. I. Meshcheryakov, J. Phys. III (Paris) **3**, 1563 (1993).
[10] I. A. Ovid'ko, J. Phys. D **27**, 999 (1994).
[11] R. W. Honeycombe, *The Plastic Deformation of Metals*, St. Martin, New York (1968) [Russ. transl., Mir, Moscow (1972)].
[12] B. I. Smirnov, *Dislocation Structure and Ordered Crystals* [in Russian], Nauka, Leningrad (1981).
[13] O. A. Kaĭbyshev and R. Z. Valiev, *Grain Boundaries and Properties of Metals* [in Russian], Metallurgiya, Moscow (1986).
[14] K. Samwer, H. Schroder, and K. Pampus, Mater. Sci. Eng. **97**, 63 (1988).

Translated by P. Shelnitz

Features of the two-dimensional modeling of drift injection magnetosensitive structures

M. A. Glauberman, V. V. Egorov, N. A. Kanishcheva, and V. V. Kozel

Instructional, Scientific, and Industrial Center, I. I. Mechnikov Odessa State University, 270063 Odessa, Ukraine.

(Submitted May 7, 1996)
Zh. Tekh. Fiz. **67**, 39–41 (July 1997)

The correctness of the two-dimensional description of the transfer of injected carriers in the base regions of drift injection magnetosensitive structures is substantiated. A criterion for selecting one of these models in accordance with the technological–design parameters and the electrical regime of the structures is obtained. © *1997 American Institute of Physics.* [S1063-7842(97)00807-6]

Magnetotransistors incorporate all the basic physical features of drift injection magnetosensitive structures and are, at the same time, the simplest magnetosensitive structures in this category. Therefore, we shall analyze the physical processes in them in the case of a magnetotransistor or, more precisely, a geometrically symmetric modification of a two-collector magnetotransistor using the scheme shown in Fig. 1.

A rigorous mathematical model requires solving the equation for the current density of the carriers of both signs together with the continuity equation of the minority carriers (holes in the present case). For the stationary case, a weak magnetic field ($\mu_p, \mu_n \ll B^{-1}$), and a low injection level (under the condition of uniform conductivity in the base) it can be written in the form[1]

$$j_p = ep\mu_p \mathbf{E} - e\varphi_T \mu_p \nabla p - \frac{3\pi}{8} ep\mu_p^2 (\mathbf{B}\times\mathbf{E}) + \frac{3\pi}{8} e\varphi_T \mu_p^2 [\mathbf{B}\times\nabla p],$$

$$j_n = en\mu_n \mathbf{E} + e\varphi_T \mu_n \nabla n + \frac{3\pi}{8} en\mu_n^2 (\mathbf{B}\times\mathbf{E}) + \frac{3\pi}{8} e\varphi_T \mu_n^2 (\mathbf{B}\times\nabla n),$$

$$\operatorname{div} \mathbf{j}_p + p/\tau = g. \tag{1}$$

Here e is an elementary charge; p, n, μ_p, and μ_n are the hole and electron number densities and mobilities; $\mathbf{E}$ is the resultant vector of the electric fields in the base; $\varphi_T \equiv kT/e$ is the temperature potential; $\mathbf{B}$ is the magnetic induction; and τ and g are the lifetime and flux density of the minority-carrier source.

It is very difficult to obtain a solution of system of equations (1) in a general form. However, as the experiment in Ref. 2 shows, the influence of the components of the magnetic induction that are parallel to the surface of the structure can be neglected, and the treatment can thereby be simplified. Then, in the XYZ rectangular coordinate system, in which the $0Z$ axis is perpendicular to that surface and the $0X$ axis is parallel to the accelerating field E_0, we can write the following expression for the excess concentration of minority carriers on the basis of system (1) under the assumption that the base is electroneutral

$$L^2\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2}\right) - 2\eta L \frac{\partial p}{\partial x} + 2\eta L(\mu_n^* + \mu_p^*) B \frac{\partial p}{\partial y} + \frac{E_z L^2}{2\varphi_T} \frac{\partial p}{\partial z} - P = -g\tau, \tag{2}$$

where L is the diffusion length of the holes, μ_n^* and μ_p^* are the Hall conductivities, and $\eta \equiv E_0 L^2/2\varphi_T$ is the field coefficient.

The determination of the three-dimensional concentration field, according to Eq. (2), is a difficultly solvable (totally unsolvable in analytic form) problem; therefore, a two-dimensional approach has been taken universally to model drift magnetotransistors (DMTs).[1,3–5] It was assumed without rigorous proof in the publications just cited that $\partial^n/\partial z^n \equiv 0$ $(n=1,2)$. As a result, several three-dimensional effects that are obviously significant were completely ignored.

In two-dimensional modeling, instead of the bulk (three-dimensional) concentration of the injected holes we naturally consider their surface (two-dimensional) concentration

$$P(x,y) \equiv \int_0^{z_0} p(x,y,z)\,dz \tag{3}$$

and automatically eliminate the problem of taking into account the hole concentration distribution along $0Z$. Then Eq. (2) in coordinate form becomes

$$L^2\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2}\right) - 2\eta L \frac{\partial P}{\partial x} + 2\eta L(\mu_n^* + \mu_p^*) B \frac{\partial P}{\partial y} - P = -\Gamma\tau, \tag{4}$$

where

$$\Gamma = -\mu\varphi_T \frac{\partial p}{\partial z}\Big|_{z=0} \equiv j_s/e \tag{4a}$$

is the surface (two-dimensional) flux density of the injection sources (the number of injected carriers created on a unit of

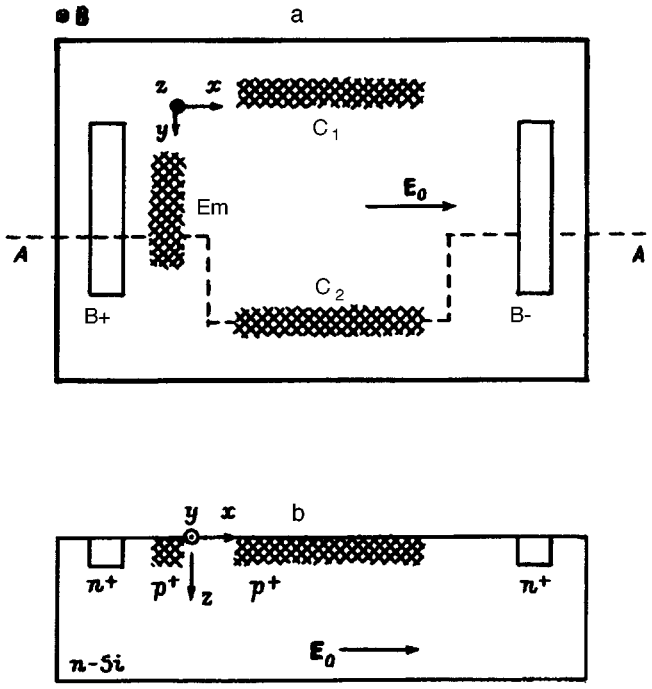

FIG. 1. Drift horizontal magnetotransistor with a transverse magnetic axis: a — top view; b — cross section at $A-A$; B^+, B^- — contacts of the base; Em — emitter; C_1, C_2 — collectors; $\mathbf{E}_0$ — accelerating electric field.

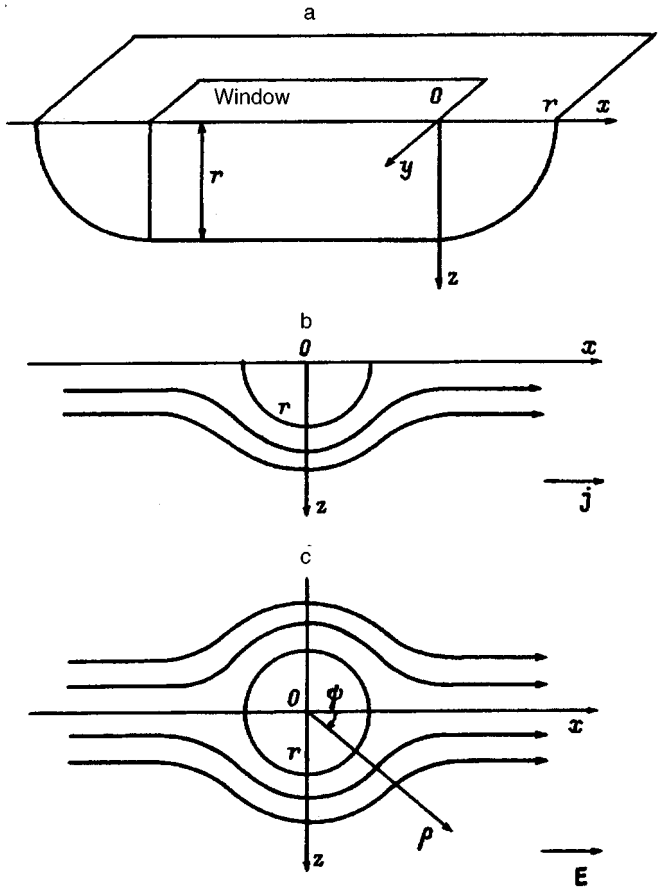

FIG. 2. Calculation of the potential in the near-emitter region of the base: a — cross section of the emitter in the $V=$const plane, b — emitter model, c — model for calculating the potential.

the surface $z=0$, or, stated differently, the component j_s of the injection current density normal to that surface per elementary charge).

Here it is natural to consider the case of the absence of bulk injection sources ($g\equiv 0$). In addition, it is assumed in Eq. (4a) that $\partial p/\partial z=0$ at $z=z_0$. Under real conditions this corresponds to a DMT structure fabricated on a single crystal or in a semiconductor layer on an insulating substrate. For DMTs fabricated in an epitaxial layer on a semiconductor substrate of the opposite type of conductivity, the extracting influence of the layer-substrate pn junction can be taken into account by introducing an effective lifetime instead of the bulk lifetime.

To correctly use Eq. (4) in practice we must now describe the injection process.

Figure 2a presents a cross section of the emitter of a DMT in the $Y=$const plane. It is assumed here that the emitter is formed by diffusion or implantation of an impurity through the window $(0;x_1)$. The regions where the impurity penetrates beneath the mask are modeled here by the sectors of radius r in accordance with the generally accepted approach.[6,7] To determine the injection parameters in the region $0\leqslant x\leqslant r$, following Ref. 8, we model the emitter by an injecting semicylinder (Fig. 2b).

To determine the potential φ of the accelerating field in the near-emitter region of the base, we supplement the configuration depicted in Fig. 2b by its mirror image relative to the $z=0$ plane. We do not obtain the problem of the flow of a current around an insulating cylinder (Fig. 2c). The transition from one configuration to the other in Fig. 2 is perfectly legitimate, because $\partial j/\partial z=0$ at $z=0$. Under the assumptions that the base is quasielectroneutral and homogeneous, for the current density of the majority carriers we have $j\propto E$, which allows us to rewrite the continuity equation in the form

$$\nabla^2\varphi=0.$$

Supplementing it with the boundary conditions (ρ and ϕ are, respectively, the radius vector and the polar angle in cylindrical coordinates)

$$\frac{\partial\varphi}{\partial\rho}\Big|_{\rho=r}=0,\quad \varphi(\rho=\infty)=-E_0\rho\cos\phi,$$

we obtain a mathematical formulation of the problem posed. Its solution has the form

$$\varphi=-[1+r^2/\rho^2]E_0\rho\cos\phi$$

or in Cartesian coordinates for the surface of a cylinder $(\rho=r)$

$$\varphi=-2E_0x.$$

Taking into account the exponential dependence between the bulk carrier concentration p and the bias in the emitter junction (we neglect the component of the emitter current created by recombination in the space-charge region), we obtain

$$p(x)=p(0)\exp(2E_0x/\varphi_T). \tag{5}$$

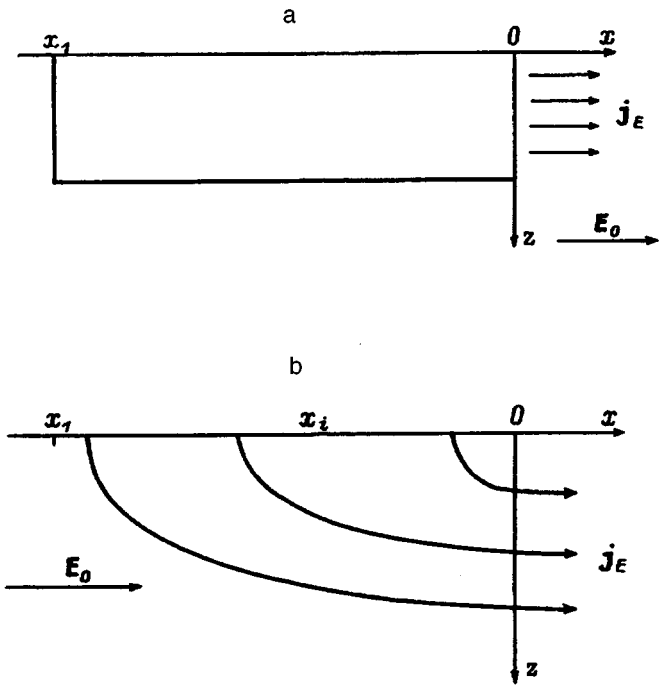

FIG. 3. Models of the emitter of a drift horizontal magnetic structure: a — vertical model, b — horizontal model.

Let us determine the number of injected carriers in the base region at the boundary with the emitter junction per unit of its length along $0Y$ (the linear density) for the cylindrical ($0 \leqslant x \leqslant r$) and flat ($x<0$) parts of the emitter. For this purpose, in the former case we integrate Eq. (5) in the range indicated, and in the latter case, assuming that the electric field is unperturbed at $x<0$, we integrate the expression

$$p(x)=p(0)\exp(E_0 x/\varphi_T)$$

in the range $-\delta x<x<0$, where $\delta x \equiv V_0/E_0$ is the width of the injecting portion of the emitter and V_0 is the bias on the emitter junction at the point $x=0$.

As a result, for the ratio χ between the linear densities in the two cases when $V_0 \gg \varphi_T$ we obtain

$$\chi=\frac{\exp(2E_0 r/\varphi_T)-1}{2}. \qquad (6)$$

When $\chi \gg 1$, practically the entire flux of injected carriers will clearly emerge from the cylindrical part of the emitter, and the flux from the flat part can be neglected. Since r, which amounts to a single micron or a fraction of a micron in practice, is small compared with the other parameters of the structure (which are equal to tens or hundreds of microns), the error in the values of x in the range from 0 to r can be neglected, and it can be assumed that the emitter is confined to the portion of the $x=0$ plane at $0 \leqslant z \leqslant r$ (the vertical emitter model, Fig. 3a).

When $\chi \ll 1$, we can neglect the carrier flux emerging from the cylindrical part of the emitter and consider only the injection from the flat portion, assuming that the emitter is concentrated in the $z=0$ plane (the horizontal emitter model, Fig. 3b).

It is clearly advisable to use the former emitter model to calculate structures with heavy doping of the emitter region and the latter model to calculate structures with shallow doping. Assigning the values $\chi=10$ and 0.1 in the former and latter cases, respectively, to consider concrete cases and using Eq. (6), we find that the vertical emitter model can be used for structures with a diffusion emitter when the accelerating fields $E_0>100$ V/cm (we selected $r \approx 3.8$ μm) and the horizontal emitter model can be used for structures with an implanted emitter when $E_0<100$ V/cm ($r \approx 0.23$ μm).

[1] I. M. Vikulin, M. A. Glauberman, and N. A. Kanishcheva, Fiz. Tekh. Poluprovodn. **11**, 645 (1977) [Sov. Phys. Semicond. **11**, 377 (1977)].
[2] I. M. Vikulin, M. A. Glauberman, G. A. Egiazaryan *et al.*, Fiz. Tekh. Poluprovodn. **15**, 479 (1981) [Sov. Phys. Semicond. **15**, 274 (1981)].
[3] L. W. Davies and M. S. Wells, Proc. IREE Austr. (6), 235 (1971).
[4] W. Allegretto, A. Nathan, and H. P. Baltes, in *NACECODE V, Proceedings of the 5th International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits*, Ireland, 1987, pp. 87–92.
[5] I. M. Vikulin, M. A. Glauberman, and V. V. Egorov, Elektron. Tekh., Ser. 2: Poluprovodn. Prib. (1), 9 (1990).
[6] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. Wiley-Interscience, New York (1969) [Russ. transl., Mir, Moscow (1984), Vol. 1].
[7] S. V. Gumenyuk, in *Abstracts of Reports to the Scientific-Technical Conference ''Sensors Based on Microelectronics Technologies''* [in Russian], Moscow, 1989, pp. 85–87.
[8] V. S. Lysenko, R. N. Litovskii, Ch. S. Roumenin, and N. D. Smirnov, Rev. Phys. Appl. **18**, 87 (1983).

Translated by P. Shelnitz

Formation of multilayer strained-layer heterostructures by liquid epitaxy. I. Theoretical aspects of the problem and mathematical model

R. Kh. Akchurin

M. V. Lomonosov Moscow State Academy of Fine Chemical Technology, 117571 Moscow, Russia

D. V. Komarov

Institute of Chemical Problems in Microelectronics, 117571 Moscow, Russia
(Submitted July 13, 1995; resubmitted May 14, 1996)
Zh. Tekh. Fiz. **67**, 42–49 (July 1997)

Problems concerned with the formation of multilayer strained-layer heterostructures by ''capillary'' liquid-phase epitaxy with forced hydraulic replacement of the solutions in the growth channel are analyzed. It is shown for short contact times between the solutions and the crystallization surface that the character of their flow in the channel plays an important role in the achievement of uniformity in the physical characteristics of the layers grown. Theoretical estimates of the hydrodynamic stability of solutions moving in narrow channels are performed for several III–V systems. A mathematical model, which permits simulation of the conditions under which strained-layer heterostructures are fabricated, is developed. It takes into account diffusive and convective mass transport in the liquid for various flow regimes in the capillary and the displacement of the heterogeneous equilibria in the system under the influence of elastic stresses. [S1063-7842(97)00907-0]

INTRODUCTION

The fabrication of multilayer strained-layer heterostructures is one of the effective methods for controlling the fundamental physical parameters of the semiconductor materials used, in particular, in optoelectronics.[1–3] The thicknesses of the epilayers forming such heterostructures must not exceed the limit corresponding to the transition from the elastically stressed state to the relaxed state with the formation of misfit dislocations and thus lie in the range from several to hundreds of nanometers in most cases. Ultrathin epilayers are presently widely obtained by metalorganic vapor-phase epitaxy, as well as by molecular-beam epitaxy. The development of processes for depositing ultrathin layers from a liquid phase has run into a number of serious problems. Nevertheless, the possibilities of the successful growth of such layers by liquid-phase epitaxy (LPE) have been demonstrated in numerous experimental studies (see, for example, Refs. 4–6). However, some theoretical aspects of LPE pertaining to the deposition of ultrathin epilayers and the fabrication of multilayer strained-layer heterostructures have not yet been adequately developed.

The obvious areas for developing LPE for use in the solution of the problems indicated are confined to lowering the growth temperature, shortening the growth time, and reducing the volume of the liquid phase used for epitaxy. The most effective way to realize such processes is epitaxy from solutions placed in the thin gap between two parallel substrates with replacement of the solutions by forcing them through capillary channels (capillary epitaxy[4]). The problems of the uniformity and reproducibility of the composition and thickness of the epilayers over their area and across their thickness are directly related to the hydrodynamic conditions in the solution and the influence of convective phenomena on the mass-transfer processes accompanying heteroepitaxy. The need to maintain elastic stresses in the epilayers requires consideration of their contribution to the displacement of the heterogeneous equilibria in the systems under consideration, and the absence of thermodynamic equilibrium between the contacting solid and liquid phases during heteroepitaxy requires the use of preliminarily supercooled solutions to preclude local dissolution of the preceding layers. In the case of low-temperature epitaxy preliminary supercooling of the solution is also conducive to more successful formation of a continuous layer under the conditions of short contact times between the liquid and solid phases.

The purpose of the present work was to develop a mathematical model, which would make it possible to find the optimum (according to the criteria for achieving the required characteristics of the epilayers) conditions for the reproducible fabrication of multilayer heterostructures with consideration of the phenomena just described.

DESCRIPTION OF THE PROCESS

Let us consider the scheme of the process leading to the formation of a multilayer structure consisting of alternating layers of the compound AB and the solid solution $A_{1-x-y}B_xC_yD$ on substrates of AB. We assume that the epitaxial deposition process takes place under isothermal conditions on two parallel, closely arranged substrates (Fig. 1), the gap between which (the growth channel) is successively filled by preliminarily supercooled solutions of appropriate composition having the same temperature T. The amount of initial supercooling of the solutions must be sufficient to prevent local dissolution of the substrates or the previously grown epilayers over the entire contact surface (in low-temperature epitaxy this condition can be satisfied for most semiconductor systems). The solutions are replaced under the action of the hydraulic pressure created by pumping them

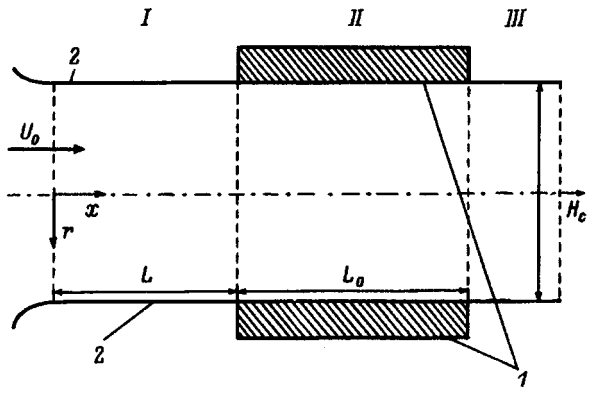

FIG. 1. Schematic representation of epitaxy in a growth channel: *1* — substrates, *2* — channel wall. Here L is the length of the segment for hydrodynamic stabilization, L_0 is the length of the substrates, U_0 is the flow velocity of the liquid phase, and H_c is the thickness of the growth channel.

through a capillary channel (forced convection), whose cross section is similar in size and shape to the growth channel. Contact between the supersaturated solution and the substrate surfaces results in elimination of the supersaturation in its near-surface regions and the growth of epilayers. A stream of the liquid phase forms in region *I*, and epitaxial growth occurs in region *II*. The latter can take place either in a relaxation regime with a cessation of flow or in a regime of continuous flow of the solution in the channel. The used (spent) solution enters region *III*, in which the processes taking place no longer have any influence on epitaxy. Let us assume that the growth channel is symmetric relative to the plane which is perpendicular to the normal between the substrates and passes through its midpoint. This makes it possible to consider the phenomena taking place in it only in the portion between the substrate and the symmetry plane. As a whole, epitaxial deposition can be described by the following set of processes: a) supply of the original solution to the growth channel by forced convection in accordance with the assigned temporal parameters; b) mass transport of the components of the material being deposited from the bulk of the solution to the substrates; c) epitaxy proper, i.e., passage of the solute onto the substrate surfaces; and d) removal of the depleted solution from the growth channel.

The description of the epitaxial process should thus include mass transport in the liquid (natural and forced convection, as well as diffusion) and distribution of the components of the material being deposited between the original liquid and the crystallizing solid phases during epitaxial deposition on the substrate. Let us consider the role of each of these phenomena in the process under consideration.

1) Natural convection. Since the variation of the temperature at the phase boundary during epitaxy can be neglected because of the negligibly small latent heat of the phase transition, the process under consideration can be considered isothermal. Therefore, in evaluating the probability of the appearance of natural convection, only the concentration variations in the solution must be taken into account. The analysis performed in Ref. 7 of the relative contributions of the temperature and concentration gradients to the stability of the liquid phase and the appearance of natural convective fluxes in LPE supports this assumption. If the concentration gradients are much stronger than the thermal gradients, the hydrodynamic stability of a quiescent liquid is determined by the Rayleigh number (Ra_c), which, as applied to our case, can be expressed in the following form

$$\mathrm{Ra}_c = g\beta_c H_c^3 \Delta\tilde{N}/\nu D, \tag{1}$$

where g is the acceleration of gravity, $\beta_c = (\partial\rho/\partial C)\rho^{-1}$ is the coefficient of the concentration dependence of the density of the solution, H_c is the thickness of the growth channel, ΔC is the variation of the concentration of the dissolved component, D is its diffusion coefficient in the liquid phase, and ν is the kinematic viscosity.

Using the similarity criteria (Ra_c/H_s) for binary III–V semiconductor systems in Ref. 7, in a first approximation we can estimate the limiting thicknesses of the growth channels, values above which can lead to the appearance of natural convection in the type of epitaxy under consideration. It should be noted here that since the values of the similarity criteria in Ref. 7 were obtained for growth temperatures that are traditionally employed in the LPE of III–V compounds and the cooling temperature difference was assumed to be 30 K, these values are in need of considerable refinement. The use of low-temperature regimes for epitaxy carried out from preliminarily supercooled solutions under isothermal conditions, should decrease the values of these criteria by factors of 10–15, according to approximate estimates. This is because the amount of initial supercooling of the solutions used in LPE is generally[8] approximately three times smaller than the cooling temperature difference adopted in Ref. 7. Moreover, most III–V systems are characterized by a three- to fivefold relative decrease in $\partial\tilde{N}/\partial T$, which causes a corresponding decrease in ΔC in Eq. (1), in the low-temperature portion of their phase diagrams that is acceptable for solving the problem under consideration.

Plots of the dependence of Ra_c on H_s calculated with consideration of this situation for various III–V systems are presented in Fig. 2. It is seen that the range of growth-channel thicknesses that conform with the absence of natural convection lies at the values $H_s < 200-500$ μm in most cases. However, since a regime of stationary regular convection that causes a significant difference between the growth rates on the upper and lower substrates appears only when $\mathrm{Ra}_c > 10^5$, appreciable natural convection should hardly be expected in the process under consideration even when the calculated values of H_c are exceeded somewhat.

It is noteworthy that the presence of an interface in the capillary channel between two solutions of different composition intended for alternate supply to the growth channel can, in principle, create the conditions for the appearance of another type of natural convection, viz., interfacial convection (for example, Marangoni convection).[9] However, with consideration of the brevity of such contact, the small size of the contact area, and, as a rule, the similarity of the physicochemical characteristics of the solutions used to create heterostructures, in most cases the phenomenon of interfacial convection can be regarded as a minor process and ignored.

2) Forced convection. Under the conditions of brief contact with the substrates the hydrodynamic features of the

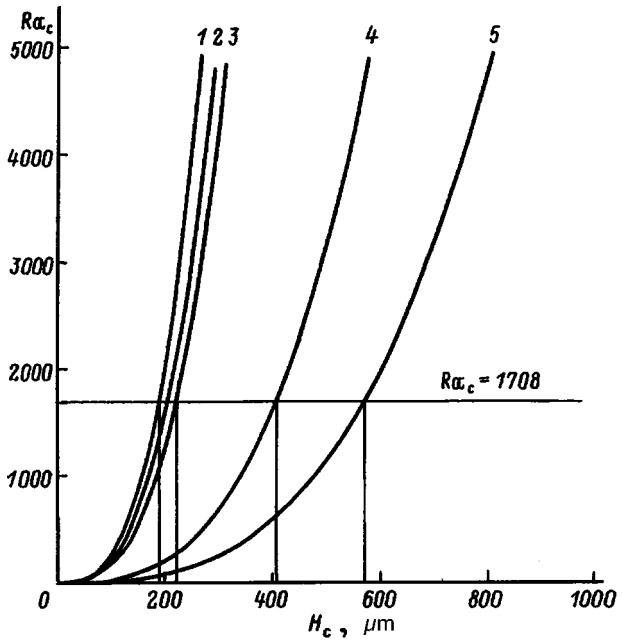

FIG. 2. Values of Ra_c as a function of the thickness of the growth channel for various binary systems: *1* — In–InAs, *2* — In–InSb, *3* — Ga–GaAs, *4* — Ga–GaP, *5* — Ga–GaSb and In–InP.

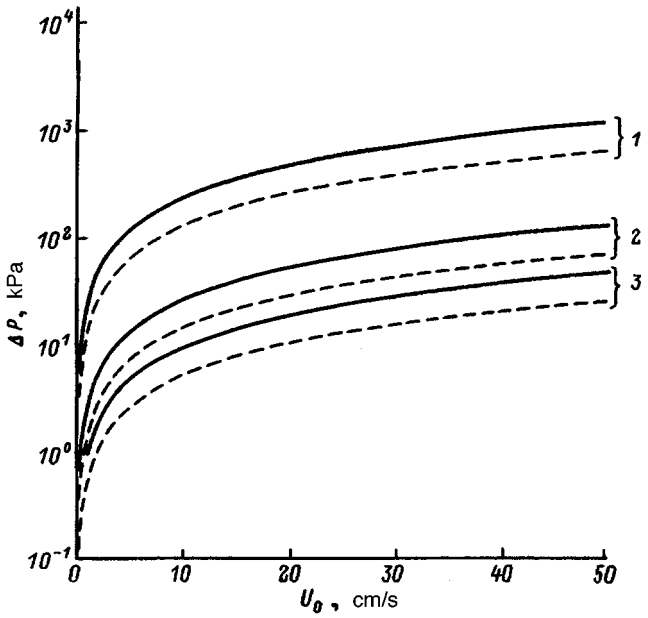

FIG. 4. Relation between the flow velocity of the liquid phase in a growth channel and the difference between the pressures at its entrance and exit. Calculation for $T=573$ K, $L_+, L_0=3$ cm, and $H_c=100$ (*1*), 300 (*2*), and 500 μm (*3*); solvents for III–V: solid lines — Ga; dashed lines — In, Bi.

solution stream entering the growth channel play an exceptionally important role in ensuring the uniformity of the properties of the epilayers being deposited. However, hitherto practically no attention has been focused on these issues in the literature devoted to LPE processes.

The character of the solution stream entering the channel is determined both by the physicochemical parameters of the liquid phase and by the flow velocity. Here the stability criterion is Ra_c, and the dimensionless characteristic of the stream is the Reynolds number ($Re_d = U_0 H_c / \nu$, where U_0 is the mean flow velocity of the solution in the channel). The relationship between these numbers determines the type of flow (Fig. 3). It should be noted that the characteristic values and relations between the criteria were obtained from the general laws of hydromechanics and do not depend on the physical parameters of the system and its components.

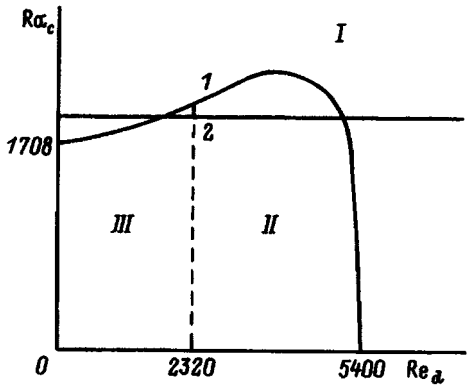

FIG. 3. Stability regions of a moving stream:[10,11] *I*, *II* — regions of nonstationary and stationary turbulent flow, respectively; *III* — region of laminar flow; wave vector: *1* — not equal to 0, *2* — equal to 0.

The ranges of the parameters of the liquid stream that are most acceptable for use in LPE processes are obvious from Fig. 3. For example, the region of nonstationary flow outside of the two curves (region *I*) is unsuitable for carrying out epitaxial processes by its very nature. Processes which require vigorous mixing of the liquid phase, for example, low-temperature epitaxy under the conditions of poor wettability by the substrate melt, can be carried out in the region of stationary turbulent flow (region *II*). In this case the epitaxy channel is like an ideally stirred reactor, since, with the exception of the thin near-wall region of the laminar sublayer, the concentrations of the components are uniformly distributed throughout the volume of the channel.[10] However, when Re is of the order of 2000, even if a relatively narrow epitaxy channel ($H_c \approx 300$ μm) is used, the flow velocities in this region are found to be very large ($U_0 > 100$ cm/s) and lead to unjustified consumption of the materials. Therefore, such a regime can be useful only in exceptional cases. At small values of the criteria (region *III*) the flow is laminar, and the flow profile does not vary with time. This region can be unequivocally recommended for application in LPE. This regime is described using relations from boundary-layer theory.[10,11]

It can be shown that the mean flow velocity of the liquid in our case is given by the relation

$$U_0 = \Delta P H_c^2 / 32 \eta L, \tag{2}$$

where η is the dynamic viscosity of the liquid and ΔP is the difference between the pressures at the entrance to the capillary channel and its exit.

Taking into account the strong degree of dilution of the solutions in low-temperature LPE, for approximate calculations we can set the viscosity of the solution equal to the viscosity of the solvents used. Figure 4 shows plots of

$U_0(\Delta P, H_c)$ for the main solvents employed in the epitaxy of III–V semiconductors (the viscosities of Ga, In, and Bi were calculated for $T=573$ K on the basis of the data in Ref. 12). It is seen that flow velocities sufficiently high to provide for repeated replacement of the solution in the growth channel during short time intervals develop already at $\Delta P=10^5$ Pa in channels of thickness $H_c>200$ μm.

The problem of determining the velocity profile in a channel is not a direct problem of boundary-layer theory, but can be solved by its methods. The so-called internal problem, i.e, flow within a channel, is considered here, and the relations presented below, including the semiempirical relations, are valid for channels of any shape, not just for the pipes for which they were derived.[10] As the thickness of the boundary layers increases (the distance along the normal from the channel wall to the point where the velocity differs by no more than 1% from the velocity in the region of potential flow is taken as the thickness of the boundary layer), the layers formed by the resistance of the channel walls become closer until they merge. The length of the segment for hydrodynamic stabilization is defined as

$$L_0=0.065\sqrt{\mathrm{Re}_d}H_c. \tag{3}$$

If L (Fig. 1) is greater than this quantity, the flow is stabilized, and the distribution of the velocity U over the radius r of the channel (in the present case r is the distance from the channel axis along the normal to the growth surface) has the form

$$U_r=U_0(1-4r^2/H_c^2), \tag{4}$$

which corresponds to a Poiseuille velocity profile. As is seen from Eq. (4), the distribution of the flow velocity in the channel does not depend on the properties of the liquid phase. It is determined explicitly only by the mean flow velocity and the geometry of the channel.

If $L<L_0$, the boundary layers next to the substrates do not merge, and the model can be represented in the form of two independent plates over which the solution stream flows. The thickness of the boundary layer δ on each horizontal wall of the channel along the coordinate x is[11]

$$\delta=4.64x/\sqrt{\mathrm{Re}_x}, \tag{5}$$

where $\mathrm{Re}_x=U_0(L_0+L)/\nu$.

The thickness of the boundary layer increases regularly with increasing distance from the entrance to the channel to its limiting value, which corresponds to the moment when the boundary layers formed by the opposite walls of the channel merge. In this case the distribution of the velocity with respect to the coordinates x and r is described by the expression[11]

$$U_{xy}(\chi)=0.5\chi U_0(3-\chi^2), \tag{6}$$

where $\chi=(H_c-2r)/\delta$.

As the distance from the point of entry into the channel increases, the boundary-layer thickness δ increases, and the flow velocity at a constant distance from the substrate surface decreases.

From the standpoint of practical implementation, this regime provides poorly for uniformity of the distribution of the thickness of the epilayers along the substrate, since the solution at sites more distant from the entrance to the channel is more depleted, and the replacement by fresh solution is slowed. The mathematical model developed in the present work takes into account both profiled and unprofiled flow regimes. Only the alternative in which the velocity profile completes its formation in the segment containing the substrates is excluded, since it leads to high nonuniformity not only of the thickness, but also of the composition of the epilayers.

3) Diffusive mass transport. The driving forces of diffusive mass transport in a solution are the concentration gradients that appear with the onset of crystallization. It is described by diffusion equations written in the differential form

$$D_i\frac{\partial^2 C_i}{\partial r^2}+v\frac{\partial C_i}{\partial r}=\frac{\partial C_i}{\partial t}, \tag{7}$$

where D_i is the diffusion coefficient of component i in the melt, C_i is the concentration of that component, and v is the velocity of the phase boundary along the coordinate r for the one-dimensional case.

The boundary conditions for the system under consideration (Fig. 1) are as follows

$$\left.\frac{\partial C_i}{\partial t}\right|_{r=+0.5H_c,\,t>0}=\left.\frac{\partial C_i}{\partial t}\right|_{r=-0.5H_c,\,t>0}=-f, \tag{8}$$

where f is the rate of epitaxy.

Let D_i be independent of the concentrations of the other components in the solution. Then the number of diffusion equations is one less than the number of components. Assuming that the concentration gradients in planes parallel to the substrate surfaces are negligibly small, we can legitimately confine ourselves to the one-dimensional case with respect to the spatial coordinates.

4) Heterogeneous equilibria. The formation of strained-layer heterostructures initially presumes that there is no relaxation of the stressed state in the epilayers. One consequence of this is the displacement of the heterogeneous equilibria in the system, which is caused by the variation of the activities of the components in the elastically stressed solid phase and requires taking into account the epitaxial process in the model. It is taken into account by introducing additional terms for the ''elastic'' component of the activity coefficients into the equations describing the equilibrium between the phases.[13]

A detailed description of the method for calculating the influence of elastic stresses on equilibrium processes was given in Ref. 14. The system of equations for calculating the activities of the components of a solid solution $A_{1-x-y}B_xC_yD$ that is coherently associated with the substrate in the case under consideration has the form

$$\begin{aligned}RT\ln(\gamma^s_{AB}x^s_{AB})=&RT\ln(1-x-y)+\alpha^s_{AB-AC}x^2+\alpha^s_{AB-AD}y^2\\&+(\alpha^s_{AB-AC}+\alpha^s_{AB-AD}-\alpha^s_{AC-AD})xy\\&+\sigma(a-a_s)(2a_{AB}-a-a_s),\end{aligned}$$

$$RT\ln(\gamma^s_{AC}x^s_{AC}) = RT\ln(x) + \alpha^s_{AB-AC}(1-x-y)^2 + \alpha^s_{AC-AD}y^2 + (\alpha^s_{AB-AC} + \alpha^s_{AC-AD} - \alpha^s_{AB-AD})y(1-x-y) + \sigma(a-a_s)(2a_{AC} - a - a_s),$$

$$RT\ln(\gamma^s_{AD}x^s_{AD}) = RT\ln(y) + \alpha^s_{AB-AD}(1-x-y)^2 + \alpha^s_{AC-AD}x^2 + (\alpha^s_{AB-AD} + \alpha^s_{AC-AD} - \alpha^s_{AB-AC})x(1-x-y) + \sigma(a-a_s)(2a_{AD} - a - a_s), \quad (9)$$

where the γ^s_{ij} are the activity coefficients of the respective binary components and the parameters of their interaction in the solid phase; the α^s are the mole fractions and crystal lattice periods of the respective binary components; a, a_s, and a_{ij} are the lattice periods of the unstrained solid solution, the substrate, and a binary component, respectively.

The parameter σ is defined by the relation[14]

$$\sigma = \frac{3\tilde{N}_{44}(\tilde{N}_{11} + 2\tilde{N}_{12})N_A a}{2(\tilde{N}_{11} + 2\tilde{N}_{12} + 4\tilde{N}_{44})}, \quad (10)$$

where N_A is Avogadro's number, and C_{11}, C_{12}, and C_{44} are the elastic constants of the respective material.

It should be stressed that it is correct to use Eqs. (9) only for epilayers of subcritical thickness.

CONSTRUCTION OF THE MATHEMATICAL MODEL OF THE PROCESS

The combined epitaxial process described can be represented by the system of equations of the concentration in Cartesian coordinates

$$\rho\left(\frac{\partial C_i}{\partial t} + w_k\frac{\partial C_i}{\partial x_k}\right) = -\operatorname{div} I + \mathfrak{N}_i, \quad (11)$$

where ρ is the density of the liquid phase, C_i is the concentration of the ith component in the liquid phase, w_k is the velocity of the medium along the coordinate k, $k \in \{x,y\}$, and $\mathfrak{N}_i$ is the resultant flux due to mass exchange of the ith component with the medium.

The total diffusion flux I of the components is defined as

$$I = -\rho\sum_{k=1}^{n} D_{ik}\nabla C_{ik}, \quad (12)$$

where n is the number of components in the system.

Bearing in mind the unit matrix of the diffusion coefficients and the zero values of the velocity components that are orthogonal to the direction of motion of the stream, we obtain

$$\rho\left(\frac{\partial C_i}{\partial t} + w_x\frac{\partial C_i}{\partial x}\right) = \rho D_i\Delta C_i + \mathfrak{N}_i. \quad (13)$$

The mass flux $\mathfrak{N}_i$ of the ith component to the drain is determined by the epitaxial process and the hydrodynamic transport of the material in the epitaxy channel.

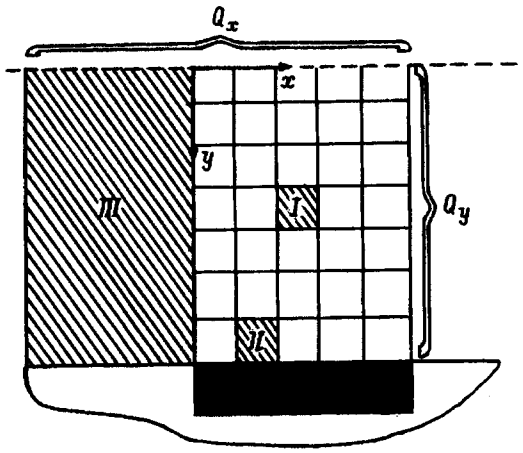

FIG. 5. Introduction of a working net in the space of the growth channel from the substrate to the symmetry axis.

Assigning the corresponding initial and boundary conditions for the concentrations of the components, we obtain the correct boundary-value problem with respect to the concentrations of the input components for the growth process from a homogeneous liquid phase entering the growth channel. For the growth of each succeeding layer we must assign the initial conditions as the distribution of the components in the preceding liquid phase. Because of the complexity of such a formulation of the problem (one of the boundary conditions is the epitaxial process), it is very difficult to obtain an exact analytical solution. Therefore, to construct a model it is helpful to use a finite-difference approximation, which is the simplest route in the present case.

Since there are several independent processes in one system, by dividing the entire time interval into short time periods we can consider the processes independently after a number of such divisions that is sufficiently large to ensure the assigned accuracy of the output parameters. The spatial coordinates can be treated in the same manner.

We introduce a uniform three-dimensional spatiotemporal net with the dimensions $(Q_x; Q_y; Q_\tau)$, where τ is the time coordinate. Let the origin of coordinates be on the symmetry axis of the growth channel (Fig. 5). There is no need to consider the processes with respect to the third spatial coordinate, since the distribution of all the parameters of the system is homogeneous with respect to it.

The introduction of the net divides the entire space of interest into a large set of cells located in three characteristic regions (Fig. 5). In region *I* molecular transport of the material by forced convection and mass transport due to diffusion operate. Region *II* (the surface layer) is subject, in addition, to the mass-draining action of the substrate or the growing epilayer. In region *III*, which is before the entrance to the growth channel, only the effect of molecular transport is manifested. The number of cells Q_y in the net in the direction of the normal to the substrates is determined from the solution of model diffusion problems with complete relaxation of the supersaturation in the liquid phase. The approximation $\Delta y = \sqrt{D\Delta\tau}$ is too rough for the present model; therefore, it is helpful to use the results of direct detailed-

balance calculations of the relaxation to determine the relaxation time.

When the processes taking place in the solution and on the substrate surfaces are systematically modeled, the following approaches are used for each longitudinal cell in the net in accordance with its classification. Diffusion within the growth channel is treated on the basis of the already existing distribution of the concentrations of the components in the liquid phase. Only processes governed by concentration gradients that are orthogonal to the substrates are taken into account. Epitaxy is a general boundary condition of the diffusion equations. The combined solution of the heterogeneous equilibrium and diffusion equations is used to find the concentrations of the components of the solution in the near-boundary cells of the growth channel and in the epilayers deposited, as well as the thicknesses of the layers growing during a step along the time coordinate. The convective transport of the material in the growth channel is considered for two hydrodynamic regimes: profiled flow corresponding to a Poiseuille velocity profile and independent flow at the substrates. A combined modification is not allowed. The possibility of epitaxial deposition is considered for both continuous flow of the solution in the channel (the continuous pumping regime) and for flow with a pause to achieve more complete removal of the supersaturation (the ''relaxation'' regime).

Let us examine the model of the processes which are included in the scheme. Thus, diffusive mass transport is considered only in the direction orthogonal to the substrates for each of the Q_x cells along the substrate. Let i be the spatial label of a cell measuring Δy, and let j be a label, which reflects the number of layers along the time coordinate. Then we can construct an approximation relation for an explicit or implicit scheme that is bilayer with respect to j. The use of the explicit scheme, in which the concentration values in each step $\Delta\tau$ along the time coordinate are expressed explicitly in terms of the corresponding values in the preceding step, imposes restrictions on the length of the step

$$\lambda_{Df}=D_{\hat{e}}\Delta\tau/\Delta y^2<1/2. \tag{14}$$

A violation of this condition influences the convergence of the method. Approximation (11) is absolutely stable when the implicit scheme is used,[15] and unconditional convergence follows from the approximation and the stability.[16] The recurrence relation of such a difference scheme for the concentrations has the form[17]

$$\lambda C_{i-1}^{j+1}-(1+2\lambda)C_i^{j+1}+\lambda C_{i+1}^{j+1}=-C_i^j. \tag{15}$$

The closing conditions on the boundary without consideration of the epitaxial processes, i.e., for drainless flow, are

$$C_{i-1}^{j+1}|_{i\leqslant 1}=C_i^{j+1},$$
$$C_{i+1}^{j+1}|_{i\geqslant Q_y}=C_i^{j+1}. \tag{16}$$

System (15) with boundary conditions (16) and the known initial concentration distribution is a three-diagonal matrix equation, which can be solved by the elimination method,[16] a modification of the Gauss method for systems of three-diagonal equations.

Solving system of equations (15) and (16) for each of the elementary sets of Q_x and Q_y cells (Fig. 5) with variation of the time τ, we obtain the concentration distribution, which varies only as a result of diffusive mass transport.

To calculate the heterogeneous equilibria, the ''direct problem'' is solved, i.e., the parameters of the solid phase and the liquidus temperature are determined from the concentrations of the components in the liquid phase. This process serves as the boundary condition for diffusion influenced by the interface. To describe the compositions of the liquid and solid phases on the phase boundary we use the system of equations of the ''coherent'' phase diagram mentioned above, which is closed by the mass balance equations at the crystallization front[14]

$$x'_{\infty i}=x'_{0i}+N_G(x^s_{0i}-x'_{0i}), \tag{17}$$

where $x'_{\infty i}$, x^s_{0i}, and x'_{0i} are the atomic fractions of the ith component in the respective phases (the subscripts 0 and ∞ refer to the phase boundary and the bulk of the liquid phase, respectively), and N_G is the mole fraction of the deposited solid phase.

In our case we obtain a system of six equations, which is solved by the Newton-Kantorovich linearization method by means of expansion into a Taylor series.[17] The solution is determined in each of the Q_x elementary surface cells.

Since the epitaxy and diffusion processes are assumed to be independent, such an approach is applicable to a discrete surface layer of the liquid phase. The thickness Δy of this layer is determined from the required accuracy of the calculation, and the time for complete relaxation of such a layer, which depends on it, can be expressed as

$$\tau_{\text{relax}}=k\Delta y^2/D^{\min}, \tag{18}$$

where k is a coefficient, which can be determined from the results of the calculation described and $D^{\min}$ is the smallest of the diffusion coefficients of the components of the solution.

This relaxation time is also taken as the step along the time coordinate: $\Delta\tau=\tau_{\text{relax}}$.

The forced convection process used to replace the solutions proceeds in accordance with the assigned temporal parameters. It imposes a restriction on the step Δx along the longitudinal coordinate, which is parallel to the velocity vector. This step must be such that during the time $\Delta\tau$ the elementary cell located at the substrate (i.e., where the flow velocity is smallest) moves over the distance Δx. Then all the remaining cells move over a larger distance. Additivity of the motion operates here, i.e., if the solution does not manage to traverse a whole number of cells during the time $\Delta\tau$, this fact is taken into account in the next temporal step, and the accuracy of the calculation is thereby increased.

The boundary condition for forced convection is the supply of a solution stream in which the concentrations are known *a priori* (region *III* in Fig. 5). Before the beginning of the process, the velocity profile is determined from the geometric configuration of the epitaxy channel using Eq. (3). The distribution of the flow velocity of the melt along the spatial coordinates of the net is constructed from Eq. (4) in the case of profiled flow or from Eq. (6) in the case of un-

profiled flow. Each of the $(Q_x;Q_y)$ elementary cells has its own linear flow velocity, and it depends only on the vertical coordinate in the case of a Poiseuille profile.

CONCLUSIONS

Thus, the relative roles of the various types of mass transport in a liquid have been analyzed in reference to the isothermal modification of ''capillary'' LPE with forced hydraulic replacement of the solutions in the growth channel. The important role of the character of the forced convective fluxes in achieving uniformity of the parameters of epilayers has been demonstrated. Theoretical estimates of the hydrodynamic stability of the flow of solutions moving in narrow channels have been made for several III–V systems, and the ranges of growth channel thicknesses and flow velocities which ensure laminar flow have been established.

A mathematical model, which permits simulation of the conditions for the fabrication of strained-layer heterostructures, has been developed. It takes into account diffusive and convective mass transport in the liquid in various flow regimes in a capillary and the effect of displacement of the heterogeneous equilibria under the influence of elastic stresses.

The model provides for introduction of a spatiotemporal net in the (x,y,τ) coordinate system, which makes it possible to consider the processes taking place in the growth channel with the required accuracy along the coordinates indicated. A finite-difference approximation is used to solve the computational problems.

The processes considered in the model can occur in any sequence. Test measurements of the mutual errors appearing on cyclic permutation of the processes showed that permutation leads to results that are equivalent to within the rounding errors.

[1] G. S. Osborn, J. Vac. Sci. Technol. **2**, 176 (1984).
[2] C. P. Kuo, S. K. Vong, R. M. Cohen *et al.*, J. Appl. Phys. **57**, 5428 (1985).
[3] S. R. Kurtz, R. M. Biefeld, L. R. Dawson *et al.*, Appl. Phys. Lett. **53**, 1961 (1988).
[4] Zh. I. Alferov, V. M. Andreev, A. A. Vodnev *et al.*, Pis'ma Zh. Tekh. Fiz. **12**, 1089 (1986) [Sov. Tech. Phys. Lett. **12**, 450 (1986)].
[5] V. M. Andreev, O. O. Ivent'eva, S. G. Konnikov *et al.*, Pis'ma Zh. Tekh. Fiz. **12**, 533 (1986) [Sov. Tech. Phys. Lett. **12**, 217 (1986)].
[6] G. T. Aĭtieva, V. N. Bessolov, A. T. Denisova *et al.*, Zh. Tekh. Fiz. **56**, 910 (1986) [Sov. Phys. Tech. Phys. **31**, 552 (1986)].
[7] V. I. Polezhaev, A. V. Buné, N. A. Verezub *et al.*, *Simulation of Convective Heat and Mass Transfer on the Basis of the Navier-Stokes Equations* [in Russian], Nauka, Moscow (1987).
[8] V. B. Ufimtsev and R. Kh. Akchurin, *Physicochemical Principles of Liquid-Phase Epitaxy* [in Russian], Metallurgiya, Moscow (1983).
[9] *Hydrodynamics of Interfaces (A Collection of Articles)* [Russian translation], Mir, Moscow (1984).
[10] V. A. Kirpikov and G. N. Shorin, *Introduction to Boundary-Layer Theory* [in Russian], MIKhM, Moscow (1974).
[11] G. Z. Gershuni and E. M. Zhukhovitskiĭ, *Convective Stability of Incompressible Liquids* [in Russian], Nauka, Moscow (1972).
[12] *Landolt-Bornstein*, New Series, Vol. 17d, Springer-Verlag, Berlin (1984).
[13] A. S. Jordan and M. Ilegems, J. Phys. Chem. Sol. **36**, 329 (1975).
[14] V. V. Kuznetsov, P. P. Moskvin, and V. S. Sorokin, *Nonequilibrium Phenomena in the Liquid Heteroepitaxy of Semiconductor Solid Solutions* [in Russian], Metallurgiya, Moscow (1991).
[15] *Mathematical Models and Numerical Methods in Transport Theory* [in Russian], A. A. Samarskiĭ (ed.), Izd. Akad. Nauk BSSR, Minsk (1982).
[16] S. K. Godunov and V. S. Ryabenkii, *Difference Schemes: An Introduction to the Underlying Theory*, North-Holland, Amsterdam–New York (1987).
[17] L. I. Turchak, *Principles of Numerical Methods* [in Russian], Nauka, Moscow (1987).

Translated by P. Shelnitz

Formation of multilayer strained-layer heterostructures by liquid epitaxy. II. Simulation of the fabrication of heterostructures based on indium–arsenic–antimony–bismuth solid solutions

R. Kh. Akchurin

M. V. Lomonosov Moscow State Academy of Fine Chemical Technology, 117571 Moscow, Russia

D. V. Komarov

Institute of Chemical Problems in Microelectronics, 117571 Moscow, Russia
(Submitted July 13, 1995; resubmitted May 14, 1996)
Zh. Tekh. Fiz. **67**, 50–56 (July 1997)

The formation of $InAs_{1-x-y}Sb_xBi_y/InSb$ and $InAs_{1-x-y}Sb_xBi_y/InSb_yBi_y$ strained-layer heterostructures by ''capillary'' LPE is simulated. The laws governing the dependence of the gap width E_g and the thickness d of the epilayers on the conditions of the process are revealed. It is shown that because of the sharp increase in the rate of epitaxial deposition as the LPE temperature is raised, the successful growth of epilayers of subcritical thickness is possible only up to $T<550$ K. The influence of the rate of laminar flow of the liquid in the growth channel in a relaxation regime and in a continuous pumping regime on the uniformity of the distribution of E_g and d in the epitaxial heterostructures is analyzed. Effective combinations of parameters for carrying out the process, which ensure the achievement of $E_g \approx 0.1$ eV (77 K) in the active layers with variable-band-gap layers of minimal thickness, are established.
 [S1063-7842(97)01007-6]

INTRODUCTION

Solid solutions based on indium arsenide antimonide bismuthide ($InAs_{1-x-y}Sb_xBi_y$) are of considerable interest as a material for far-infrared photosensitive devices. Epitaxial layers (epilayers) of $InAs_{1-x-y}Sb_xBi_y$ with a long-wavelength fundamental optical absorption edge at 8.9 μm at 77 K have been obtained for the composition corresponding to $x=0.82$ and $y=0.0018$ on indium antimonide (InSb) substrates by liquid-phase epitaxy.[1] According to the theoretical estimates in Ref. 2, the fabrication of $InAs_{1-x-y}Sb_xBi_y/InSb$ strained-layer heterostructures can bring about a significant decrease in the gap width of the solid solutions and displacement of the fundamental optical absorption edge at 77 K to 12–14 μm.

The purpose of the present work was to evaluate the influence of the principal phenomena accompanying the epitaxial process, as well as the technological parameters of the process, on the output parameters of the epitaxial heterostructures on the basis of the mathematical model described in Ref. 3. Heterostructures containing alternating principal (active) $InAs_{1-x-y}Sb_xBi_y$ layers and auxiliary (buffer) layers of InSb or $InSb_{1-y}Bi_y$ are considered.

The model has the form of a ''direct substitution problem,'' i.e., the assigned values are the input parameters, and thus corresponds to the processes in reality. A solution of the inverse problem might have advanced the search for the best parameters, but in view of its great complexity such an approach is not the optimum method for solving the problem as a whole. This, however, does not rule out the existence of inverse frequency problems.

The assigned parameters of the process generally include: a) thermodynamic parameters, viz., the temperature of the process, the amount of supercooling in the system, the initial concentrations of the components in the liquid phases (solutions), and the substrate material; b) hydrodynamic parameters, viz., the geometry of the growth space and the velocity profile of the solutions in the growth channel; c) the transient times of convective mass transport and nonconvective relaxation for the growth of the active and buffer layers. The output data include the spatial distribution of the physical parameters of interest to us (in the present case they are the thicknesses of the epilayers and the gap widths of the solid solutions) in the multilayer heterostructure.

ORIGINAL ASSUMPTIONS

a) Preliminarily supercooled solutions are used in the process. The amount of initial supercooling is sufficient for ensuring that the driving forces of crystallization exceed the driving forces of dissolution[4] over the entire contact surface between the liquid and solid phases.

b) The rate-limiting step of the epitaxial process is the supply of the components to the interface. Although the boundary of the kinetic region, in which the main rate-limiting role is played by the processes on the growth surface, can be established only on the basis of experimental investigations, the basis for this assumption is the use of solutions having a sufficiently high degree of initial supersaturation in the process.

c) Diffusive mass transport takes place only in the direction perpendicular to the substrates. Its other orthogonal component is negligibly small throughout the epitaxy channel, with the exception of the thin near-wall region, where a difference in the diffusion fluxes along the substrate is possible. Therefore, allowance for the longitudinal component of the diffusion can be reduced to a smoothing of the calcu-

lated data. The diffusion coefficients are concentration-independent and, therefore, remain constant over the course of the entire isothermal process.

d) The only component of molecular transport in the system under consideration is the transport of the material by forced convection, whose rate vector is parallel to the substrates. Actually this implies the neglect of interfacial convection and the local phenomena caused by the nonuniformity and imperfect symmetry of the epitaxy channel: in particular, the phenomena appearing on the substrate edges are not considered (it is assumed that the channel walls and the working surfaces of the substrate are smooth and parallel to their respective partners).

e) The distance between the substrates does not vary with time, i.e., the epilayer is assumed to be infinitely thin in comparison to the thickness of the liquid phase. It should be noted that this assumption can be less than fully correct when a large number of epilayers must be created in a heterostructure. In such a case the motion of the phase boundary in the diffusion description and the corresponding variation of the velocity profiles must be taken into account. This can be done quite simply when it is assumed that the thickness of the epilayers is uniform along the substrate.

f) The flow of the solution at the entrance is laminar. This condition is always satisfied for reasonable values of the velocity of the liquid phase.[3]

g) The flow in the channel is stabilized, i.e., when the solution enters the space between the substrates a Poiseuille velocity profile is formed. This is not a restriction of the calculation, since the model contains alternative algorithms for distributing the velocity;[3] however, a dependence of the profile on the longitudinal coordinate causes higher nonuniformity in the distribution of the output parameters along the substrate, which is undesirable in the case of the growth of thin epitaxial structures.

h) The processes occurring in the liquid phase which has passed through the growth channel do not influence the output parameters of the layers deposited.

CHOOSING THE WORKING REGION OF MODEL PARAMETERS

To ascertain the character of the dependence of the output parameters of the epilayers on the input parameters, a working region must be chosen in the multidimensional space of the data on the basis of preliminary calculations.

The limits of the temperature range selected were 523 and 573 K. The upper limit is stipulated by the strong increase in the growth rate of the active and buffer layers as the growth temperature is raised and, therefore, by the equivalent decrease in the precision of the method. The lower limit is specified by the significant decrease in the arsenic concentration in the solution, which has the same consequences. In addition, the appearance of kinetic restrictions on the growth surface, as well as restrictions associated with expansion of the immiscibility region in the solid phase, is possible at low temperatures.[5]

The choice of the amount of initial supercooling of the solution reduces to determining the minimum value needed to prevent local dissolution of the substrates or the preceding epilayers, which are coherently associated with the growing layers. For solid phase compositions that produce layers with the required output parameters it has a value of about 10 K at $T=673$ K and is smaller at lower temperatures.[6] The permissible amount of supercooling has an upper bound of the order of 15–20 K due to the increase in the probability of spontaneous crystallization in the bulk of the solution, as well as the large growth rates of the epilayers.

The (111) orientation of the substrate is most stable toward random fluctuations of the crystallization conditions. A small normal and a large tangential component of the growth rate are then ensured, which have a favorable effect on the quality of the surface of the strained layers.

The dimension of the substrate in the direction of motion of the solutions (the length of the substrate) plays an important role in the process under consideration, since the probability of the appearance of heterogeneous distributions of the output parameters of the epilayers is highest specifically along the substrate. A length of 1 cm was chosen here as the model value.

The thickness of the epitaxy channel is restricted from above by the possibility of the appearance of free convection[3] and was assumed to be equal to 500 μm. If all the other parameters are constant, the effective volume of the solution increases as the thickness decreases. Therefore, it is technologically advantageous to carry out the process in narrow capillaries; however, when the process is carried out in practice, it is more difficult to ensure a parallel arrangement of the substrates in this case. A value of 100 μm was taken as the lower limit of the capillary thickness, although systems with a distance between the substrates equal to 70 (Ref. 7) and even 50 μm (Ref. 8) have been described in individual experiments.

The mean flow velocity of the liquid phase at the entrance to the growth channel is formally restricted only from above, but in practice only from below. When the velocity of the incoming stream is high, it undergoes turbulization, and undesirable inertial phenomena appear. However, the values of such critical velocities are very high: for channels with a thickness of 500 μm they are in the vicinity of 200 cm/s, and they increase even more with decreasing channel thickness.[3] Such high velocities lead to an extremely low efficiency for the process and can hardly be employed in practice. On the other hand, as will be shown below, the flow velocity must be sufficient to ensure a homogeneous distribution of the output parameters along the substrate. Moderate velocity values in the range 2–100 cm/s were used in the numerical experiments. The appearance of nonuniformity along the substrate at smaller velocities and a low deposition efficiency at higher velocities demonstrated that the range of optimum velocities is completely covered.

The time of the epitaxial process is restricted from below by the accuracy and reproducibility of creating an assigned time interval and by the response time of the instruments through the relationship of the ratio between the thicknesses of the active and buffer layers to the required gap width of the solid solutions in the active regions of the heterostructure. From practical considerations it should be less than 1 s. The upper bound is set by the need to grow thick pseudo-

TABLE I. Thermodynamic parameters of binary systems.[5,9,10]

System	Melting point, K	Entropy of fusion, J/mol·K	Interaction parameters, J/mol
In–As	1215	60.80	$1717-26.63T$
In–Sb	798	59.95	$13565-49.45T$
In–Bi	383	37.67	$8680-18.5T$
As–Sb	–	–	3140
As–Bi	–	–	6985
Sb–Bi	–	–	2700
InAs–InSb	–	–	$6450+3.85T$
InAs–InBi	–	–	46515
InSb–InBi	–	–	12700

TABLE II. Crystal lattice periods (a), coefficients of thermal expansion (α), and elastic constants (C_{11}, C_{12}, and C_{44}) of binary compounds.[9,10]

Compound	a, nm	$\alpha\cdot10^6$, K^{-1}	C_{11}, GPa	C_{12}, GPa	C_{44}, GPa
InAs	0.60584	5.04	83.3	45.3	39.6
InSb	0.647937	5.19	62.7	36.7	30.2
InBi	0.667 [a]	–	–	–	–

[a]The period relative to that of the sphalerite lattice[2] was used. The values of the elastic constants and the lattice periods are given for a growth temperature of 573 K.

morphic epilayers at comparatively high flow velocities.

The physicochemical data used in the calculation are presented in Tables I and II.

DEPENDENCE OF THE DISTRIBUTION OF THE OUTPUT PARAMETERS OF THE EPITAXIAL LAYERS ON THE LPE CONDITIONS

Calculations were performed according to the following scheme to analyze the dependence of the epitaxy results on the initial data and parameters. A certain intermediate point, for which the results of the calculation were at least acceptable in the context of the problem posed, was selected in the parameter space. The corresponding set of parameters defined a basic calculation regime. Then the changes in the outcome of the problem following local variation of each parameter at constant values of the other parameters were analyzed. An analysis of the data on the nature of the influence of each parameter revealed the appropriate relations between them in the context of the solution of the problem posed.

The simulation of the formation of strain-layer $InAs_{1-x-y}Sb_xBi_y/InSb$ and $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ heterostructures was aimed at reaching a long-wavelength fundamental optical absorption edge of the active layers corresponding to 12 μm at 77 K. We note that the solution of this problem would make it possible to obtain wider-gap materials. The data in Ref. 2 on the critical thicknesses of the epilayers of the required composition and the dependence of the gap width of the active layers on the concentration and geometric parameters of the heterostructures were used in the calculation.

The assigned calculated parameters of the basic calculation regimes for the heterostructures indicated are presented in Table III. Relaxation-free epitaxy regimes (regimes with continuous pumping of the solution through the growth channel) were considered in both cases. The distribution of $E_g(d)$ for these regimes is presented in Fig. 1.

Only a slight difference in the distribution of the output parameters between the first and ith heterojunction of the multilayer structure was revealed during the calculation; therefore, the resultant distribution of the gap width across the thickness of the epilayers at the beginning and end of the

TABLE III. Data on the basic calculation regimes.

Parameter	$InAs_{1-x-y}Sb_xBi_y$/InSb system	$InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ system
Growth temperature T, K	523	
Growth channel thickness H_c, μm	300	
Flow velocity at channel entrance U_0, cm/s	20	
Diffusion coefficient D, cm^2/s:		
for As	1.0×10^{-4}	
for Sb and Bi	8.0×10^{-5}	
Thickness ratio of the active and buffer layers	0.32	0.52
Solution pumping time for formation of the active layer, s	1	
Solution pumping time for formation of the buffer layer, s	0.25	1.04
Initial concentrations in the solution, atomic fraction: In–As–Sb–Bi:		
As	0.00005	
Sb	0.00720	
Bi	0.49270	
In–Sb–[Bi]:		
Sb	0.0235	0.0092
Bi	0	0.4420
Critical thickness of the active layer, nm	140.2	
Plot of the resultant distribution of E_g across the thickness d of the heterostructure	Fig. 1a	Fig. 1b

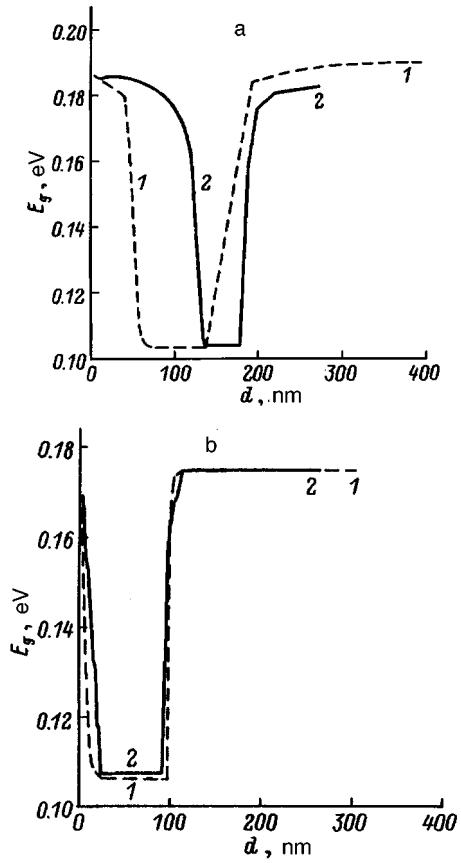

FIG. 1. Distribution of E_g (77 K) across the thickness of $InAs_{1-x-y}Sb_xBi_{yy}/InSb$ (a) and $InAs_{1-x-y}Sb_xBi_y/Sb_{1-y}Bi_y$ (b) heterostructures for the basic calculation regimes at the beginning (*1*) and end (*2*) of the substrate.

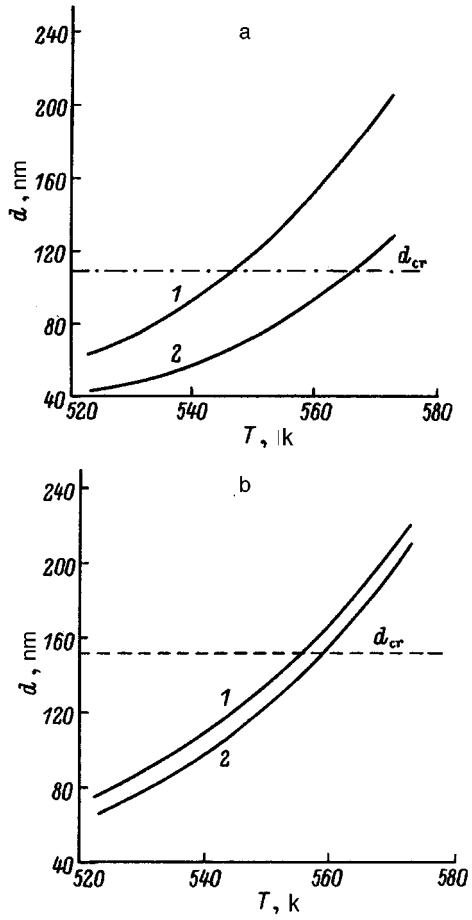

FIG. 2. Dependence of the thickness of the active layer on the LPE temperature for $InAs_{1-x-y}Sb_xBi_y/InSb$ (a) and $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ (b) heterostructures at the beginning (*1*) and end (*2*) of the substrate.

substrate is shown for only one cycle of the process.

It is seen that the results of the calculations for the two regimes differ significantly. In one case (Fig. 1a) considerable nonuniformity of the values of the output parameters of the layers along the substrate is observed. The variation of the depth of the active layer in the portion of the substrate farthest from the entrance to the growth channel is attributed to the small near-surface longitudinal rate of molecular transport. Consequently, at the end of the substrate the layer grows from the onset exclusively as a result of diffusion from the central part of the channel, where the rate of molecular transport is several orders of magnitude greater. The ratio between the thicknesses of the epilayers at the opposite ends of the substrate amounts to ≈ 0.6, which can be attributed to the small degree of relaxation of the liquid phase. For the same reason there is some variation in E_g, which is caused by the influence of the ratio between the thicknesses of the active and buffer layers on the variation of the gap width due to the pseudomorphic effect.

In the other case (Fig. 1b) the uniformity of the thickness of the active epilayers increases significantly, and the thickness of the variable-band-gap layers decreases. The obvious reason for this is the sharp decrease in the absolute value of the concentration gradient of the components in the liquid phase when the solution is replaced because of the similarity in composition (Table III). It should be noted that the physical parameters of the renewed solutions, which are maintained practically constant, also sharply decrease the probability of the appearance of interfacial convection.[3]

The calculations showed that an increase in the temperature of the process leads to a sharp increase in the growth rate of the epilayers. For example, in the case of the formation of $InAs_{1-x-y}Sb_xBi_y/InSb$ structures, the thickness of the active layers reaches the critical value already at 550 K even after very short contact times (Fig. 2). The nonuniformity of the total thickness of the epilayers along the substrate increases significantly (Fig. 3). In the case of the formation of $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ heterostructures the uniformity of the deposition process along the substrate is considerably higher, but the dependence of the growth rate on the temperature remains practically unchanged.

Thus, it is clear from the calculated data presented that it is wise to fabricate the required strained-layer heterostructures at temperatures no higher than 530–550 K using $InSb_{1-y}Bi_y$ solid solutions as the material of the buffer epilayers. Therefore, in the further discussion we shall confine ourselves to this group of heterostructures. It should be noted here that lowering the temperature of the process to a level below 500 K could hardly be beneficial both because of the possibility of the kinetic limitation of epitaxial growth[11] and because of the poorer wetting of the substrate. In addition,

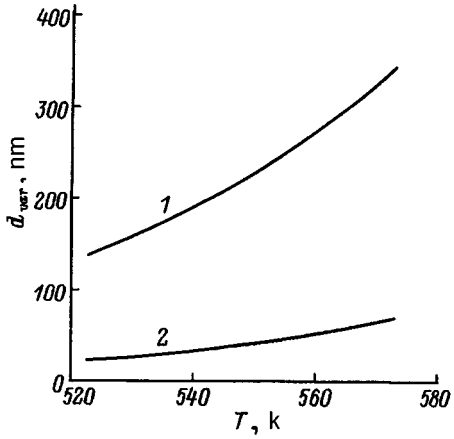

FIG. 3. Dependence of the nonuniformity of the thickness of the active layer along the substrate on the temperature of the process for $InAs_{1-x-y}Sb_xBi_y/InSb$ (*1*) and $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ (*2*) heterostructures.

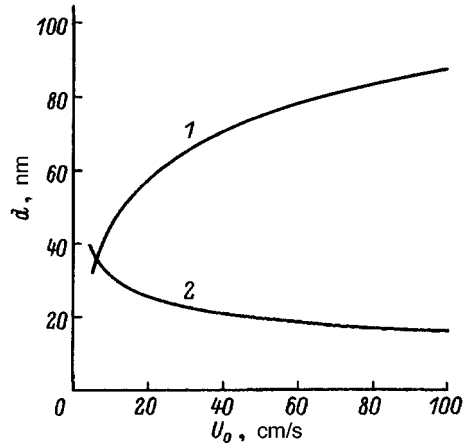

FIG. 5. Dependence of the thickness of the active layer of an $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ heterostructure at the end of the substrate (*1*) and the nonuniformity of its thickness along the substrate (*2*) on the flow velocity of the liquid phase at the entrance to the growth channel.

the absence of reliable data on the parameters of the interaction at low temperatures lowers the reliability of the calculation results.

The dependence of the thickness of the epilayers on the thickness of the growth channel for the continuous pumping regime is illustrated in Fig. 4. The increase in the thickness of the layers as the distance between the substrates decreases is caused by the alteration of the flow velocity profile in the channel. As the thickness of the capillary decreases, the linear rate of molecular transport of the solution at the same relative distance from the interface increases, the absolute values of the concentration gradients of the components in the liquid phase increase, and the total rate of epitaxial deposition consequently increases. The decrease in the nonuniformity of the thickness of the layers along the substrate is expected, since this nonuniformity is caused by the small rates of molecular transport near the crystallization surface. From the practical standpoint, a decrease in the thickness of the growth channel also promotes an increase in the effective volume of the solution, since the epitaxial process is characterized by a higher deposition efficiency.

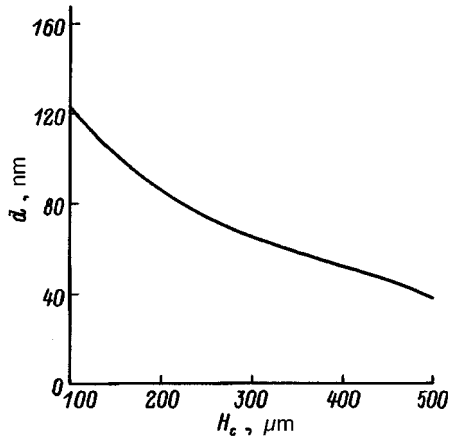

FIG. 4. Dependence of the thickness of the active layer of an $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ heterostructure on the thickness of the growth channel.

The variation of the mean flow velocity at the entrance to the growth channel is also noticeably reflected in the thickness of the active layers (Fig. 5, curve *1*) and in the uniformity of its distribution along the substrate (Fig. 5, curve *2*).

It is noteworthy that the distribution of the output parameters across the thickness of the heterostructures has a sharper form at high flow velocities. This is due to the dominance of the convective transport over the weaker diffusive transport, which consequently leads to comparatively small thicknesses for the variable-band-gap layers. At the same time, the thickness of the variable-band-gap layers varies only slightly as the volume rate of flow of the solutions, which is proportional to the mean flow velocity, increases. Therefore, in the context of the problem under consideration the smallest flow velocity that provides for the assigned homogeneity of the distribution of the output parameters of the epilayer should be chosen.

Processes in which epitaxy takes place in a regime with continuous flow of the solutions in the channel have not been considered hitherto. The simulation of relaxation-free processes as simpler process will help us to better describe the character of the variation of the output parameters as the input parameters are varied. At the same time, as the calculations showed, these processes are plagued by several significant deficiencies. They are characterized, in particular, by fairly high nonuniformity of the thickness of the active and buffer epilayers along the substrate and, as a consequence, by a spread of values of E_g. The more complicated relaxation processes with stopping of the flow (a pause) after the replacement of each successive portion of the solution in the growth channel for additional removal of the supersaturation have their own special features. We shall now examine some of them, calling the duration of the pause the relaxation time by convention.

Let the conditions of the process be distinguished from the basic regime for the formation of $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ structures only by the presence of a period for relaxation of the supersaturated solution

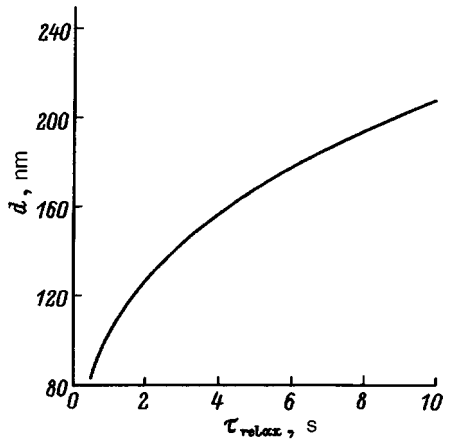

FIG. 6. Dependence of the thickness of the active layer of an $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ heterostructure on the relaxation time τ_{relax}.

after it is supplied to the growth channel. The general form of the dependence of the thickness of the active epilayer on the relaxation time is shown in Fig. 6. It follows from it, in particular, that in the present case this time is restricted from above by the possibility of exceeding the critical thickness of the active layer. Plots of $E_g(d)$ for the maximum permissible relaxation time (≈ 3 s) are presented in Fig. 7 (curves *1* and *2*). When the active layers are grown according to a scheme with complete relaxation and maintenance of the elastic stresses in them (with consideration of the approximate nature of the calculations of d_{cr}), a characteristic ''well'' is observed on the calculated plot of $E_g(d)$ (Fig. 7, curve *3*). It is seen that there is a decrease in E_g at the very beginning of

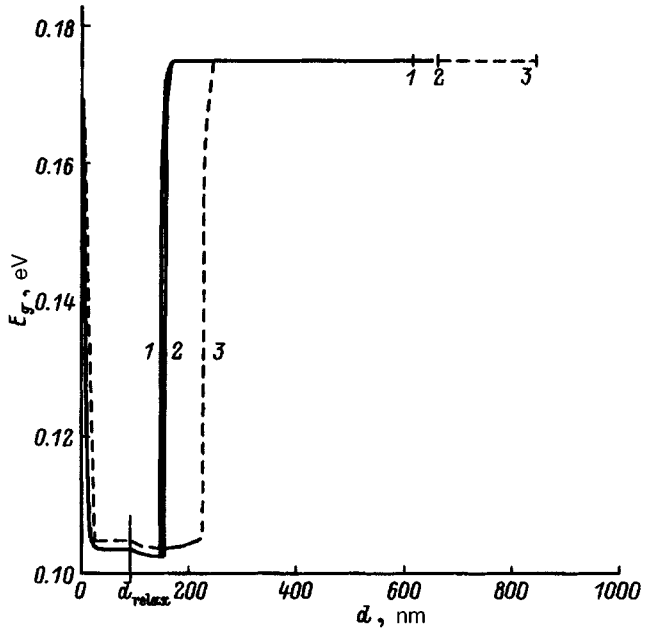

FIG. 7. Distribution of E_g (77 K) across the thickness of an $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ heterostructure in the relaxation epitaxy regime: $\tau_{relax}=3$ s (*1*, *2*) and $\tau_{relax}=10$ s (*3*). Curves *1* and *2* refer to the end and beginning of the substrate, respectively; when $\tau_{relax}=10$ s, the curves for the beginning and end of the substrate coincide; d_{relax} is the beginning of the relaxation period.

the relaxation period as a result of enrichment of the solid phase with arsenic due to the higher rate of diffusion of its atoms in the solution from the central region of the growth channel. As the liquid phase becomes depleted of arsenic, its concentration in the solid phase decreases with a resultant increase in E_g, and the growth rate of the layer slows. As a result, the growth of the layers according to the complete relaxation scheme leads to the formation of the characteristic ''well'' on the $E_g(d)$ curve.

It should be noted that the use of relaxation regimes provides for a considerably more homogeneous distribution of the output parameters along the substrate even when $InAs_{1-x-y}Sb_xBi_y/InSb$ structures are formed. Another positive feature of these processes is their high efficiency with respect to the materials being consumed. On the other hand, the relaxation regime requires an increase in the contact time of the liquid phase with the crystallization surface, which (with consideration of the restrictions with respect to the maximum permissible thickness of the epilayers) can create definite difficulties in their practical implementation.

One of the ways to decrease the thickness of the epilayers in this case may be to lower the flow velocity at the entrance to the channel. The inhomogeneities of the thicknesses along the substrate appearing in this case (Fig. 5) are smoothed to a considerable degree in the relaxation regime of the epitaxial process, and the nonuniformity along the substrate becomes negligibly small when $InSb_{1-y}Bi_y$ is employed as the material of the buffer layers. It is also expedient to carry out the epitaxial growth in a regime of incomplete relaxation by selecting a pause duration which ensures that the layer achieves an assigned thickness smaller than the critical value.

CONCLUSIONS

The laws governing the variation of the physical parameters of the epilayers (E_g and d) on the conditions of the formation process have been revealed and analyzed as a result of the simulation of the formation of $InAs_{1-x-y}Sb_xBi_y/InSb$ and $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ strained-layer heterostructures by LPE.

It has been shown that as a result of the sharp increase in the rate of epitaxial deposition with increasing growth temperature, the successful growth of the epilayers of subcritical thickness is possible only at $T<550$ K.

An increase in the rate of laminar flow of the liquid phase in the growth channel and the use of relaxation growth regimes promote a more homogeneous distribution of the thickness of the epilayers along the structure. On the other hand, to increase the degree of utilization of the solutions it is best to restrict the thickness of the growth channel to 100–300 μm.

The formation of $InAs_{1-x-y}Sb_xBi_y/InSb_{1-y}Bi_y$ heterostructures is preferable for achieving uniformity of the physical parameters of the active layers. A significant decrease in the thickness of the variable-band-gap layers is provided in this case, as in the case of an increase in the rate of laminar flow of the liquid in the growth channel.

In conclusion, it should be noted that the main problems in the practical implementation of the process considered are

the phenomena associated with the poor wettability of the substrate surface at low temperatures and the need to precisely control several technological parameters (the concentrations of the dissolved components in the liquid phase and the contact time with the substrate) due to their small absolute values.

We thank V. A. Zhegalin and T. V. Sakharova for their assistance in writing the programs for the individual stages of the calculation. This work was performed with the support of the State Committee of the Russian Federation for Higher Education (Grant No. 460-6 of 1994 for ''Physicochemical Research in Chemical Technologies'').

[1] R. Kh. Akchurin and T. V. Sakharova, Pis'ma Zh. Tekh. Fiz. **18**(10), 16 (1992) [Sov. Tech. Phys. Lett. **18**, 307 (1992)].
[2] R. Kh. Akchurin and O. V. Akimov, Fiz. Tekh. Poluprovodn. **29**, 362 (1995) [Semiconductors **29**, 183 (1995)].
[3] R. Kh. Akchurin and D. V. Komarov, Zh. Tekh. Fiz. **67** (7), 42 (1997) [Tech. Phys. **42**, 755 (1997)].
[4] R. Kh. Akchurin, in *Physicochemical Processes in Microelectronics* [in Russian], MITKhT, Moscow (1990), pp. 318–342.
[5] R. Kh. Akchurin, T. V. Sakharova, and V. A. Zhegalin, Izv. Vyssh. Uchebn. Zaved. Tsvetn. Metall. (7), 23 (1994).
[6] R. Kh. Akchurin, V. A. Zhegalin, and T. V. Sakharova, Izv. Ross. Akad. Nauk Ser. Neorg. Mater. (1995) (in press).
[7] G. T. Aĭtieva, V. N. Bessolov, S. E. Klimenko *et al.*, Pis'ma Zh. Tekh. Fiz. **11**, 465 (1985) [Sov. Tech. Phys. Lett. **11**, 192 (1985)].
[8] Yu. Yu. Abdurakhmanov, S. E. Klimenko, and V. E. Korsukov, Pis'ma Zh. Tekh. Fiz. **8**, 762 (1982) [Sov. Tech. Phys. Lett. **8**, 331 (1982)].
[9] S. S. Strel'chenko and V. V. Lebedev, *III–V Compounds (A Handbook)* [in Russian], Metallurgiya, Moscow (1984).
[10] *Physicochemical Properties of Semiconductor Materials (A Handbook)* [in Russian], A. V. Novoselova and V. B. Lazarev (eds.), Nauka, Moscow (1979).
[11] V. B. Ufimtsev and R. Kh. Akchurin, *Physicochemical Principles of Liquid-Phase Epitaxy* [in Russian], Metallurgiya, Moscow (1983).

Translated by P. Shelnitz

Acoustooptic 2×2 switch for radiation with different wavelengths as an element of a fiber-optic gyroscope

V. M. Kotov

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 141120 Fryazino, Russia
(Submitted September 6, 1995; resubmitted May 28, 1996)
Zh. Tekh. Fiz. **67**, 57–62 (July 1997)

Planar acoustooptic 2×2 directional couplers that switch optical rays with different wavelengths are considered. A method for calculating the angular and frequency characteristics of such switches is developed, and the parameters of a switch based on a planar Ti-$LiNbO_3$ structure are calculated. Experiments employing bulk acoustooptic diffraction in TeO_2 are performed, confirming the basic theoretical assumptions. It is shown experimentally that the best conditions for switching optical rays directed into optical fibers can be provided nearly always by varying the diffraction angles and the frequencies of the sound waves. The functional possibilities of the 2×2 directional coupler investigated in a fiber-optic gyroscope are described.
 [S1063-7842(97)01107-0]

Acoustooptic (AO) 2×2 switches are considered the most promising devices for a number of fiber-optic systems (laser Doppler anemometers, regulated multiplexor-demutiplexors, fiber-optic gyroscopes, etc.), since they make it possible not only to switch optical rays, but also to shift their frequency by the frequency of the acoustic wave.

In Refs. 1–3 AO 2×2 directional couplers based on the AO diffraction of two rays with identical wavelengths on two intersecting acoustic beams were investigated. The functional possibilities of such switches in a fiber-optic gyroscope were investigated in Ref. 4. A modified AO switch, in which the optical and acoustic waves propagate in a single plane (the ''planar'' modification of an AO 2×2 switch) was described in Ref. 5. This modification is the most interesting from the standpoint of implementation in integrated optics. The development of such switches is urgently needed, because the employment of integrated-optical elements in fiber-optic gyroscopes is presently considered highly prospective.[6]

This paper presents the theory of planar AO 2×2 directional couplers, which make it possible to switch optical radiations with different wavelengths. Such switches permit the transmission of two different radiations through a single loop, i.e., the production of two independent Sagnac signals associated with passage around the loop. A method for calculating the switches in widely used planar structures is described. Emphasis is placed on the investigation of switching in a planar Ti-$LiNbO_3$ structure, which is considered the most promising material for tasks involving the AO switching of optical radiation,[7] since it has been found to have very small light losses (less than 1 dB/cm), as well as a high AO quality constant.

We note that switches which handle two-color radiation can be created only on the basis of the AO interaction.

Let us consider the interaction in Y-cut Ti-$LiNbO_3$ inclined at a small angle β relative to the $0Z$ optical axis of the crystal. The proposed modification of an AO 2×2 directional coupler is based on anisotropic Bragg diffraction, which permits diffraction of the original radiation in two different directions when it interacts with two collinearly propagating acoustic waves. Figure 1 shows a vector diagram of the interaction in Ti-$LiNbO_3$. The original TE-polarized optical rays with the wavelengths λ_1 and λ_2 ($\lambda_1>\lambda_2$), whose wave vectors are denoted by **K** and **T**, respectively, interact with acoustic waves propagating orthogonally to $0Z'$. The ray **K** diffracts on the acoustic waves $\mathbf{q}_1$ and $\mathbf{q}_2$ in the $\mathbf{K}_1$ and $\mathbf{K}_2$ directions, respectively, and **T** diffracts on the acoustic waves $\mathbf{q}_3$ and $\mathbf{q}_4$ in the $\mathbf{T}_1$ and $\mathbf{T}_2$ directions, respectively. The diffracted rays $\mathbf{K}_1$ and $\mathbf{K}_2$, as well as $\mathbf{T}_1$ and $\mathbf{T}_2$, have TM polarization. We note that $\mathbf{K}_1$ and $\mathbf{T}_1$ (as well as $\mathbf{K}_2$ and $\mathbf{T}_2$) are collinear to one another. This is the condition for the 2×2 AO switching of two rays **K** and **T** with different wavelengths that overlap in the $\mathbf{K}_1$ ($\mathbf{T}_1$) and $\mathbf{K}_2$ ($\mathbf{T}_2$) directions.

Let us determine the parameters of such switching. The cross sections of the wave vector surfaces of a negative uniaxial crystal are described in our case by the following equations:[8,9]

$$\frac{K_z^2}{K_\beta^2}+\frac{K_x^2}{K_e^2}=1 \quad \text{for TE polarization,}$$

$$K_z^2+K_x^2=K_0^2 \quad \text{for TM polarization.} \tag{1}$$

For a positive uniaxial crystal Eqs. (1) change places. Here K_z and K_x are the projections of the light wave vector onto the $0Z'$ and $0X$ axes, where the $0Z'$ axis is the projection of the $0Z$ optical axis onto the waveguide surface of the crystal; the angle between $0Z'$ and $0Z$ equals β; $K_0=2\pi n_0/\lambda$; $K_e=2\pi n_e/\lambda$, where n_o and n_e are the principal refractive indices of the film; and $K_\beta=2\pi n_\beta/\lambda$, where

$$n_\beta=n_0 n_e(n_0^2\sin^2\beta+n_e^2\cos^2\beta)^{-1/2}. \tag{2}$$

We shall henceforth assume that the optical radiation with a wavelength λ_1 has the refractive indices n_0 and n_e and that the radiation with the wavelength λ_2 has the refractive indices N_0 and N_e.

 1063-7842/97/070769-06$10.00

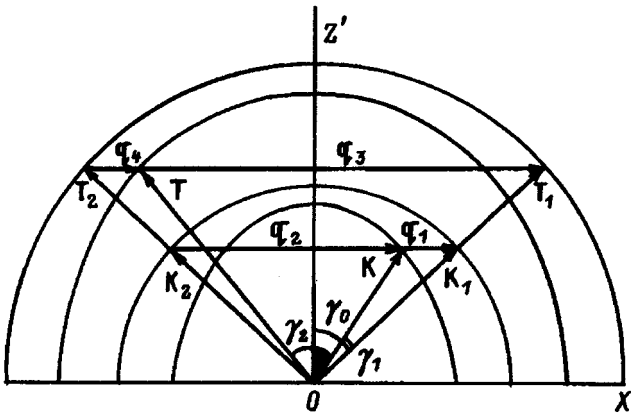

FIG. 1. Vector diagram of an AO 2×2 directional coupler switching optical rays with different wavelengths.

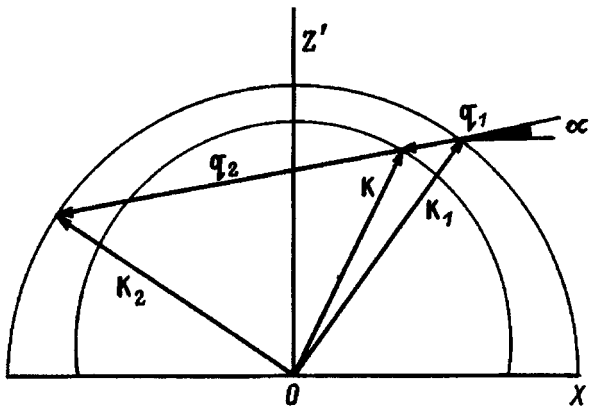

FIG. 2. Vector diagram of the AO interaction of optical radiation with acoustic waves propagating at an angle α to the $0Z'$ axis.

Let the angle γ_0 between **K** and the $0Z'$ axis be assigned. Then, from Eq. (1) it is not difficult to obtain the wave vectors K, K_1, K_2, q_1, and q_2 participating in AO diffraction (Fig. 1):

$$K=\left(\frac{\cos^2\gamma_0}{K_\beta^2}+\frac{\sin^2\gamma_0}{K_e^2}\right)^{-1/2},$$

$$K_1=K_2=K_0, \tag{3}$$

$$\sin^2\gamma_1=1-\frac{K_1^2}{K_0^2}\cos^2\gamma_0,$$

$$q_1=K_1\sin\gamma_1-K\sin\gamma_0, \quad q_2=K_1\sin\gamma_1+K\sin\gamma_0.$$

After γ_1 has been determined, it is not difficult to obtain the AO diffraction parameters for the radiation with the wavelength λ_2:

$$T=\sqrt{T_0^2\cos^2\gamma_1+T_e^2\left(1-\frac{T_0^2}{T_\beta^2}\cos^2\gamma_1\right)},$$

$$T_1=T_2=T_0,$$

$$\tan\gamma_2=T_e\sqrt{1-\frac{T_0^2}{T_\beta^2}\cos^2\gamma_1}\cdot(T_0\cos\gamma_1)^{-1},$$

$$q_3=T_1\sin\gamma_1+T\sin\gamma_2, \quad q_4=T_1\sin\gamma_1-T\sin\gamma_2, \tag{4}$$

where $T_0=2\pi N_0/\lambda_2$, $T_2=2\pi N_e/\lambda_2$, $T_\beta=2\pi N_\beta/\lambda_2$, N_β is specified according to expression (2) after the replacements $n_0\to N_0$ and $n_e\to N_e$, and the angles γ_0, γ_1, and γ_2 are indicated in Fig. 1.

As an example, let us calculate the parameters for the 2×2 switching of radiations with the wavelengths $\lambda_1=1.0$ μm and $\lambda_2=0.8$ μm in a Ti-$LiNbO_3$ structure. Since the refractive indices of the Ti-containing layer on the $LiNbO_3$ crystal do not differ very strongly from the refractive indices of $LiNbO_3$ itself (the differences are of the order of 0.01), we shall use the values of n_0 and n_e of a $LiNbO_3$ crystal in the calculations. Let us consider AO diffraction with $TE_0\rightleftarrows TM_0$ mode conversion. Diffraction takes place on the transverse wave. Let $\beta=5°$. Then, assigning, for example, $\gamma_0=0.85°$, we can easily find the remaining parameters of the interaction: $\gamma_1=1.65°$, $f_1=121$ MHz, $f_2=374$ MHz, $f_3=468$ MHz, and $f_4=156$ MHz. Here f_1, f_2, f_3, and f_4 are the frequencies of the acoustic waves q_1, q_2, q_3, and q_4, respectively, which are related by the spectrum $f_i=q_iu/2\pi$, where u is the velocity of an acoustic wave. It was assumed in the calculations that $u=3.84\times10^5$cm/s. We note that all these frequencies can be generated by a single piezoelectric transducer, as is done, for example, in the AO elements used in laser printers[10] to independently control six optical channels. As for the frequency band of the transducer, there are presently planar AO deflectors, whose band is at ~360 MHz on the 3 dB level (from 120 to 480 MHz)[7] and thus completely covers all the frequencies needed for the functioning of the 2×2 switch described here for handling two-color radiation.

Another modification of the AO 2×2 switching of two-color radiation, in which two, instead of four, switching frequencies are fully satisfactory, is possible. This modification is realized in the case of oblique propagation of the acoustic waves relative to the $0X$ axis. Figure 2 shows the vector diagram of the AO diffraction that is the basis of such transducers. Incident optical radiation with the wave vector **K** diffracts on the collinearly propagating acoustic waves $\mathbf{q}_1$ and $\mathbf{q}_2$ in the $\mathbf{K}_1$ and $\mathbf{K}_2$ directions, respectively. The waves $\mathbf{q}_1$ and $\mathbf{q}_2$ propagate in the crystal at the angle α to the $0X$ axis.

The AO diffraction parameters are determined using the following procedure. We draw the straight line $K_z=K_x\cdot\tan\alpha+K_p$ (where α is the angle between the acoustic vectors and the $0X$ axis and K_p is a certain parameter). This straight line intersects the wave vector surfaces (1) at points, whose projections onto the $0X$ axis equal

$$K_{x_{1,4}}=-\frac{K_p\tan\alpha}{\tan^2\alpha+1}\pm\sqrt{\frac{K_p^2\tan^2\alpha}{(\tan^2\alpha+1)^2}-\frac{K_p^2-K_0^2}{\tan^2\alpha+1}},$$

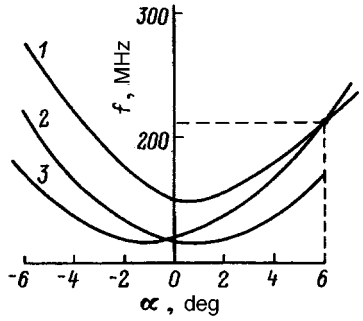

FIG. 3. Dependence of the frequency f_1 of the first harmonic on the angle α.

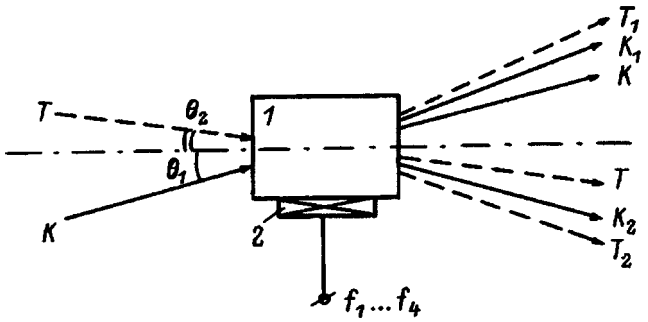

FIG. 4. Optical scheme of the experiment.

$$K_{x_{2,3}} = -\frac{K_p K_e^2 \tan\alpha}{K_e^2 \tan^2\alpha + K_\beta^2} \pm \sqrt{\frac{K_p^2 K_e^4 \tan^2\alpha}{(K_e^2 \tan^2\alpha + K_\beta^2)} - \frac{(K_p^2 - K_\beta^2)^2 K_e^2}{K_e^2 \tan^2\alpha + K_\beta^2}}. \quad (5)$$

Here the K_{xi} are numbered so that the relations $K_{x_1} > K_{x_2} > K_{x_3} > K_{x_4}$ would hold. From (5) we find the amplitudes of the acoustic wave vectors

$$q_1 = |K_{x_3} - K_{x_4}|/\cos\alpha, \quad q_2 = |K_{x_1} - K_{x_3}|/\cos\alpha. \quad (6)$$

As in Ref. 5, we can set $q_2 = 3q_1$, i.e., both acoustic waves are generated by the same transducer: q_1 is generated at the frequency of the first harmonic of the transducer, and q_2 is generated at the frequency of the third harmonic. An analysis reveals that the condition $q_2 = 3q_1$ can always be satisfied by adjusting the parameter K_p for any *a priori* assigned value of the angle α.

Figure 3 presents plots of the dependence of the frequency f_1 of the first harmonic on α for the type of diffraction under consideration when the condition $q_2 = 3q_1$ holds and $\beta = 5°$. Curve *1* was constructed for radiation with $\lambda_2 = 0.8$ μm, and curve *2* was plotted for $\lambda_1 = 1.0$ μm. These values are perfectly adequate for determining the parameters of the AO 2×2 switching of two-color radiation on the same acoustic waves. For this purpose, we construct curve *3*, which is a mirror image of curve *2* relative to the vertical axis. This curve intersects curve *1* at a point, whose projections onto the horizontal and vertical axis give the angle α and the frequency f_1 of the first harmonic by which the 2×2 two-color switching of rays with the wavelengths $\lambda_1 = 1$ μm and $\lambda_2 = 0.8$ μm takes place (in our case $\alpha = 6°$ and $f_1 = 212$ MHz). It was assumed in the calculations that the velocity of the sound wave does not depend on α.

The proposed types of 2×2 switching of two-color radiation can be realized in most planar structures. They require that AO diffraction with mode conversion ($TE_0 \rightleftarrows TE_1$, $TE_0 \rightleftarrows TM_0$, etc.) can occur in the structure. In this case the cross sections of the wave vector surfaces will be described either by circles of the form $K_z^2 + K_x^2 = K_0^2$ or by ellipses of the form $K_z^2/K_\beta^2 + K_x^2/K_e^2 = 1$ [where K_x, K_z, K_β, and K_e were defined in Eqs. (1)], or by a combination of circles and ellipses, as in the case under consideration here, the values of K_0, K_e, and K_β varying from mode to mode. The procedure described above for determining the parameters of AO 2×2 switching is perfectly acceptable in each concrete case.

Acoustooptic diffraction with mode conversion has been observed in many planar structures. For example, anisotropic diffraction in a film of amorphous TeO_2 grown on a quartz substrate was investigated in Ref. 11. Mode conversion of the $TE_0 \rightleftarrows TM_0$ type as a result of AO diffraction in an As_2S_3 film grown on a $LiNbO_3$ substrate was considered in Ref. 12. The diffraction efficiency was 93% for an electrical power of 150 mW, a wavelength of the optical radiation equal to 1.153 μm, and an acoustic frequency equal to 200 MHz. The AO interaction in Ti-$LiNbO_3$ was investigated in Refs. 7, 13, and 14. The 2×2 switch in such a structure was investigated in Ref. 13 in a somewhat different geometry, in which the acoustic waves propagate noncollinearly to one another. A high AO diffraction efficiency was observed with a low level of interference between the channels being switched (the polarization extinction ratio was greater than 40 dB). As was noted above, Ti-$LiNbO_3$ structures are presently very promising for the AO switching of optical radiation.

The foregoing ideas were tested experimentally in the example of bulk AO diffraction in a TeO_2 single crystal. Figure 4 presents the optical scheme of the experiment. The optical radiation K with the wavelength λ_1 impinges on crystal *1* at the angle Θ_1. Acoustic waves generated by transducer *2* propagate within the crystal. The radiation K diffracts on acoustic waves with the frequencies f_1 and f_2 in the directions K_1 and K_2, respectively. The radiation T with the wavelength λ_2 ($\lambda_1 > \lambda_2$) impinges on the face of the crystal *1* at the angle Θ_2 and diffracts in the crystal on acoustic waves with the frequencies f_3 and f_4 in the T_1 and T_2 directions, respectively. The crystal measured $8\times8\times10$ mm along the [110], [$1\bar{1}0$], and [001] directions, respectively. The transverse acoustic wave generated by $LiNbO_3$ transducer *2* propagated along [110] with a direction of displacement along [$1\bar{1}0$]. The ''planar'' modification of the 2×2 switching of two-color radiation with the wavelengths $\lambda_1 = 0.5145$ and $\lambda_2 = 0.488$ μm (radiation from an Ar laser) was realized. All the AO diffraction parameters were appropriately recalculated for TeO_2. The vector diagram in Fig. 1 was used to select the frequencies f_1, f_2, f_3, and f_4, by which 2×2 switching of the two-color radiation took place. The calculations show that, if, for example, the switching

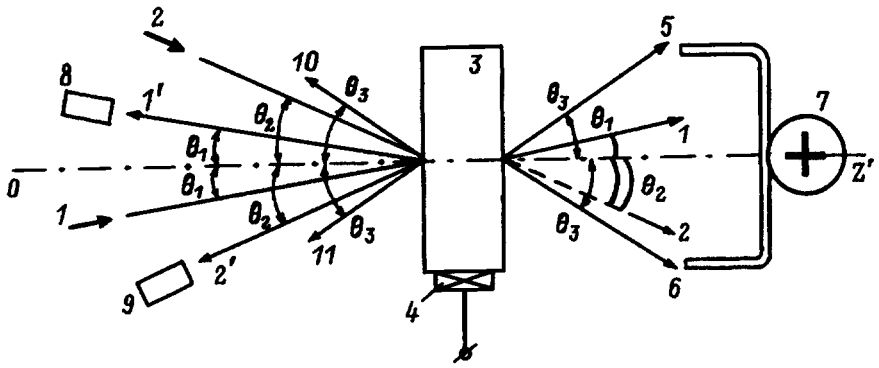

FIG. 5. An AO 2×2 directional coupler in a fiber-optic gyroscope.

frequencies $f_1=32$ MHz and $f_2=3f_1=96$ MHz are chosen for the radiation at λ_1, the radiation at λ_2 will be switched by the frequencies $f_4=41$ MHz and $f_3=97$ MHz. The frequencies f_1 and f_4 are covered by the band of the first harmonic of the piezoelectric transducer, and f_2 and f_3 are covered by the band of the third harmonic. The experiment showed that 2×2 switching by these frequencies is, in fact, observed. The corresponding diffracted rays propagate collinearly to one another in the crystal. At the exit from the crystal these rays are not parallel, the angle between them being of the order of $3-5'$. Some additional experimental investigations revealed that a situation in which the diffracted rays leaving the crystal propagate parallel to one another, i.e., K_1 and T_1, as well as K_2 and T_2, merge with one another, can be achieved fairly simply by adjusting the angles and frequencies. Such a situation was realized, in particular, for $f_1=32$ MHz, $f_2=98$ MHz, $f_3=110$ MHz, and $f_4=36$ MHz. The frequencies f_1 and f_4 are covered by the frequency band of the first harmonic of the transducer, and f_2 and f_3 are covered by the band of the third harmonic. This makes it possible to ensure collinear propagation of the switched optical rays even when there is a medium with a practically arbitrary refractive index (for example, the glass shaping optics of an optical fiber) after the AO cell. In other words, there is a possibility for creating the best conditions for switching optical rays propagating along the fiber-optic loop of a gyroscope.

Let us examine the functioning of a 2×2 directional coupler in a fiber-optic gyroscope. Figure 5 presents the optical scheme of a gyroscope. The original optical radiations with the wavelengths λ_1 and λ_2, whose frequencies are ω_1 and ω_2, are generated by sources *1* and *2*, respectively. These radiations are directed at AO directional coupler *3* at the angles Θ_1 and Θ_2 to the $0Z'$ axis of the cell, respectively. Acoustic waves generated by transducer *4* at the frequencies f_1, f_2, f_3, and f_4 propagate within the cell. Radiation *1* diffracts on f_1 and f_2, and radiation *2* diffracts on f_3 and f_4. The corresponding diffracted rays are collinear to one another; they propagate in directions *5* and *6* at the same angle Θ_3 to the $0Z'$ axis. When all the acoustic frequencies are generated simultaneously, rays with the frequencies ω_1-f_2 and ω_1+f_1, which form as a result of the AO diffraction of the radiation at λ_1, as well as rays with the frequencies ω_2+f_3 and ω_2-f_4, which form as a result of the diffraction of the radiation at λ_2, propagate in directions *5* and *6* (all the frequency changes are easily traced using the vector diagram in Fig. 1). After passage around fiber loop *7* in opposite directions these rays reenter AO switch *3* and diffract on the same acoustic waves: the radiation at λ_1 diffracts on f_1 and f_2, and the radiation at λ_2 diffracts on f_3 and f_4. As a result of the repeated diffraction, two rays with the wavelength λ_1 and the frequency ω_1 propagate in direction *1′* (the acoustic frequencies f_1 and f_2 in this case are systematically deducted as a result of double diffraction). These rays interfere in photodetector *8*, which is placed on the path of ray *1′*. Two rays at λ_1 with the frequencies $\omega+(f_2-f_1)$ also propagate toward source *1*. Two rays with the wavelength λ_2 and the frequency ω_2 propagate in direction *2′*. Here, too, the acoustic frequencies f_3 and f_4 are systematically deducted. These rays are directed to photodetector *9*. Two rays with the frequency $\omega_2+(f_3-f_4)$ propagate toward source *2*. The radiations returning to the sources do not disrupt their generation, since they have different frequencies. Thus, photodetectors *8* and *9* detect the Sagnac signals associated with passage around the loop at the different wavelengths λ_1 and λ_2. We note that rays *1* and *1′* propagate symmetrically relative to the $0Z'$ axis, i.e, at the same angles Θ_1. Rays *2* and *2′* are symmetric, and they propagate at the angles Θ_2. Diffracted rays *5* and *6*, as well as rays *10* and *11*, which do not undergo repeated diffraction, propagate at equal angles Θ_3.

The modification just described permits the creation of the best conditions for the entry (departure) of the radiation at λ_1 and the radiation at λ_2 in the fiber-optic loop by independently adjusting the intensities and frequencies of the acoustic waves and optimizing the conditions for detecting them in photodetectors *8* and *9*. This modification permits the use of the ''zero'' method for each radiation on an individual basis and correction of the drift of the zero and provides for frequency ''decoupling'' between the radiations of the sources and the radiations propagating in the fiber loop.

The question of the efficiency of the AO interaction in a planar structure is important. We assume that a Rayleigh wave propagates in *X*-cut lithium niobate along the *Y* axis. The incident and diffracted rays propagate in a planar waveguide near the *Z* axis. The efficiency of the diffraction of optical radiation propagating in a planar structure at the Bragg angle to a surface acoustic wave is given by the expression[15]

$$h=\sin^2\left[\frac{2\pi\delta_0}{\lambda_0}\frac{2\pi L}{\Lambda}\frac{(N_iN_d)^{3/2}}{4\cos\Theta}F\right]. \quad (7)$$

Here δ_0 is the amplitude of the displacement of the transverse component of the Rayleigh wave; λ_0 and Λ are the wavelengths of the light and sound, respectively; L is the AO interaction length; N_i and N_d are the refractive indices of the incident and diffracted waves, respectively; $\cos\Theta=\sqrt{\cos\Theta_i\cos\Theta_d}$, where Θ_i and Θ_d are the angles between the incident and diffracted waves and the acoustic wave front; and F is the overlap integral, which is defined as

$$F=\frac{\int_{-\infty}^{0}E_d^*(x)f(x)E_i(x)dx}{\sqrt{\int_{-\infty}^{\infty}|E_d(x)|^2dx\int_{-\infty}^{\infty}|E_i(x)|^2dx}}. \quad (8)$$

Here E_i and E_d are the electric fields of the incident and diffracted radiations, and $f(x)$ is a function which takes into account all the mechanisms (effects) that influence AO diffraction. When lithium niobate is employed, this function equals

$$f(x)=f_{A0}+f_{E0}+f_{SR}, \quad (9)$$

where f_{A0} is the component that takes into account the direct elastooptic effect, f_{E0} takes into account the piezoelectric effect, and f_{SR} takes into account the surface strain (surface perturbation) of the medium

$$f_{A0}(x)=\frac{\Lambda}{2\pi\delta_0}P_{IJ}S_J(x),$$

$$f_{E0}(x)=\frac{\Lambda}{2\pi\delta_0}r_{Ik}e_k(x),$$

$$f_{SR}(x)=\frac{\Lambda(n^2-1)}{2\pi N_iN_d}\delta(x). \quad (10)$$

Here P_{IJ} is the matrix of elastooptic coefficients; S_J is the distribution of the strain along the X axis caused by the acoustic wave; r_{Ik} is the matrix of electrooptic coefficients; $e_k(x)$ is the electric field appearing as a result of the piezoelectric effect; n is the refractive index of the medium; and $\delta(x)$ is a Dirac delta function. In most practical cases $f_{SR}\ll f_{A0}$, f_{E0}; therefore, we can set $f_{SR}=0$.

In our case (a Rayleigh wave propagating in X-cut lithium niobate along the Y axis and light propagating in a planar waveguide along the Z axis), only the strain components S_1, S_2, and S_6 and the electric field components e_1 and e_2 are nonzero. Hence, we can determine all the components of $f(x)$ using the method described in Ref. 15. In particular, on the basis of formulas (10), for TE→TE AO diffraction we have

$$f_{A0}(x)=\frac{\Lambda}{2\pi\delta_0}[P_{21}S_1+P_{22}S_2],$$

$$f_{E0}(x)=\frac{\Lambda}{2\pi\delta_0}r_{22}e_2. \quad (11)$$

For TM→TM diffraction we can write

$$f_{A0}(x)=\frac{\Lambda}{2\pi\delta_0}[P_{11}S_1+P_{12}S_2],$$

$$f_{E0}(x)=\frac{\Lambda}{2\pi\delta_0}r_{22}e_2. \quad (12)$$

For TM→TE diffraction the corresponding expressions are

$$f_{A0}(x)=\frac{\Lambda}{2\pi\delta_0}\frac{P_{11}-P_{12}}{2}S_6,$$

$$f_{E0}(x)=-\frac{\Lambda}{2\pi\delta_0}2r_{22}e_1. \quad (13)$$

To obtain the concrete form of all the coefficients in f_{A0} and f_{E0} we must know the distributions of the strains S_i and the electric field e_i with respect to X, which, however, can be obtained only by numerical methods and thus present a problem in themselves. A preliminary estimate shows that the diffraction efficiency in our case will be about 10 times smaller than the diffraction efficiency in the geometry that is presently widely used (a Rayleigh wave propagating in Y-cut lithium niobate along the Z axis and an optical wave propagating along the X axis).

This is perfectly acceptable for many problems (particularly for the switching of radiation in systems of fiber-optic gyroscopes).

The following conclusions can be drawn on the basis of the results of this work. A method for calculating the parameters of planar AO 2×2 directional couplers that switch rays with different wavelengths has been proposed. The angular and frequency characteristics of AO 2×2 switches based on a planar Ti-$LiNbO_3$ structure have been calculated. The experiments performed using bulk diffraction in TeO_2 have confirmed the basic theoretical assumptions. The functional possibilities of an AO 2×2 directional coupler that switches rays with different wavelengths in a fiber-optic gyroscope have been described. A method for calculating the efficiency of the AO interaction in planar structures, which takes into account both the direct elastooptic effect and the piezoelectric effect, has been described.

[1] S. N. Antonov, Yu. V. Gulyaev, V. M. Kotov, and P. V. Poruchikov, Radiotekh. Elektron. (Moscow) **32**, 623 (1987).
[2] S. N. Antonov and V. M. Kotov, Radiotekhnika (8), 22 (1988).
[3] V. M. Kotov, Zh. Tekh. Fiz. **63**(1), 180 (1993) [Tech. Phys. **38**, 44 (1993)].
[4] V. M. Kotov, Pis'ma Zh. Tekh. Fiz. **19**(15), 1 (1993) [Tech. Phys. Lett. **19**, 475 (1993)].
[5] S. N. Antonov and V. M. Kotov, Zh. Tekh. Fiz. **60**(10), 166 (1990) [Sov. Phys. Tech. Phys. **35**, 1212 (1990)].
[6] A. M. Savel'ev and T. I. Solov'eva, Zarubezhn. Radioelektron. (6), 55 (1982).
[7] R. V. Shmidt, IEEE Trans. Ultrason. **SU-23**, 22 (1976).
[8] V. V. Lemanov and O. V. Shakin, Fiz. Tverd. Tela (Leningrad) **14**, 229 (1972) [Sov. Phys. Solid State **14**, 184 (1972)].
[9] T. Tamir (ed.), *Integrated Optics*, Springer-Verlag, Berlin (1975) [Russ. transl., Moscow (1978)].
[10] H. Koebner (ed.), *Industrial Applications of Lasers*, Wiley (1984).
[11] Y. Ohmachi, Electron. Lett. **9**, 539 (1973).
[12] Y. Ohmachi, J. Appl. Phys. **44**, 3928 (1973).
[13] E. M. Korablev, V. V. Proklov, G. V. Titarenko, and Yu. L. Kopylov,

Pis'ma Zh. Tekh. Fiz. **12**, 465 (1986) [Sov. Tech. Phys. Lett. **12**, 189 (1986)].
[14] *Integrated Optics: Proceedings of the 3rd European Conference ECIO'85*, Berlin, 1985, pp. 169–173.
[15] J. Xu and R. Stroud, *Acoustooptic Devices. Principles, Design, and Applications*, Wiley, New York (1992).

Translated by P. Shelnitz

Electron-beam-induced fluorescence of carbon dioxide clusters. II. Molecular beam with clusters

S. Ya. Khmel' and R. G. Sharafutdinov

Institute of Thermal Physics, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia
(Submitted February 8, 1996)
Zh. Tekh. Fiz. **67**, 63–71 (July 1997)

The fluorescence of CO_2 clusters excited by an electron beam in a molecular beam formed from a free jet expansion of carbon dioxide is investigated. Data on the fluorescence of the clusters are obtained by comparing the electron-beam-induced fluorescence with the signal appearing in an electronic-fluorescence detector as a result of reflection of the cluster beam from an obstacle placed after the electron beam, as well as with the intensity of the cluster beam. It is established that, unlike a jet expansion, for CO_2 clusters in a molecular beam the fluorescence yield from a cluster decreases significantly with increasing cluster size and is very small for clusters of large size. It is concluded on the basis of the results from an investigation of the fluorescence of CO_2 clusters that the electron-beam-induced fluorescence technique can be used for measurements in molecular beams with CO_2 clusters. [S1063-7842(97)01207-5]

INTRODUCTION

In our preceding paper[1] we examined problems concerned with the use of the electron-beam fluorescence (EBIF) technique in free jet expansions, i.e., gas flows with clusters. However, the EBIF technique is also employed in molecular beams formed from free jets to measure the gas density and the populations of the rotational levels of the molecules.[2–4]

In a molecular beam with clusters, as in a jet, the influence of the clusters on the results of electron-beam measurements, particularly on the electron-beam-induced fluorescence of the clusters, must be taken into account. On the one hand, the contribution of clusters to the emission of a molecular beam can lead to distortion of the results of measurements of the parameters of the monomer and, on the other hand, it can be used to obtain information on the clusters themselves.

This paper examines van der Waals clusters, specifically CO_2 clusters. Because of the small binding energy in such clusters, each molecule retains its own ''individuality'' within the cluster and its fluorescence can be regarded in most cases as the fluorescence of an individual molecule, on which the remainder of the cluster acts.[5,6]

As we know, the EBIF technique is based on a sequence of processes involving excitation of the molecules by electron impact and spontaneous emission. Differences between the fluorescence of monomers and clusters can be expected because of dissimilarities in any step of the sequence of excitation-emission processes. The following dissimilarities are possible. First, the excitation cross section of the molecules in a cluster can differ from the excitation cross section of the free molecules. Second, the fluorescence intensity per cluster molecule can differ from that of the free molecules for the following reasons: changes in the transition probability from the excited state of a molecule in a cluster, quenching of the fluorescence of an excited molecule in a cluster, and ejection of an excited molecule from a cluster upon fragmentation. Third, the fluorescence spectrum can vary, since the values of the terms of the ground and excited states of a certain molecule can be altered by the action of the remainder of the cluster on that molecule.

To study the fluorescence of clusters in a molecular beam, the optical signal must be compared with a quantity that is proportional to the total density of the gas-condensate mixture, i.e., the total number of molecules in a unit volume in the free and bound states. The fluorescence of clusters in a jet was investigated in a similar manner, and the x-ray signal was employed as the quantity that is proportional to the total density in Ref. 1. In the present case we propose using the intensity of the molecular beam I (the number of molecules intersecting a unit area of a transverse section of the beam per unit time). This quantity is measured by an intensity sensor,[7] even in a molecular beam with van der Waals clusters,[8]

$$I = n_m v_m + n_{\mathrm{cl}} v_{\mathrm{cl}}. \tag{1}$$

Here v_m and v_{cl} are the velocities of the molecules and clusters; n_m and n_{cl} are the number densities of the monomers in the gas phase and in the bound state (i.e., in clusters); and $n_{\mathrm{cl}} = \Sigma_N N n_N$, where N is the cluster size (the number of monomers in the cluster and n_N is the number density of the clusters of size N.

In a jet expansion the velocity increases due to the release of the heat of condensation into the flow.[8] For example, it was shown in Refs. 9 and 10 that as the backing pressure varies in the range $P_0 = 8 - 609$ kPa the velocity of the monomers in a carbon dioxide jet increases by about 22%, and the velocity of the clusters is not much less than the velocity of the monomers and also increases to a small extent. Accordingly, we have a similar picture in a molecular beam, which means that

$$I \approx (n_m + n_{\mathrm{cl}}) v_m \approx nV, \tag{2}$$

where $n = n_m + n_{\mathrm{cl}}$ is the total density and V is the velocity of the gas at low pressures before the onset of condensation.

 1063-7842/97/070775-08$10.00

Thus, the intensity of the molecular beam is, in fact, proportional to the total density of the gas-condensate mixture to within the variation of the velocity.

Let us briefly describe some features of the formation of a cluster beam from a jet expansion. The fraction of the condensate and the cluster size in a jet expansion are determined by the stagnation parameters P_0 and T_0 and by the nozzle diameter d_*.[11,12] The fraction of the condensate, even at high pressures, or, as is said, in a regime with developed condensation, is of the order of 30% and increases only slightly as the pressure rises further. The cluster size can increase without bound as the pressure is increased. These parameters depend weakly on the distance to the nozzle, if it is appreciably greater than the nozzle-condensation front distance. In the molecular beam formed from a jet expansion, the mean cluster size is somewhat greater than in the jet, and the fraction of the condensate is significantly higher due to enrichment of the beam with clusters as a result of the differences in the thermal dispersion of the monomers and clusters, i.e, the light and heavy components. Its intensity increases sharply for the same reason. The fraction of condensate in a molecular beam generally exceeds 90% (see Refs. 8 and 13) at sufficiently large values of P_0. This is the main advantage of using a molecular beam to investigate the fluorescence of the clusters in front of a free jet. The main disadvantage is the weakness of the signal.

The electron-beam-induced fluorescence of N_2 clusters in a molecular beam was first investigated in the manner indicated above in Ref. 4. This method was subsequently developed in several studies[14–18] of the fluorescence of CO_2, N_2O, H_2O, and N_2 clusters excited by a low-energy electron beam in a molecular beam. In Ref. 10 the optical signal of a CO_2 cluster beam excited by electron impact was compared with the theoretical value of the total density. The fluorescence of clusters can be excited not only by electron impact, but also by photon irradiation. Therefore, we mention a review of the work performed on the fluorescence of inert gas clusters in molecular beams excited by synchrotron radiation.[5]

The purpose of the present work is to investigate the electron-beam-induced fluorescence of CO_2 as applied to diagnostics in a molecular beam with clusters.

EXPERIMENT

The experiments were carried out on the VS-4 low-density gas-dynamic facility of the Institute of High-Temperature Physics of the Siberian Branch of the Russian Academy of Sciences.[19] The equipment used in the present work is schematically represented in Fig. 1.

Gas source *1* was an axisymmetric sonic nozzle with a diameter $d_*=0.95$ mm. The pressure in the prechamber or the backing pressure was varied in the range 8–609 kPa, the nozzle temperature being held at the room-temperature level and monitored using thermocouple *2*. The gas from the source expanded in the vacuum chamber, which was evacuated by booster pumps with a throughput of 35 000 liter/s and a liquid nitrogen cryogenic pump with a throughput for CO_2 as high as 20 g/s. This enabled us to maintain the pressure in the vacuum chamber at the 0.1–1 Pa level in the working regimes with consumption of a gas. Technical-grade carbon dioxide was employed as the working gas without additional purification.

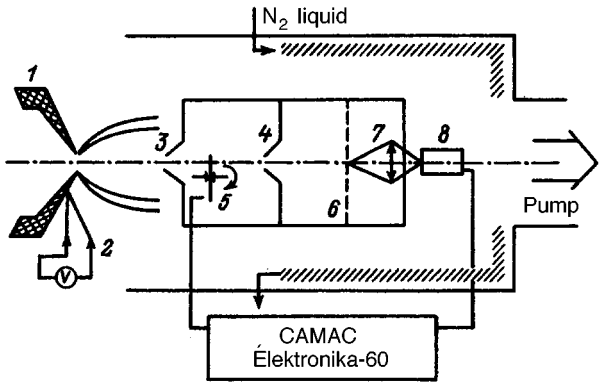

FIG. 1. Experimental setup.

A molecular-beam system was installed within the working chamber of the VS-4 facility for time-of-flight measurements. The molecular beam was formed from the jet by skimmer *3* and collimator *4* The diameter of the skimmer (the conical diaphragm) was 3.23 mm, and the collimator had a rectangular shape and measured 2.5×9.6 mm. The time-of-flight analysis was carried out according to an ordinary scheme. Beam chopper *5*, a disk with two slits rotating at 180 Hz, sliced out bunches of molecules, which were picked up by a detector after traversing a definite distance, which is called the time-of-flight base.

In the present case an electronic-fluorescence detector that we developed[3] was used instead of the traditional ionization detector. It consists of electron beam *6* (its diameter is $\approx 1-2$ mm, the energy of the electrons is 5.5 keV, the current is ≈ 20 mA, and the distance from the point where the beam emerges from the electron gun to the collector is >190 mm), optical system *7* for collecting the radiation, whose parts were made from quartz, and FEU-39A photomultiplier *8*. The distance between the beam chopper and the electron beam, or the time-of-flight base, was equal to 220 mm, and the distance between the skimmer and the beam chopper was 245 mm. The photomultiplier was mounted on a two-component coordinate placer, which made it possible to adjust it and to obtain transverse profiles of the molecular beam. The optical system and the photomultiplier provides for the detection of radiation in the spectral range 200–600 nm.

It is seen from this description that the employment of an electronic-fluorescence detector is equivalent to the use of the EBIF technique in a molecular beam. The main difficulty arising here is the low level of the effective signal, the signal-to-noise ratio usually being <1. It was overcome by a signal accumulation system. An accumulation system based on a multichannel analyzer in a standard CAMAC crate controlled by an Élektronika-60 minicomputer permits the accumulation of 1000 signals during 1.5 min, the time-of-flight signal being divided into 1000 temporal channels.

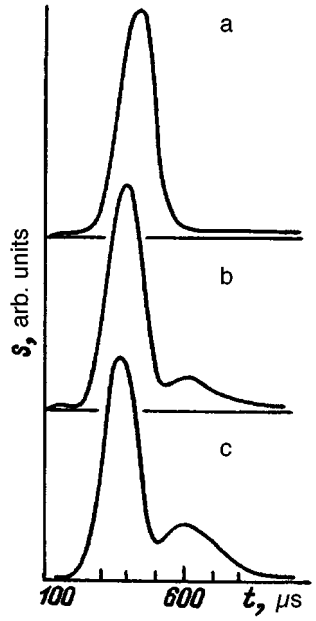

FIG. 2. Evolution of the form of the time-of-flight signal as a function of the backing pressure in a molecular beam formed from a carbon dioxide jet expansion. $d_* = 0.95$ mm; $T_0 = 293$ K; nozzle–skimmer distance x, mm (P_0, kPa): a — 65 (8.5), b — 230 (101), c — 360 (608).

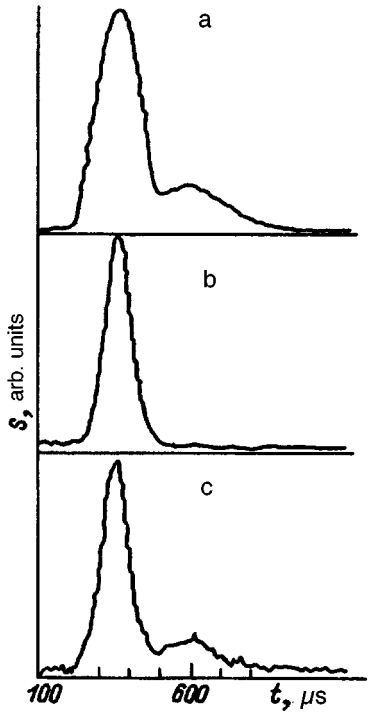

FIG. 3. Evolution of the form of the time-of-flight signal as a function of the detector geometry. Electron-beam/lens distance z, a — 80, b — 158, c — 158 mm (a plate of quartz glass was placed at a distance of 37 mm between the electron beam and the lens); $d_* = 0.95$ mm, $T_0 = 285$ K, $P_0 = 608$ kPa, $x = 360$ mm.

RESULTS

A time-of-flight system with an electronic-fluorescence detector[3] was used in the present work to investigate the electron-beam-induced fluorescence of CO_2 clusters in a molecular beam. Plots of the dependence of the amplitude of the time-of-flight signal (i.e., the radiation excited by the electron beam) on the backing pressure, the nozzle–skimmer distance, etc. were obtained. This technique is similar to modulation of a molecular beam with synchronous detection at the modulation frequency.

In the case of electron-beam excitation, the fluorescence spectrum of the rarefied carbon dioxide, including the condensed CO_2 in the jet, contains only bands belonging to the molecular ion.[20–22] The spectral characteristics of the optical system and the photomultiplier provided for the simultaneous recording of the fluorescence in the $A-X$ and $B-X$ band systems of CO_2^+.

When an electronic-fluorescence detector is used, the obstacle placed after the electron beam makes it possible to record the cluster component of the molecular beam.[9] The clusters striking the surface of the obstacle undergo fragmentation, and the monomers and lighter clusters formed fly back from the surface, enter the electron beam, undergo excitation, and emit radiation, causing the appearance of a second peak in the time-of-flight signal. This peak is also present in an ordinary molecular beam, but it is very weak, and the presence of clusters sharply enhances it. Let us demonstrate this. Figure 2 presents the dependence of the time-of-flight signal on the backing pressure. In this case the lens of the optical system for collecting the radiation served as the obstacle. At low pressures we have an ordinary time-of-flight signal.[23] As the pressure is increased, condensation begins in the jet,[1,11,16,21] and clusters appear in the molecular beam formed from the jet. Under such conditions the time-of-flight signal shifts toward shorter flight times, and, in addition, at pressures above a certain value a second signal appears at flight times longer than those of the first signal. In these measurements the nozzle–skimmer distance is varied so that the influence of the skimmer interaction and scattering on the background gas on the form of the signal would be eliminated.[24]

Figure 3 shows the variation of the shape of the bimodal time-of-flight signal as a function of the detector geometry. The first signal corresponds to a small electron-beam/lens distance, the second signal corresponds to a large distance, and in the third case this distance was also large, but a transparent quartz plate was placed between the lens and the beam. It is seen from the figure that as the electron-beam/lens distance is increased, the second peak in the time-of-flight signal vanishes. However, when there is an obstacle between the electron beam and the lens, the second peak reappears. Thus, an obstacle placed at a small distance after the electron beam can, in fact, serve as a probe for the cluster component of the molecular beam.

The first peak of the time-of-flight signal is formed by the emission of the monomers and clusters, the contribution of the latter, unlike that of the monomers, being unknown. The second peak in the time-of-flight signal reflects only the behavior of the cluster component. By comparing the dependences of the amplitudes of the first and second peaks on the backing pressure, the nozzle–skimmer distance, the background pressure, etc. we can qualitatively estimate the con-

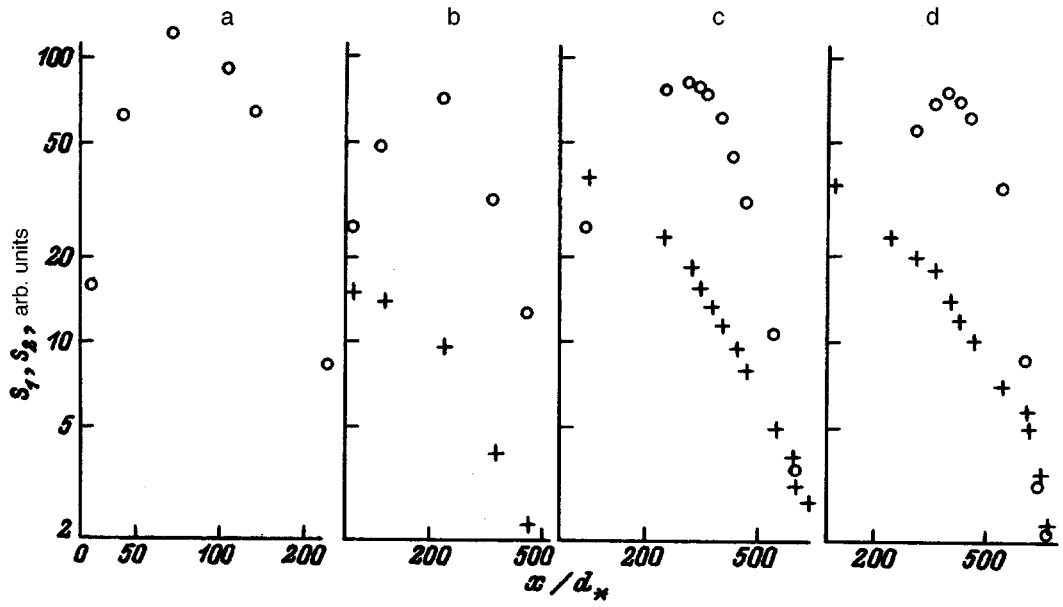

FIG. 4. Dependence of the amplitude of the first (S_1, ○) and second (S_2, +) peaks of the time-of-flight signal on the distance x. $d_*=0.95$ mm, $T_0=293$ K; P_0, kPa: a — 8, b — 101, c — 203, d — 608.

tribution of the clusters to the emission of the molecular beam.

Figure 4 presents plots of the amplitudes of the first (S_1) and second (S_2) peaks as functions of the distance for several pressures. All the values are presented in relative, but comparable units of measure. The first of these plots (Fig. 4a) corresponds to an ordinary time-of-flight signal consisting of one peak. Its amplitude at first increases with increasing distance, but after a maximum has been achieved at a certain distance, it begins to decrease quite sharply. Such behavior of the amplitude of the time-of-flight signal corresponds well with the literature data[23,25] and is attributed to the influence of the skimmer interaction at small distances and the influence of scattering on the background gas at large distances. A bimodal time-of-flight signal is observed for the remaining pressures (Figs. 4b–d). The behaviors of the amplitude of the first peak of the bimodal signal and the amplitude of an ordinary time-of-flight signal are similar to one another, but differ fundamentally from the behavior of the amplitude of the second peak of the bimodal signal over the entire range of pressures. The amplitude of the second peak decreases monotonically with increasing distance and scarcely depends either on the skimmer interaction or on scattering on the background gas.

Figure 5 presents plots of the dependence of the amplitude of the first (S_1) and second (S_2) peaks of the time-of-flight signal on the pressure of the background gas P_H. Conditions with a fairly pronounced bimodal type of signal were selected here, the skimmer being at a distance from the nozzle at which scattering on the background gas already begins to influence the amplitude of S_1. For clarity, the figure shows the evolution of the shape of the time-of-flight signal as P_H is varied. Just as in the preceding case (Fig. 4), the second peak is significantly more conservative toward scattering on the background gas than the first peak. Also, while S_2 decreases by about 1.5 fold as the pressure rises, S_1 decreases by nearly 10 fold. We recall that the second peak of the bimodal signal reflects the behavior of the cluster component of the molecular beam. Thus, it follows from the data presented above that the first peak of the bimodal time-of-flight signal is formed predominantly by the emission of the monomers, while the contribution of the clusters is insignificant. Thus, when a molecular beam with CO_2 clusters is excited by a high-energy electron beam, it can be concluded that the clusters make a smaller contribution to the emission detected than do the monomers. If it is taken into account that the fraction of the condensate in the molecular beam exceeds 90% at fairly large values of P_0, the intensity of the electron-beam-induced fluorescence per cluster molecule is much smaller than the analogous value for the free molecules.

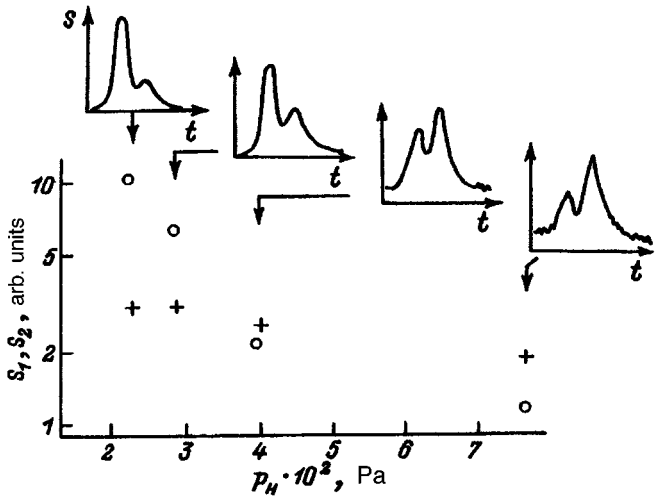

FIG. 5. Dependence of the amplitude of the first (S_1, ○) and second (S_2, +) peaks of the time-of-flight signal on the pressure of the background gas P_H. $d_*=0.95$ mm, $T_0=293$ K; $P_0=354.6$ kPa, $x=550$ mm.

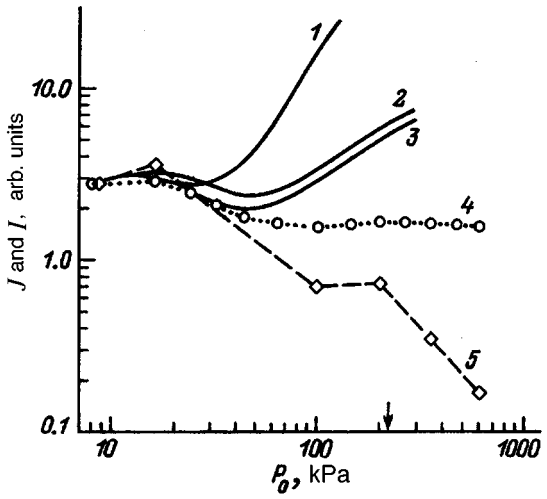

FIG. 6. Dependence of the intensity (*1*) and the optical signal (*2–5*) of the molecular beam formed from a CO_2 jet expansion on the backing pressure (*2* — visible region of the spectrum, *3–5* — ultraviolet region of the spectrum, arrow — P_0 for $N\approx 600$; x, mm: *3* — 100, *4* — 65–360, *5* — 100; *1–3* — data from Refs. 14–16, *4* and *5* — data from the present work.

This result can be obtained by another, more general and exact method, which was mentioned in the Introduction, viz., by comparing the optical signal with the intensity of the molecular beam. The latter was not measured in the present work. However, the optical signal that we obtained can be compared with the intensity data of other investigators.[14–16] In general, such a comparison must be made for conditions that are free of the skimmer interaction and scattering on the background gas. In those measurements the nozzle–skimmer distance was adjusted so that the influence of both factors on the form of the signal would be eliminated. A procedure for determining the conditions under which the influence of these factors is insignificant was described in Ref. 24. Approximately the same conditions exist at the distances corresponding to the positions of the maxima on the curves in Fig. 4.

Figure 6 presents plots of the dependence of the optical signal J of the molecular beam measured in the present work (curve *4*) and the total intensity I of a molecular beam formed from a CO_2 jet (curve *1*)[14–16] on the backing pressure P_0. The zeroth moment of the time-of-flight signal served as the optical signal.[3,26] This is preferable to using its amplitude. An electronic-fluorescence detector scheme which eliminates the appearance of the second peak was chosen for these measurements.

In the figure all the amplitudes ''merge'' with one another at $P_0=8$ kPa in the absence of condensation. The experiments in Refs. 14–16 were carried out under the following conditions: $d_*=1$ mm, $T_0\approx 280$ K, $P_0=8-300$ kPa, $d_s=3.5$ mm, and $x=100$ mm. In the present work the conditions were as follows: $d_*=0.95$ mm, $T_0\approx 283$ K, $P_0=8-609$ kPa, $d_s=3.23$ mm, and $x=65-360$. As we see the conditions are fairly similar.

It was noted in the Introduction that the total intensity is approximately proportional to the total density of the molecular beam. Thus, the only obstacle to correctly comparing the values of I and J in Fig. 6 is the difference between the nozzle–skimmer distances at which they were measured. This distance is variable for curve *4*: it increases with increasing pressure, and it is greater than the distance in Refs. 14–16 above $P_0\approx 24$ kPa. This means that in Refs. 14–16 the molecular beam formed under conditions with a significant skimmer interaction at $P_0>24$ kPa (compare Fig. 4). However, this does not introduce a large error into the measurements of the intensity in the cluster beam, since more than 90% of it is determined by the cluster component at sufficiently high values of P_0,[8,13] and the cluster component is conservative toward the skimmer interaction.

To correctly compare the values of J and I (curves *1* and *4*) in Fig. 6, we correct the values of the optical signal, i.e., we adjust them to the distance $x=100$ mm. This correction can easily be made using the results of measurements similar to those presented in Fig. 4, but taken for $T_0\approx 285$ K. We take the amplitudes of the first peak S_1 of the time-of-flight signal at the distance $x=100$ mm as the optical signal. Curve *5*, which corresponds to the data indicated, is appreciably lower than curve *4*. This is a consequence of the strong skimmer interaction, which primarily influences the light component of the molecular beam and significantly reduces the density of the monomers in it.

It follows from the results presented that, as P_0 increases, its intensity (curve *1*) increases sharply when condensation begins in the jet and clusters appear in the molecular beam formed from that jet, while the optical signal is approximately constant or even decreases (curves *4* and *5*).

DISCUSSION

The intensity of the optical emission from a cluster beam $J_{\Delta\lambda}$ is specified by the expression

$$J_{\Delta\lambda}=\beta(n_m+\alpha n_{\rm cl}). \tag{3}$$

Here β is the fluorescence intensity per molecule of the gas phase; $\beta=(J_{\Delta\lambda}/n)_0$, where the subscript 0 refers to the gas phase; α is the fluorescence yield of the clusters (the fluorescence intensity per cluster molecule normalized to the value for the gas phase); and $\alpha=(J_{\Delta\lambda})_{\rm cl}/n_{\rm cl}\beta$, where $(J_{\Delta\lambda})_{\rm cl}$ is the cluster component of the intensity of the optical emission of the gas-condensate mixture.

Then we can express the ratio of the optical signal to the intensity of the molecular beam in the following manner, taking into account Eqs. (3) and (1) and the normalization used in Fig. 6:

$$\frac{J_{\Delta\lambda}}{I}=\frac{(n_m+\alpha n_{\rm cl})V}{n_m v_m+n_{\rm cl}v_{\rm cl}}. \tag{4}$$

If the insignificant variation of the velocity with P_0 is neglected, then $V\approx v_m\approx v_{\rm cl}$, and Eq. (4) can be rewritten in the form

$$\frac{J_{\Delta\lambda}}{I}\approx\frac{n_m+\alpha n_{\rm cl}}{n_m+n_{\rm cl}}>\alpha. \tag{5}$$

This ratio can be used to obtain an upper estimate of α, since it follows from the data presented above that $\alpha<1$.

Thus, it follows from the plots presented in Fig. 6 that $\alpha \ll 1$ at sufficiently large values of P_0 and that for clusters of sufficiently large size the intensity of the electron-beam-induced fluorescence per cluster molecule is much smaller than the analogous value for the free molecules. An estimate based on the stagnation parameters, the use of the similarity relations in Ref. 11, and the data from the electron diffraction measurements in Refs. 26 and 27 shows that the mean cluster size at the maximum value of P_0 is $N \approx 7000$. The cluster size is known to increase with increasing P_0.[11] Therefore, it also follows from Fig. 6 that the fluorescence yield of the clusters in the molecular beam decreases as their size increases. This finding is qualitatively consistent with the data in Refs. 14–17 on the fluorescence yield of CO_2 clusters in a molecular beam. We note that the only significant difference between our data and the data in Refs. 14–17 is in the energy of the exciting electron beam: 5.5 keV as opposed to 90 eV.

It was established in the preceding study[1] that the CO_2 clusters in a jet expansion emit at the wavelengths of the monomers with a fairly high efficiency per molecule ($\alpha \approx 0.5-0.7$) when they are excited by an electron beam with an energy of 14 keV. The size dependence is insignificant. For example, as follows from Ref. 1, the fluorescence yield of CO_2 clusters with an average size $\overline{N} \approx 600$ is equal to $\approx 0.5-0.7$ in a jet, but according to the plot in Fig. 6, the value for clusters of the same size in a molecular beam is much less than unity. In both cases the mean size was estimated from the stagnation parameters using the similarity relations and results in Ref. 27.

Thus, there is an inconsistency between the results in Ref. 1 and the present work. Let us try to account for it. It was shown in Ref. 1 for a jet expansion that a mechanism for the appearance of emission by a CO_2 cluster involving the ejection of an excited molecule from the cluster upon fragmentation operates when it is excited by electron impact. The same mechanism for the appearance of fluorescence from clusters apparently operates in a molecular beam. Just this mechanism accounts for the results of the investigations of the fluorescence of CO_2 clusters in a molecular beam in Refs. 14–17. It was established in Ref. 4 that the fluorescence spectrum of a cluster beam of nitrogen excited by electron impact does not differ from the fluorescence spectrum of the gas phase. This is also an argument in favor of the mechanism with fragmentation. In our opinion, the strong difference between the fluorescence yields of CO_2 clusters in a jet and a molecular beam is due to the change in the phase state of the clusters (they pass from the liquid to the solid state) and possibly to the change in their temperature.

Carbon dioxide clusters in a molecular beam have a crystalline structure when their size exceeds $\overline{N} \approx 50$, and their temperature falls in the range $T \approx 110-120$ K and does not depend on the stagnation parameters or the cluster size, if the molecular beam formed from a jet of the pure gas.[27] On the other hand, van der Waals clusters form in the liquid state in a jet, and only afterwards can they become solid as they cool in the jet as a result of heat exchange with the gas and evaporative cooling,[28] while in a molecular beam they are generally in the solid state due to evaporative cooling. This was confirmed experimentally in Ref. 29, where the kinetics of the phase transition of nitrogen clusters in a jet were investigated using coherent anti-Stokes Raman spectroscopy (CARS), as well as in Ref. 30, where similar investigations were carried out for carbon tetrachloride clusters using electron diffraction analysis in a jet. Thus, CO_2 clusters apparently form in the liquid state in a jet. Then, as they move away from the place where they formed, a phase transition begins in them, and in the molecular beam they exist in the solid state.

Therefore, the observed difference in the fluorescence yield can be described within the proposed process for the mechanism of the appearance of emission by clusters with fragmentation. In a jet expansion an excited molecular ion formed within a cluster has sufficient energy to leave the ''liquid'' (weakly bound) cluster upon fragmentation, as occurs. In a molecular beam this energy is insufficient for a cluster of the same size, since the cluster is in the solid state and additional energy is needed to deform the lattice. As a result, the ion remains within the cluster, and its electron excitation decays nonradiatively. However, if a molecular ion forms from the surface of a cluster, it can apparently detach itself from the cluster and make a contribution to the emission. Under the condition of equally probable positioning of the charge in the bulk of a cluster, its fluorescence yield should decrease as its size increases, since the ratio of the surface of the cluster to its volume decreases in inverse proportion to the size. Thus, the fluorescence yield of the clusters in a molecular beam dependence on the cluster size and is small for fairly large clusters. On the other hand, if such a dependence exists at all in a jet, it is significantly weaker, and the fluorescence yield depends primarily on the phase state and the temperature of the clusters.

Apart from these two reasons for the fluorescence yield to be smaller in a molecular beam than in a jet, there is one more reason. The mean cluster size is greater in a molecular beam than in a jet at the same backing pressure. This is due to the dissimilar thermal dispersion of clusters of different size in a molecular beam, which becomes enriched with heavy clusters.[31] However, in our opinion, this effect is not strong enough to cause such a radical difference in the fluorescence yield. For example, the size of Ar clusters is $\approx 50\%$ greater in a molecular beam than in a jet at the same backing pressure in the range of cluster sizes from 500 to 2000.[31]

It was pointed out above that our results are qualitatively consistent with the data from Refs. 14–17. Let us now try to compare them quantitatively. However, we first point out a methodical error which was made in Refs. 14–16 in calculating the fluorescence yield of the clusters α (the latter is an intermediate quantity, which was used in Refs. 14–16 to determine the emission probability of a molecule excited in a cluster φ after taking into account the excitation cross section σ: $\alpha \sim \varphi\sigma$). Expression (5) was used to calculate it. The value of the monomer density used was obtained in the following manner: the gas-dynamic parameters (primarily, the density of the molecules) in the jet expansion were first calculated, and then the usual model for the formation of a molecular beam[23,25] was used to determine the monomer

density in it. However, this cannot be done for the following reasons. First, the model of the formation of a molecular beam presumes the absence of perturbing factors. In Refs. 14–16 the beam formed under the conditions of a strong skimmer interaction. This has already been pointed out in describing Fig. 6. Second, it is known from the same model that

$$n_b = n_s(r^2/L^2)S^2,$$

where n_b is the density of the monomers in the molecular beam in the region of the detector, n_s is the density of the gas in the jet before the skimmer, r is the radius of the skimmer, L is the skimmer-detector distance, S is the velocity ratio in the jet in the region of the skimmer (equilibrium or ''perpendicular,'' depending on the flow regime), $S^2 = mV^2/2kT$, m is the mass of the monomer, k is Boltzmann's constant, and T is the temperature.

The calculations of the gas-dynamic parameters in a jet expansion performed in Refs. 14–16 were one-dimensional and did not take into account the lack of equilibrium between the monomers and clusters. This could have led to a large error in the value of the velocity ratio and thus in the calculated value of n_b. Apparently, taking this circumstance into account, Vostrikov *et al.*[14–16] have abandoned this procedure in their more recent work.[17,18] Therefore, the value for large CO_2 clusters in Ref. 17 differs from the data in Refs. 14–16 by an order of magnitude. Unfortunately, the specific method used to determine the monomer density in the cluster beam was not indicated in Refs. 17 and 18.

In our opinion, this problem can be solved using the time-of-flight method proposed in Refs. 8 and 13. If v_m and $v_{\rm cl}$ are additionally measured by the methods from Refs. 8–10, the fluorescence yield of the clusters can be correctly determined using expression (4).

Thus, only the primary data in Refs. 14–16 and the present investigation, particularly the optical signals obtained in the two cases, can be compared. As we have already pointed out, this possibility exists, because the experimental conditions are sufficiently similar. In Fig. 6 the corresponding optical signals for a nozzle–skimmer distance $x = 100$ mm are represented by curves *2* and *3* (Refs. 14–16, the visible and ultraviolet regions of the spectrum), as well as curve *5* (the present work). All three values ''merge'' with one another at a minimal pressure in the absence of condensation. As is seen from this figure, the data obtained in Refs. 14–16 are markedly greater in amplitude than the results of the measurements in the present work. In our opinion, such a situation is possible for the following reasons: first, because of the difference between the energies of the exciting electrons, and, second, because of the background illumination. In Refs. 14–16 the electron gun unit was located in the field of vision of the photomultiplier. This could have resulted in background illumination due to the cathodoluminescence caused by collisions of the electrons with parts of the electron gun unit. These electrons can be primary electrons scattered on clusters and secondary electrons formed in clusters, as well as other types of electrons. In this case the background cannot be ''cut off'' using modulation of the cluster beam and synchronous detection at the modulation frequency, as was done in Refs. 14–16. In view of the small amplitude of the effective signal of the molecular beam, this can be a serious problem. We note that a long-focus electron gun was used in the present work; therefore, there were no solid surfaces in the region of the electron beam in the field of vision of the optical system.

Let us now turn to the electron-beam-induced fluorescence diagnostics of molecular beams with CO_2 clusters. As follows from the present work and Refs. 14–17, the clusters make a contribution to the emission at the wavelengths of the monomers; however, the fluorescence yield decreases significantly as the cluster size increases, and it is very small for clusters of large size. The small fluorescence yield of the clusters in a molecular beam makes the use of the EBIF technique in it far more productive than in a jet.[1] Here the main contribution to the fluorescence is made by the monomers.

When the EBIF technique is used in cluster beams, we can take advantage of the fact that, roughly speaking, clusters do not emit, while monomers do. Thus, the problem of separating the monomer and cluster signals is solved. In time-of-flight measurements the signal obtained corresponds to the monomer component of the cluster beam. This permits measurement of the velocity distribution function of the monomers in a molecular beam with clusters and, therefore, in the jet expansion from which the beam formed.[9,10]

The measured distribution function can be used to determine the monomer component of the intensity of the cluster beam. If the total intensity of the cluster beam is recorded using an intensity sensor simultaneously with the time-of-flight measurements, the relative intensity of the monomers or clusters (the fraction of monomers or condensate in the flux) can be found, in analogy to what was done in Ref. 8.

Finally, the last example regarding the use of the EBIF technique in a cluster beam is the measurement of the velocity of the clusters according to the method proposed in Ref. 10. It was pointed out above that the obstacle placed after the electron beam can serve as a probe for the cluster component of a molecular beam.[9] Therefore, recording the time-of-flight signal with and without an obstacle permits determination of the velocity of the clusters.

CONCLUSIONS

The fluorescence of CO_2 clusters excited by electron impact in a molecular beam with clusters was investigated in the present work. For this purpose, the electron-beam-induced fluorescence was compared with the signal appearing in an electronic-fluorescence detector as a result of reflection of the cluster beam from an obstacle placed after the electron beam, as well as with the intensity of the cluster beam.

It has been established for CO_2 clusters in a molecular beam, as opposed to a jet expansion, that the fluorescence yield of the clusters decreases significantly with increasing cluster size and is very small for clusters of large size. This difference between the fluorescence properties of the clusters in a molecular beam and a jet is apparently attributable to the change in their phase state (they undergo a transition from

the liquid state to the solid state), and possibly to the change in the temperature of the clusters.

On the basis of the data obtained on the fluorescence of CO_2 clusters it has been included that the EBIF technique can be used in a molecular beam with clusters to measure the velocity and temperature of the monomers, as well as the velocity of the clusters within the time-of-flight method.

We thank P. A. Skovorodko for following this work and for some fruitful discussions. This work was performed with support from the Russian Interdisciplinary Scientific-Technical Program ''Fullerenes and Atomic Clusters.''

[1] S. Ya. Khmel' and R. G. Sharafutdinov, Zh. Tekh. Fiz. **67**(3), 43 (1997) [Tech. Phys. **42**, 291 (1997)].
[2] M. Faubel and E. R. Weiner, J. Chem. Phys. **75**, 641 (1981).
[3] A. E. Belikov, E. S. Voronel', Ya. Ya. Tomsons *et al.*, Zh. Prikl. Mekh. Tekh. Fiz., No. 2, 18 (1986).
[4] R. J. Holland, G. O. Xu, A. Robertson Jr. *et al.*, J. Chem. Phys. **88**, 7952 (1988).
[5] T. Moller, Z. Phys. D **20**, 1 (1991).
[6] E. Shalev, N. Ben-Horin, and J. Jortner, Chem. Phys. Lett. **177**, 161 (1990).
[7] J. Deckers and J. B. Fenn, Rev. Sci. Instrum., No. 1, 96 (1963).
[8] J. Cuvellier and A. Binet, Rev. Phys. Appl. **23**, 91 (1988).
[9] A. E. Belikov, S. Ya. Khmel', and R. G. Sharafutdinov, in *13th International Symposium on Molecular Beams. Book of Abstracts*, El Escorial, Madrid (1991), p. A10.
[10] S. Ya. Khmel' and R. G. Sharafutdinov, in *15th International Symposium on Molecular Beams. Book of Abstracts*, Berlin, 1993, pp. E8.1–E8.4.
[11] O. F. Hagena and W. Obert, J. Chem. Phys. **56**, 1793 (1972).
[12] N. G. Gorchakova, P. A. Skovorodko, and V. N. Yarygin, Inzh. Fiz. Zh. **49**(1), 5 (1985).
[13] J. P. Visticot, J. M. Mestdagh, C. Alcaraz *et al.*, J. Chem. Phys. **88**, 3081 (1988).
[14] A. A. Vostrikov, V. P. Gilyova, and D. Yu. Dubov, in *Rarefied Gas Dynamics. Proceedings of the 17th International Symposium*, Aachen, 1990, pp. 1197–1204.
[15] A. A. Vostrikov, V. P. Gilyova, and D. Yu. Dubov, Z. Phys. D. **20**, 205 (1991).
[16] A. A. Vostrikov, V. P. Gileva, and D. Yu. Dubov, Zh. Tekh. Fiz. **62**(1), 60 (1992) [Sov. Phys. Tech. Phys. **37**, 28 (1992)].
[17] A. A. Vostrikov, I. V. Samoilov, and D. Yu. Dubov, in *15th International Symposium on Molecular Beams. Book of Abstracts*, Berlin, 1993, pp. E.13.1–E.13.4.
[18] A. A. Vostrikov and V. P. Gileva, Pis'ma Zh. Tekh. Fiz. **20** (15), 40 (1994) [Tech. Phys. Lett. **20**, 625 (1994)].
[19] A. A. Bochkarev, E. G. Velikanov, A. K. Rebrov *et al.*, in *Experimental Methods in the Dynamics of Rarefied Gases* [in Russian], Novosibirsk, 1974, pp. 6–29.
[20] A. A. Bochkarev, V. A. Kosinov, A. K. Rebrov, and R. G. Sharafutdinov, *ibid.*, pp. 98–137.
[21] A. E. Beylich, Phys. Fluids. **14**, 898 (1971).
[22] A. E. Belikov and S. Ya. Khmel', Izv. Sib. Otd. Akad. Nauk SSSR. Ser. Tekh. Nauk (1), 84 (1990).
[23] J. B. Anderson, R. P. Andres, and J. B. Fenn, ''Supersonic nozzle beams,'' in *Molecular Beams (Advances in Chemical Physics, Vol. 10)*, J. Ross (ed.), Interscience, New York (1966), pp. 275–317 [Russ. transl., Mir, Moscow (1969), pp. 299–345].
[24] A. E. Belikov, S. Ya. Khmel', and R. G. Sharafutdinov, in *Rarefied Gas Dynamics. Proceedings of the 17th International Symposium*, Aachen, 1990, pp. 1220–1226.
[25] A. E. Zarvin and R. G. Sharafutdinov, Zh. Prikl. Mekh. Tekh. Fiz. (6), 107 (1979).
[26] Élkeleĭ and Nat, Prib. Nauchn. Issled., No. 3, 43 (1960) [Russ. transl. of Rev. Sci. Instrum. (different date)].
[27] G. Torchet, J. Farges, M. de Ferandy, and B. Raoult, Ann. Phys. (Paris) **14**, 245 (1989).
[28] L. S. Bartell, L. Harsanyi, and E. J. Valente, J. Phys. Chem. **93**, 6201 (1989).
[29] R. D. Beck, M. F. Hineman, and J. W. Nibler, J. Chem. Phys. **92**, 7068 (1990).
[30] L. S. Bartell and T. S. Dibble, J. Phys. Chem. **95**, 1159 (1991).
[31] S. V. Nedosekova and P. A. Skovorodko, in *Dynamics of Rarefied Gases. Proceedings of the 10th All-Union Conference* [in Russian], Moscow, 1991.

Translated by P. Shelnitz

Thermoluminescence of corundum containing anion defects following ultraviolet laser and x irradiation

V. S. Kortov, A. I. Syurdo, and F. F. Sharafutdinov

Ural State Technical University, 620002 Ekaterinburg, Russia
(Submitted February 14, 1996)
Zh. Tekh. Fiz. **67**, 72–76 (July 1997)

The thermoluminescence of single crystals of corundum containing anion defects following x-ray and laser excitation is investigated. Its features in the luminescence bands of F, F^+, and Cr^{3+} centers are studied. Synchronous measurements of the thermoluminescence and thermally stimulated exoelectron emission are performed by the fractional glow technique following x-ray and laser excitation of the samples. It follows from the results obtained that several traps are active in the temperature range of the principal dosimetric peak (400–500 K). The spectral sensitivity curve contains maxima corresponding to absorption bands of F, F^+, and Al_i^+ centers. A possible mechanism for the recombination luminescence of F centers is discussed. It is found that the material exhibits high sensitivity to small doses of ultraviolet laser radiation.
 [S1063-7842(97)01307-X]

INTRODUCTION

The role of optical, including laser, radiation in technological processes, medical practice, and scientific research has increased considerably in the recent period. The broad spectral range and the great diversity of the temporal and power characteristics of the radiation employed make it difficult to measure them by the most widely used electronic methods and have stimulated the development of miniature integrating detectors. Various physical effects can be used to detect optical radiation using such detectors. Measurements of high-intensity optical fluxes can be performed by utilizing the photochromic effect, photorefraction, and electron spin resonance. The detection of small doses (scattered radiation presents a great danger in practical work) requires highly sensitive detector materials. Since one of the most sensitive and relatively simple methods for obtaining dosimetric information is thermally stimulated luminescence (TSL), it is preferable to select one of the crystal phosphors employed in TSL dosimetry as the detector material.

During the last 20–25 years aluminum oxide has been repeatedly proposed as a sensitive medium for ultraviolet (UV) radiation.[1,2] It was noted in Ref. 3 that detectors based on corundum containing anion defects can be used in the dosimetry of UV radiation. Some experimental evidence supporting this was obtained in Ref. 4.

The purpose of the present work is to study the features of the TSL and thermally stimulated exoelectron emission (TSEE) of corundum containing anion defects after optical (including laser) and x-ray excitation, as well as its dosimetric properties on UV laser irradiation.

SAMPLES AND MEASUREMENT METHOD

Single crystals of α-Al_2O_3 without specially introduced dopants were investigated in this work. Oxygen vacancies were created by growing the crystals in a strong reducing medium. Their concentration in the samples investigated amounted to $\sim 10^{18}$ cm^{-3}.

X irradiation was carried out at room temperature (Co anode, 50 kV, 10 mA). An LGI-21 nitrogen laser ($\lambda = 337$ nm) and a DDS-30 deuterium lamp were employed for optical excitation.

The spectral distribution of the TSL and its excitation spectra was measured by previously described methods[5] using an SDD-2 system and an MSD-1 monochromator. The TSL excitation spectra were corrected using a fluorescent source of quanta. The luminescence spectra were corrected using a standard lamp. The thermoluminescence measurements were performed with a constant heating rate on a specially adapted ENDOVA-50 system. In the comparative experiments the heating rate was 2 K/s. An FÉU-106 (or FÉU-39A) photomultiplier operating in the photon-counting mode served as the radiation detector. The luminescence bands at 1.8, 3.0, and 3.8 eV in the TSL spectrum were isolated by interference filters with a maximum transmission equal to 30–35% and a half-width of the transmission band equal to 10–15 nm, as well as by an MSD-1 monochromator. Synchronous TSL and TSEE measurements were performed using the fractional glow technique on a previously described experimental apparatus.[6]

RESULTS

Crystals of corundum containing anion defects exhibit TSL with a dominant peak at 460 K after laser irradiation at 337 nm. Its intensity and total light yield correlate with the dose of laser radiation. The spectral distribution of the TSL peak at 460 K is predominated by violet (3.0 eV) and red (1.8 eV) luminescence. The luminescence at 3.0 eV is caused by F centers, and the luminescence at 1.8 eV is caused by the chromium impurity.[1,7]

The spectral composition of the principal dosimetric peak at 430 K after x-ray excitation is dominated by the luminescence of F centers. In addition, there is luminescence from F^+ and Cr^{3+} centers (3.8 eV).[7] Figure 1 presents the TSL curves recorded in each of these luminescence bands of α-Al_2O_3 after x-ray excitation (curves *1*, *2*, and *4*) and in the

 1063-7842/97/070783-05$10.00

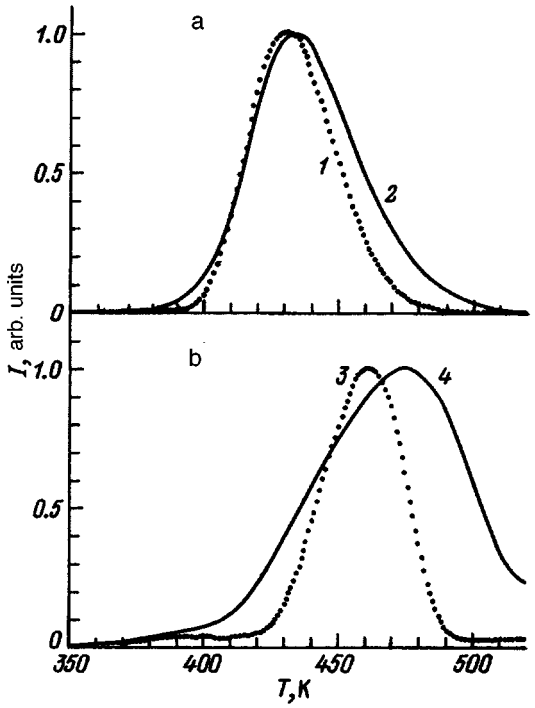

FIG. 1. Thermoluminescence curves of α-Al_2O_3 in the temperature range of the principal dosimetric peak for a constant heating rate of 2 K/s. *1, 2, 4* — x-ray excitation, *3* — laser irradiation. Position of the band, eV: *1, 3* — 3.0; *2* — 3.8; *4* — 1.8.

band at 3.02 eV after excitation by laser radiation (curve *3*). The TSL peaks of the x-irradiated sample in the bands at 3.0 eV ($XTSL_{3.0}$) and 3.8 eV ($XTSL_{3.8}$) exhibit second-order kinetics (the geometric factor $\mu=0.56$). The shape of the TSL curves recorded in the band at 1.8 eV after x irradiation ($XTSL_{1.8}$) and in the band at 3.0 eV after photoexcitation ($PTSL_{3.0}$), including laser excitation, corresponds to first-order kinetics ($\mu=0.42-0.46$). It should be noted that at equal total emitted light yields the $PTSL_{3.0}$ peaks have a temperature at the maximum that is 10–25 K higher than the corresponding value for the $XTSL_{3.0}$ peaks, and the $PTSL_{3.8}$ yield is considerably lower than the $XTSL_{3.8}$ yield.

Figure 2 presents plots of the temperature dependence of the maximum $TSL_{3.0}$ (curve *1*) and TSEE (curve *3*) intensities in cycles obtained by the fractional glow technique, as well as plots of the temperature dependence of the mean activation energy for TSEE ($\langle E_{TSEE}\rangle$, curve *4*) and TSL ($\langle E_{TSL}\rangle$, curve *2*) after excitation of the sample by x radiation. The TSEE peak is shifted 10 K toward higher temperatures relative to the TSL peak, $\langle E_{TSEE}\rangle$ equals 1.56 eV and remains constant over the entire temperature range of the peak, and $\langle E_{TSL}\rangle$ reaches a maximum at $T=385-395$ K (1.56 eV) and decreases to 0.57 eV at $T=455$ K. In the temperature range 440–450 K the $\langle E_{TSL}\rangle(T)$ curve has a segment with a practically constant activation energy ($\langle E_{TSL}\rangle\approx 0.7$ eV). It should be noted that at comparable TSL intensities the samples irradiated by laser radiation have lower exoelectron emission activity than the x-irradiated samples, making it difficult to perform measurements using the fractional glow technique to determine the kinetic parameters. The $\langle E_{TSL}\rangle(T)$ curve recorded after laser irradiation is

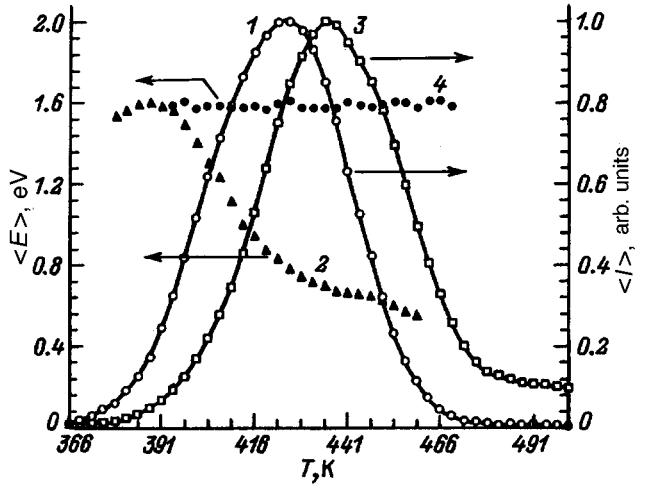

FIG. 2. Temperature dependence of the maximum intensity and the mean activation energy in fractional glow cycles following x-ray excitation of the sample.

similar to the $\langle E_{TSL}\rangle(T)$ curve for the x-irradiated sample, but the mean activation energy decreases smoothly on the high-temperature side of the peak at 440–450 K (Fig. 3, curve *2*).

An examination of the reproducibility of the results of the TSL measurements with a constant heating rate revealed a decrease in the intensity of the peak (460 K) for the samples subjected to laser irradiation. Figure 4 shows the decrease in the $PTSL_{3.0}$ yield in successive measurement cycles (the dose of laser radiation was 250 mJ/cm^2). After eight cycles, the intensity of the PTSL peak decreased more than 20-fold. The sensitivity was partially restored after the samples were annealed at 900–1000 K. However, complete restoration of the sensitivity is possible, if the samples are subjected to x irradiation, which leads to filling of all the trapping centers, before each photoexcitation cycle. The shallow traps are emptied when the samples are heated to 800 K. During subsequent photoexcitation, the charge carriers migrate from the trapping centers that have a large ther-

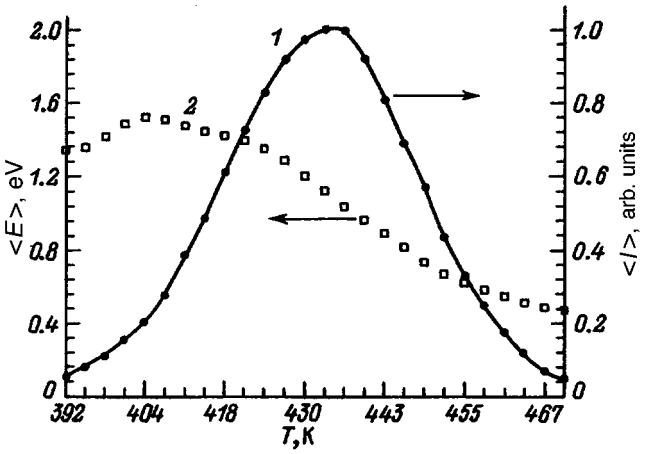

FIG. 3. Temperature dependence of the maximum intensity $\langle I\rangle$ (curve *1*) and the mean activation energy $\langle E\rangle$ (curve *2*) for TSL in the 3.0 eV band in fractional glow cycles following laser excitation of the sample.

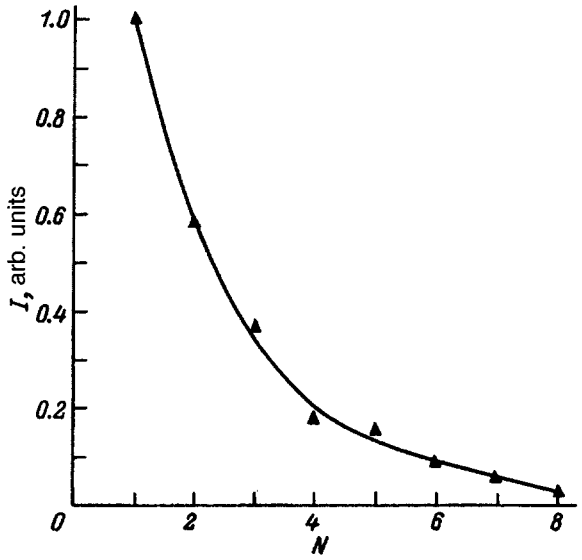

FIG. 4. Dependence of the TSL intensity (I) following laser excitation on the number of measurement cycles (N).

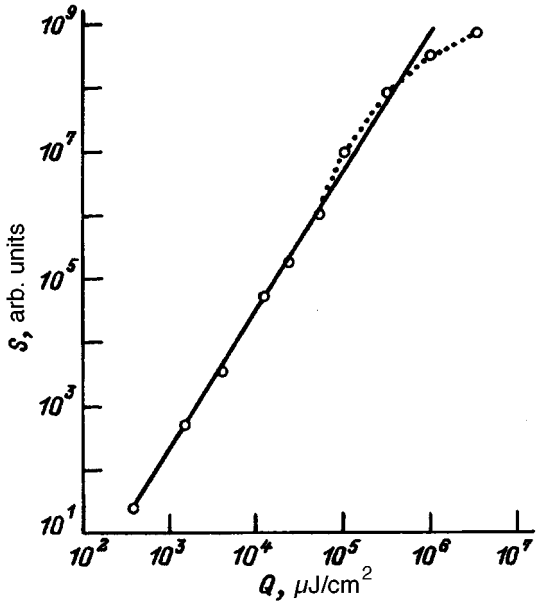

FIG. 5. Dependence of the emitted total light yield (S) on the dose of laser radiation (Q).

mal activation energy to the levels responsible for the principal dosimetric peak of corundum containing anion defects.

To determine the x-irradiation dose needed to completely fill the deep traps, we investigated the dependence of the TSL yield of laser-irradiated (250 J/cm^2) samples on the preliminary x-ray excitation dose. It was found that the samples preliminarily irradiated by a dose greater than 10^5 R have the greatest sensitivity.

The dose dependence was investigated on samples prepared in accordance with the conditions indicated above. Scattered laser radiation was used to measure doses smaller than 25 mJ/cm^2. In the range of doses of laser radiation at 337 nm from 0.4 to 1000 mJ/cm^2, the intensity and the total light yield in the TSL maximum at 460 K were found to be linearly dependent on the dose in log–log coordinates (Fig. 5). This finding attests to the fundamental possibility of using corundum crystals containing anion defects to detect laser radiation in the UV range.

To ascertain the functional possibilities of the proposed detectors of UV laser radiation, we studied the dependence of the TSL yield at 460 K on the wavelength of the UV light for a constant dose of UV irradiation (Fig. 6). A significant increase in the sensitivity of the detectors was observed at wavelengths smaller than 280 nm. The features recorded on the TSL spectral sensitivity curve at 205, 225, 255, and 305 nm coincide with the maxima of the absorption bands of corundum containing anion defects and are produced by F(205 nm), F^+(225, 255 nm), and Al_i^+(305 nm) centers.[1,8]

DISCUSSION

The plots of the dependence of the $XTSL_{1.8}$ peak on the x-irradiation dose and the heating rate, as well as the position of the maximum and the shape of the peak, differ from the analogous dependences of the $XTSL_{3.0}$ and $XTSL_{3.8}$ peaks. It can be theorized that another trap is responsible for this peak. In addition, it was discovered that after several hundred cycles of x irradiation, thermal annealing, laser excitation, and recording of the thermoluminescence, there is a significant decrease in the concentration of F centers (according to optical absorption measurements) along with a decrease in the TSL yield in the band at 3.0 eV, which is possibly attributable to destruction of the F centers by the electromagnetic field of the laser radiation. Displacement of the temperature of the $XTSL_{1.8}$ maximum to 495–500 K is observed. The trap responsible for the $XTSL_{1.8}$ peak is apparently a V^2 center,[9,10] whose thermal activation energy decreases, if another defect (presumably an F center) is located nearby.

The nature of the traps responsible for the principal dosimetric peak at 430 K cannot yet be considered precisely

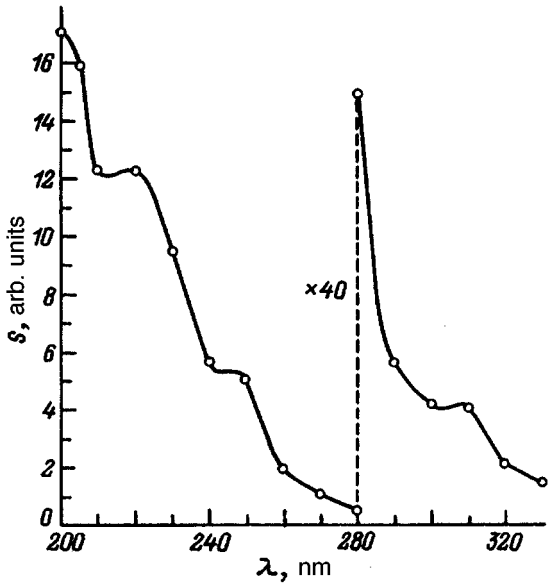

FIG. 6. Dependence of the emitted total light yield (S) on the wavelength of the exciting light (λ).

established. Additional difficulties are caused by the anomalous behavior of the mean activation energy and the frequency factor.

The hole nature of the trapping centers postulated in the model of a ''two-step'' Auger process[10] is not consistent with the experimental results indicating photoinduced migration of the charge carriers.[11]

Gimadova *et al.*[12] proposed a model associated with structural changes induced by thermal fluctuations, in which the recombination and trapping centers are single complexes and charge transfer takes place without the participation of bands. Within this model there are difficulties in interpreting the differences in the temperature dependences of $\langle E_{\mathrm{TSEE}} \rangle$ and $\langle E_{\mathrm{TSL}} \rangle$, the shapes of the $\mathrm{PTSL}_{3.0}$ and $\mathrm{XTSL}_{3.0}$ peaks, and the temperature positions of their maxima.

Anomalously high values of the activation energy E and of the frequency factor s are observed not only for corundum containing anion defects. It was shown in Refs. 13 and 14, where samples of LiF:Mg and LiF:Ti were investigated, that recombination followed by the trapping of charge carriers of opposite sign with simultaneous emptying of electron and hole traps can, in principle, lead to overestimated values of E and s. It can be presumed that conditions under which the recombination process can be described on the basis of such a model are created in corundum containing anion defects in the temperature range of the principal dosimetric peak.

The interaction of an F center with a hole leads to the formation of an F^+ center in the ground or excited state (F^+*). An F^+* center has a mean lifetime of several nanoseconds, and it is unlikely that it can recombine with an electron during that time. Conversely, if an F center captures an electron, an F^- center forms. The existence of such a center has been postulated in experimental research,[11] and it was concluded in the theoretical study in Ref. 15 that the additional electron occupies a local level located at a depth of 1 eV or less from the bottom of the conduction band. At 400–500 K such a center is probably thermally unstable. An estimate of the mean lifetime of an F^- center using the formula

$$\tau = s^{-1} \cdot \exp(E/kT) \tag{1}$$

for $E = 0.5$ eV, $s = 10^{12}\mathrm{s}^{-1}$, and $T = 450$ K gives $\tau \approx 0.4$ μs. If over the course of this time the F^- center captures a hole, the reaction $F^- + e^+ \rightarrow F^* \rightarrow F + h\nu$ (3.0 eV) will take place. Recombination of one of the s electrons of the F^- center with the incoming hole is assumed here. The possibility of a similar interaction of the F* center with a hole with the formation of an F^+* center has not been ruled out. Another recombination mechanism, which leads to the formation of an F* center is the capture of an electron by an F^+ center. Obviously, the reaction $F + e^- + e^+ \rightarrow F^- + e^+ \rightarrow F^*$ cannot alter the concentration of F centers, while the other mechanism $F^+ + e^- \rightarrow F^*$ should lead to an increase in the number of F centers and a decrease in the number of F^+ centers during thermoluminescence. However, measurements of the corresponding optical absorption bands during isochronous annealing[11] show that this does not occur at 400–500 K. Therefore, it can be assumed that the competitive reaction $F + e^+ \rightarrow F^+ * \rightarrow F^+ + h\nu$ (3.8 eV) takes place.

The nonelementary dependence of the TSL yield on the wavelength of the exciting light (Fig. 6), as well as the complicated spectral composition of the PTSL peak, indicate that both electrons and holes are released when a sample preliminarily excited by x radiation is irradiated with light. Apart from the release of the charge carriers from the deep traps, the F centers undergo photoionization, which causes the TSL spectral sensitivity to rise as the energy of the quanta of the exciting light approach the absorption maximum of the F centers. The features found near the absorption maxima of the F^+ centers can be attributed to the conversion of the F^+ centers into F centers, which is accompanied by the release of the holes[16] that are subsequently captured in the traps responsible for the principal dosimetric peak. The maximum at 305 nm is attributed to optical destruction of the corresponding color centers.[17]

The decrease in the emitted total light yield in the $\mathrm{PTSL}_{3.8}$ peak in comparison to the $\mathrm{XTSL}_{3.8}$ peak with identical PTSL and XTSL yields in the band at 3.0 eV can be attributed to the fact that the ratio between the concentrations of holes and electrons captured in traps upon the photoexcitation of α-Al_2O_3 is smaller than the ratio for x irradiation.

The shift of the $\mathrm{PTSL}_{3.0}$ peaks toward higher temperatures in comparison with the $\mathrm{XTSL}_{3.0}$ peaks (with equality between the emitted total light yields) is attributed mainly to the competitive influence of the deeper traps, which are emptied by the laser radiation. Additional experimental confirmation is provided by the fact that the $\mathrm{XTSL}_{3.0}$ peaks can have maxima at higher temperatures, if the sample is subjected to laser irradiation before x irradiation.

CONCLUSIONS

The following conclusions can be drawn on the basis of the results of the investigations performed:

1. Detectors for picking up optical, including laser UV, radiation (in both the direct and scattered forms) can be developed on the basis of single crystals of corundum containing anion defects.

2. The principal dosimetric TSL peak at 430 K is nonelementary. It is associated with the emptying of at least three traps.

3. When the samples are optically excited, the same traps which cause thermoluminescence after x-ray excitation are responsible for the TSL peak with a maximum at 450 K, but their relative populations are different, and the competitive influence of the deeper trapping centers increases.

4. A number of experimental findings can be explained, if the possibility of the capture of a charge carrier of opposite sign is allowed.

[1] G. P. Summers, Radiat. Prot. Dosim. **8**, 69 (1984).
[2] W. G. Buckman, Health Phys. **22**, 402 (1972).
[3] M. S. Akselrod, V. S. Kortov, D. J. Kravetsky *et al.*, Radiat. Prot. Dosim. **32**, 15 (1990).
[4] L. Oster, D. Weiss, and N. Kristianpoller, J. Phys. D. **27**, 1 (1994).
[5] L. V. Levshin and A. M. Saletskiĭ, *Luminescence and it Measurement* [in Russian], MGU, Moscow (1989).
[6] V. I. Kirpa, A. S. Kuz'minykh, and V. V. Popov, in *Radiation-Stimulated Phenomena in Solids. An Interinstitute Collection, No. 5* [in Russian], Sverdlovsk, 1983, pp. 33–38.
[7] J. H. Crawford Jr., Nucl. Instrum. Methods Phys. Res. B **229**, 159 (1984).

[8] M. J. Springis and J. A. Valbis, Phys. Status Solidi B **123**, 335 (1984).
[9] K. H. Lee, G. E. Holmberg, and J. H. Crawford Jr., Phys. Status Solidi A **39**, 669 (1977).
[10] M. S. Aksel'rod, V. S. Kortov, and I. I. Mil'man, Ukr. Fiz. Zh. **28**, 1053 (1983).
[11] A. S. Portnyagin, Candidate's Dissertation, Sverdlovsk (1989).
[12] T. I. Gimadova, T. S. Bessonova, I. A. Tale *et al.*, Radiat. Prot. Dosim. **33**, 47 (1990).
[13] F. G. Fairchild, P. L. Mattern, K. Lengweiler *et al.*, J. Appl. Phys. **49**, 4523 (1978).
[14] V. S. Kortov, T. S. Bessonova, M. S. Akselrod *et al.*, Phys. Status Solidi A **87**, 629 (1985).
[15] A. Stashans, E. Kotomin, and J.-L. Calais, Phys. Rev. B **49**, 14 854 (1994).
[16] E. A. Kotomin, A. I. Popov, and A. Stashans, J. Phys.: Condens. Matter. **6**, L569 (1994).
[17] V. I. Baryshnikov, E. F. Martynovich, L. I. Shchepina *et al.*, Opt. Spektrosk. **64**, 455 (1988) [Opt. Spectrosc. (USSR) **64**, 273 (1988)].

Translated by P. Shelnitz

Amplification of monochromatic short-wavelength radiation during the stochastic deceleration of a relativistic electron stream in an incoherent pump field

Ya. L. Bogomolov, N. S. Ginzburg, and E. R. Golubyatnikova

Institute of Applied Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia
(Submitted March 5, 1996)
Zh. Tekh. Fiz. **67**, 77–81 (July 1997)

The use of incoherent multiwave pump radiation or randomly varying magnetostatic fields (stochastic undulators) for improving the energy conversion efficiency in free-electron lasers based on stimulated wave scattering and the stimulated undulator emission of relativistic electron beams is proposed. It is shown within the quasilinear approximation that the electronic efficiency increases in proportion to the width of the pump spectrum due to enrichment of the spectrum of combination waves which are synchronous with the electron beam and realization of a mechanism of stochastic particle deceleration when the signal wave is monochromatic. At the same time, the efficiency scarcely depends on the spread of the beam parameters, making the use of the method promising for improving the efficiency of free-electron lasers powered by intense relativistic electron beams. © *1997 American Institute of Physics.* [S1063-7842(97)01407-4]

INTRODUCTION

The use of coherent pump radiation to heat a plasma during stimulated scattering was proposed in Refs. 1 and 2. In this case the increase in the temperature of the electronic component is proportional to the width of the spectrum of scattered incoherent radiation. The present work (see also Ref. 3) examines a situation, which, in a certain sense, is the opposite of the case considered in Refs. 1 and 2. More specifically, the use of low-frequency incoherent pump radiation, which is backscattered on a relativistic electron beam (REB) under the conditions of the inverse Compton effect, is proposed for effectively amplifying monochromatic rf radiation. The advantage of incoherent pump radiation over the traditional monochromatic pump radiation used for such systems [which are called Compton lasers or free-electron lasers (FEL)[4–7]] is the expansion of the spectrum of combination waves that are synchronous with the electron stream and the consequent realization of a mechanism of diffusive stochastic deceleration of the beam particles. As a result, the efficiency of the conversion of the beam energy into the energy of short-wavelength scattered radiation is proportional to the width of the pump spectrum and can significantly exceed the values realized in the case of monochromatic pump radiation. It is of fundamental importance that the conversion efficiency scarcely depends on the width of the translational velocity distribution function of the electrons. The lack of criticality toward the parameter spreads renders the use of incoherent pump radiation plausible, first of all, for FELs whose power source is an intense REB formed by exploding cathodes. Having considerable powers and current densities, such beams have large parameter spreads, which generally preclude their use in the efficient generation of short-wavelength radiation.

Attention should also be focused on the following aspect of the problem. It is generally known (see, for example, Refs. 4 and 5) that to within effects associated with depletion of the pump radiation, the physical processes in FELs based on stimulated wave scattering and stimulated undulator emission are identical, since in the accompanying reference system each relativistic electron perceives a periodic (undulator) magnetic field as an electromagnetic pump wave. Developing this analogy, we arrive at the conclusion that in the case of undulator radiation the randomly periodic magnetic field, whose use has the advantages mentioned above, can serve as the incoherent pump radiation.[1)]

In the present work a quasilinear approximation is used to describe the amplification of a monochromatic wave during the scattering of incoherent pump radiation on an REB. The characteristic lengths of the interaction space, in which the stochastic electron deceleration process develops and a sufficiently high (up to 20%) efficiency is realized in converting the energy of the REB into the energy of short-wavelength (particularly submillimeter) radiation, are determined using numerical simulation.

BASIC EQUATIONS

Let a monochromatic signal wave propagating together with an electron stream be assigned by the vector potential

$$\mathbf{A}_s = \mathrm{Re}[\mathbf{x}_0 A_s(z)\exp(i(\omega_s t - k_s z))].$$

In the amplification scheme under investigation the frequency of this wave is assigned by an external source. The pump field has the form of a set of discrete lines with phases that are not correlated with one another. Thus,

$$\mathbf{A}_i = \mathrm{Re}\left[\sum_{n=0}^{\infty} \mathbf{x}_0 A_{in}(z)\exp(i(\omega_{in} t + k_{in} z))\right],$$

the distance between the individual harmonics δk_i being much smaller than the total width of the packet Δk_i. The averaged motion of the relativistic electrons in the signal and pump fields can be described using the kinetic equation for the one-dimensional distribution function

 1063-7842/97/070788-05$10.00

$$\frac{\partial f}{\partial t}+u_z\frac{\partial f}{\partial z}+F_z\frac{\partial f}{\partial p_z}=0, \tag{1}$$

where

$$F_z=-\frac{e^2}{2\varepsilon}\mathrm{Re}\left[\sum_{n=0}^{\infty} ik_{cn}A_sA_{in}^*(z)\exp(i(\omega_{cn}t-k_{cn}z))\right] \tag{2}$$

is the ponderomotive force, $k_{cn}=k_s+k_{in}$ and $\omega_{cn}=\omega_s-\omega_{in}$ are the wave numbers and frequencies of the ensemble of combination waves, and $p_z=mv_z\gamma$ and $\varepsilon=mc^2\gamma$ are the longitudinal momentum and energy of the electrons.[2)]

After averaging over the ensemble of combination waves, which is traditional for the quasilinear approximation,[9–11] the equation for the slowly varying distribution function $f(z,p)$ takes on the form

$$v_z\frac{\partial\overline{f}}{\partial z}=\frac{\partial}{\partial p_z}\left(D(z,p_z)\frac{\partial\overline{f}}{\partial p_z}\right), \quad \overline{f}|_{z=0}=\overline{f^0}(p_z) \tag{3}$$

with the diffusion coefficient

$$D=\left.\frac{8\pi^3e^4c^2k_c^2I_sI_i}{\omega_s^2\omega_i^2(v_z+c)\varepsilon^2}\right|_{k_i=k_s(c-v_z)/(c+v_z)}.$$

Here $\overline{f^0}(p_z)$ is the initial longitudinal momentum distribution function of the electrons, $I_s=|A_s|^2\omega_s^2/8\pi c$ is the intensity of the signal wave, and $I_i=|A_i|^2\omega_i^2/8\pi c\,\delta k$ is the spectral intensity of the pump radiation. The variations of these quantities are described by the equations

$$\frac{dI_s}{dz}=\frac{\pi^2e^2\omega_p^2}{mc^5}\frac{I_s}{k_s}\int_0^{\infty}\frac{I_i}{\omega_i^2}\left(\varepsilon\frac{\partial\overline{f}}{\partial p_z}\right)\Bigg|_{v_z=v_c}dk_i, \quad I_s|_{z=0}=I_s^0, \tag{4}$$

$$\frac{dI_i}{dz}=\frac{\pi^2e^2\omega_p^2}{mc^5}\cdot\frac{I_i}{k_i}\frac{I_s}{\omega_s^2}\left(\varepsilon\frac{\partial\overline{f}}{\partial p_z}\right)\Bigg|_{v_z=v_c}, \quad I_i|_{z=L}=I_i^0, \tag{5}$$

where $v_c=\omega_c/k_c$ is the phase velocity of the combination waves, $\omega_p=(4\pi e^2N_0/m)^{1/2}$ is the plasma frequency, N_0 is the electron flux density, and L is the length of the scattering region.

The system of equations (3)–(5) has the integrals

$$K-I_s+\int_0^{\infty}I_i dk_i=\text{const}, \tag{6}$$

$$\int_0^{\infty}\frac{I_i dk_i}{k_i}-\frac{I_s}{k_s}=\text{const}, \tag{7}$$

which are, respectively, the conservation laws of the energy and the number of quanta in a scattering process. Here $K=N_0\int_0^{\infty}v_z\varepsilon\,\overline{f}\,dp_z$ is the kinetic energy flux of the electrons. As follows from Eqs. (6) and (7), when the frequency change is large ($\hbar\omega_s\gg\hbar\omega_i$), the energy supplied to the radiation is drawn mainly from the electron beam ($\Delta I_s\simeq\Delta K$), and the depletion of the pump radiation (when its intensity is sufficiently large) can be neglected. In the approximation of fixed pump radiation in the ultrarelativistic limit ($\gamma\gg1$, $p_z\simeq mc\gamma$) the system of equations (4) and (5) takes on the following form after passage to dimensionless variables:

$$\frac{\partial F}{\partial z}=\frac{\partial}{\partial\gamma}\left(\widetilde{D}\,\frac{\partial F}{\partial\gamma}\right), \quad F|_{z=0}=F_0(\gamma), \tag{8}$$

$$\frac{\partial P_s}{\partial z}=\frac{\pi\omega_p^2}{2\omega_s^2}\int_1^{\infty}\widetilde{D}\,\frac{\partial F}{\partial\gamma}d\gamma, \quad P_s|_{z=0}=P_s^0, \tag{9}$$

where $\widetilde{D}=4P_sP_ik_s^2\gamma^2/\pi$, $F(\gamma,z)=\overline{f}\cdot mc$ is the electron energy distribution function, $P_s=I_s\lambda_s^2/P_*$ is the power of the signal wave in absolute power units ($P_*=m^2c^5/e^2=8.7$ GW) passing through an area equal to λ_s^2, and $P_i=I_i\lambda_s^2/P_*$ is the spectral power of the pump radiation in the same normalization.

NUMERICAL SIMULATION RESULTS

The initial electron energy distribution function was assigned by the expression

$$\hat{F}^0=\frac{2}{\delta}\cos^2\left[(u-1)\frac{\pi}{\delta}\right] \tag{10}$$

in the interval $[-\delta/2,\delta/2]$ in the vicinity of the energy of the central fraction $\overline{\gamma}$, where $\delta=\Delta\gamma/\overline{\gamma}$ is the electron energy spread, and $u=\gamma/\overline{\gamma}$ is the normalized energy variable (it is assumed that $\int_1^{\infty}Fd\gamma=\int_0^{\infty}\hat{F}du=1$).

The spectral distribution of the pump field was approximated by the Gaussian line shape: $I_i=I_i^{\mathrm{tot}}\cdot g(k_i)$, where I_t^{tot} is the total (integrated) intensity of the pump radiation, and

$$g(k_i)=\frac{1}{\sqrt{\pi}\Delta k_i}\exp\left(-\frac{(k_i-\overline{k_i})^2}{\Delta k_i^2}\right).$$

We rewrite the normalized function describing the spectral distribution of the pump field in the dimensionless variables used in the form

$$g(u)=\frac{1}{\sqrt{\pi}\sigma}\exp\left[-\left(\frac{u_i^2}{u^2}-1\right)^2/\sigma^2\right],$$

where $\sigma=\Delta k_i/\overline{k_i}$ is the width of the pump spectrum, and $u_i=\gamma_i/\overline{\gamma}$ ($\gamma_i=\sqrt{k_s/4k_i}$) characterizes the center of the spectral band of the pump field in energy space.

To carry out the numerical simulation with consideration of energy conservation law (6), it is convenient to transform Eqs. (8) and (9) into a single diffusion equation containing a minimal number of independent parameters:

$$\frac{\partial\hat{F}}{\partial\zeta}=\frac{\partial}{\partial u}\left((\eta+p_s^0)g(u)u^2\frac{\partial\hat{F}}{\partial u}\right), \tag{11}$$

where

$$\zeta=\frac{\omega_p^2\lambda_s\overline{\gamma}}{4\pi c^2}\frac{\lambda_s^2I_s^{\mathrm{tot}}}{P_*^2}zp_s^0=\frac{2P_s^0\omega_s^2}{\pi\omega_p^2\overline{\gamma}},$$

$$\eta=1-\int_0^{\infty}u\hat{F}du \tag{12}$$

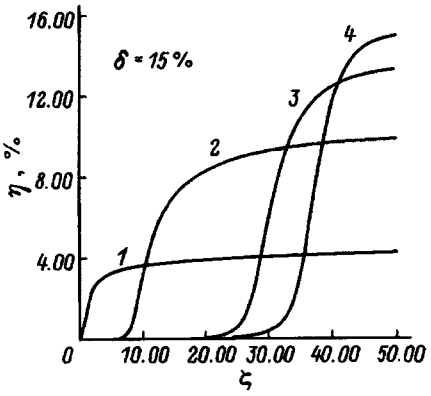

FIG. 1. Dependence of the electronic efficiency η on the length of the scattering region ζ for an initial particle energy spread $\delta=15\%$. The shift between the center of the spectral band of the pump radiation and the center of the distribution function of the particles $\varepsilon=1-u_i$ is optimal; σ, %: *1* — 10, *2* — 20, *3* — 30, *4* — 40; ε: *1* — 0.07, *2* — 0.15, *3* — 0.22, *4* — 0.28.

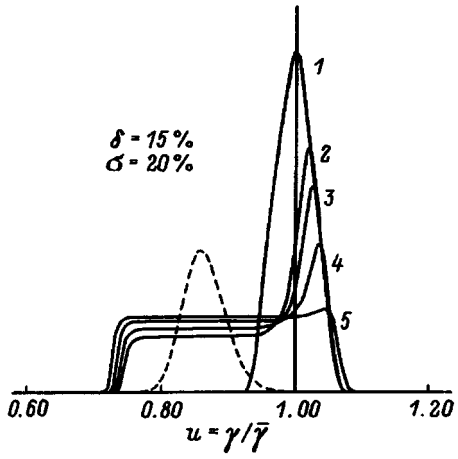

FIG. 3. Evolution of the distribution function of the electrons; $\delta=15\%$; $\sigma=20\%$; ζ: *1* — 0, *2* — 10, *3* — 12, *4* — 18, *5* — 20.

is the electronic efficiency of the conversion of the REB energy into signal wave energy.

The results of the numerical simulation presented in Figs. 1 and 2 show that the energy-conversion efficiency increases as the characteristic width σ of the pump spectrum increases. Thus, when the relative width of the pump spectrum is 40%, the maximum efficiency reaches 15%. At the same time, an increase in the width of the pump spectrum leads to a drop in the growth rates, and, accordingly, the length of the interaction space, in which the maximum efficiency is achieved, increases. It is important to stress that the results of the numerical simulation confirm the assumption made above that the method for increasing the efficiency under investigation is not critical toward the quality of the electron beam over a broad range. As is seen from a comparison of Figs. 1 and 2, the maximum efficiency scarcely depends on δ, which characterizes the spread of the parameters of the electron stream. We also note that in the stochastic deceleration regime the efficiency is very weakly dependent on the integrated intensity of the pump radiation and the electron flux density. In fact, gain saturation is caused by the formation of a plateau on the electron distribution function (Fig. 3), whose width and, therefore, maximum efficiency

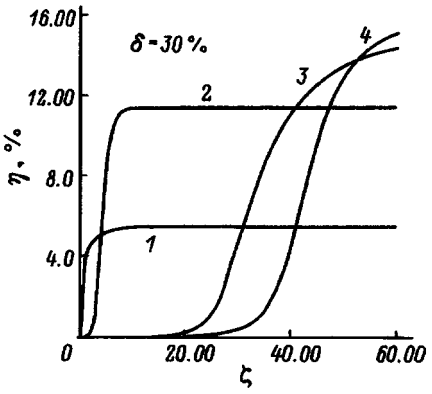

FIG. 2. Same as in Fig. 1; $\delta=30\%$; σ, %: *1* — 20, *2* — 30, *3* — 40, *4* — 50; ε: *1* — 0.1, *2* — 0.23, *3* — 0.3, *4* — 0.36.

are determined by the width of the pump spectrum, as well as by the distance (in energy space) between the center of the pump field line and the center of the particle distribution function. The pump intensity and the flux density mainly influence the growth rate and the length of the scattering region.

Let us evaluate the possibilities of utilizing the mechanism under consideration to efficiently amplify submillimeter radiation by an intense REB with a current density $j=10^4\,\mathrm{A/cm^2}$, a mean particle energy equal to 1.5 MeV ($\overline{\gamma}=4$), and an energy spread $\delta=30\%$. A backward-wave relativistic generator[12] operating in the self-modulation stochastic regime[13] can, in principle, serve as the source of incoherent pump radiation. Let the mean wavelength of the pump field be $\lambda_i=3$ cm, let the width of the spectrum be 20%, and let the integrated intensity of the pump field be 2 GW. In the case of monochromatic pump radiation, such an intensity corresponds to an electric field strength of 0.5 MV/cm. The maximum pump field strength was selected here on the basis of the restrictions imposed by the development of rf breakdown.[3] We assume that the waveguide in which the scattering process takes place has a cross-sectional area of $\sim 3\ \mathrm{cm}^2$. Gyrotrons, whose emission power in the submillimeter range reaches 100 kW, can serve as the source of the input signal with a wavelength $\lambda_s=0.9$ mm.[14] The dependence of the energy-conversion efficiency on the longitudinal coordinate for the case under consideration is shown in Fig. 4 (curve *I*). In the linear stage the growth rate equals 0.02 cm^{-1}. The maximum efficiency of $\sim 5\%$ is achieved when the length of the scattering region is ~ 2 m. The gain equals 40 dB, and the output power of the radiation is 0.8 GW.

As was noted above a similar mechanism for improving the efficiency can be realized using stochastic undulators whose magnets are separated by a distance that varies according to a random law. An undulator can be considered stochastic when it satisfies the condition

$$\frac{\Delta h_i}{\overline{h_i}} \gg \frac{1}{N},$$

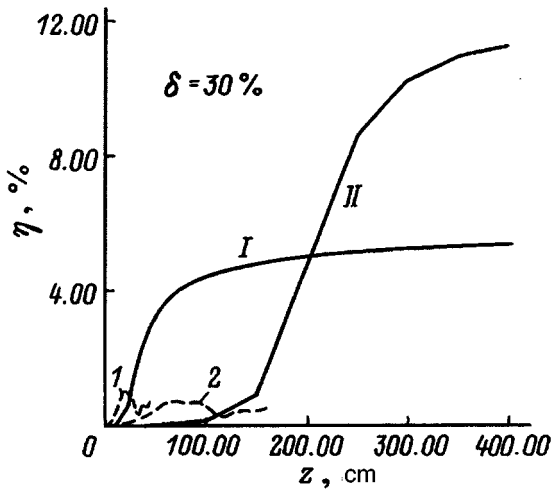

FIG. 4. Comparison of the energy-conversion efficiency of the electron stream in the case of incoherent (solid curves) and coherent (dashed) pump radiation. Pump radiation: *1*, *I* — multiwave; *2*, *II* — undulator.

where Δh_i is the width of the spectrum of wave numbers, $\overline{h_i}=2\pi/\overline{d}$ is the mean wave number, and $N=L/\overline{d}$ is the number of characteristic undulator periods $\overline{d}$ that fit into the total length L.

The amplification of a monochromatic signal wave in an FEL with a stochastic undulator is described by diffusion equations which coincide with Eqs. (8) and (9). A broader (in comparison with both microwave generators and optical lasers) wave-number spectrum of the pump field can be obtained in the case of a stochastic undulator. At the same time, because of the lack of restrictions associated with a gap, the oscillator frequency acquired by electrons in the pump field can be increased.

As an example, let us consider the case of the amplification of radiation with a wavelength $\lambda_s=0.3$ mm by an intense REB with a current density $\overline{j}=10^4$ A/cm^2 and a particle energy equal to 4.5 MeV ($\overline{\gamma}=10$) passing through a stochastic undulator with a characteristic period $\overline{d}=3$ cm and a magnetic field strength equal to 3 kOe. We choose a cross-sectional area of the interaction region equal to 0.5 cm^2. The dependence of the efficiency on the longitudinal coordinate for the case of an electron energy spread equal to 30% and a width of the longitudinal wave number spectrum of the pump field equal to 30% is depicted by curve *II* in Fig. 4. The maximum efficiency of $\sim 10\%$ is achieved when the length of the interaction region is 3 m. The output power of the short-wavelength radiation is 1.2 GW, which corresponds to a gain equal to 40 dB when the input power is 100 kW.

In conclusion, let us make a comparison with the results of the simulation of the amplification of monochromatic radiation by an REB having a broad energy spread and moving in a monochromatic pump field: $\mathbf{A}_i=\mathrm{Re}[\mathbf{x}_0 A_i \exp(i(\omega_i t+k_i z))]$. Separating the beam into individual fractions that differ with respect to the initial energy of the particles γ_ν, we describe the amplification process using the following system of equations (compare Refs. 4,5 and 15)

$$\frac{da}{d\xi}=\int_{-\Psi_{\max}}^{\Psi_{\max}} J_\nu F(\Psi_\nu)\,d\Psi_\nu, \quad J_\nu=\frac{1}{\pi}\int_0^{2\pi} e^{-i\Theta}\,d\Theta_0,$$

$$\frac{d\Theta_\nu}{d\xi}=w_\nu+\Psi_\nu, \quad \frac{dw_\nu}{d\xi}=\mathrm{Re}(ae^{i\Theta_\nu}),$$

$$\Theta_\nu|_{\xi=0}=\Theta_0\in[0,2\pi], \quad w_\nu|_{\xi=0}=0, \quad a|_{\xi=0}=a_0. \quad (13)$$

Here we have used the dimensionless variables

$$\xi=h_s z C^{-1}, \quad a=\mu\alpha_s\alpha_i C^{-2}, \quad w_\nu=\frac{\gamma-\gamma_\nu}{\overline{\gamma}}\frac{\mu}{C};$$

$$\Psi_\nu=\frac{(\gamma_\nu-\overline{\gamma})}{\overline{\gamma}}\frac{\mu}{C}$$

is the detuning from synchronism for the fraction with the energy γ_ν;

$$C=\left(\frac{\mu}{4}\frac{\omega_p^2}{\omega_s^2\overline{\gamma}}|\alpha_i|^2\right)^{1/3}$$

is the gain; $\mu=\overline{\gamma^2}+\alpha_i^2$ is the inertial bunching parameter; and $\alpha_{s,i}=eA_{s,i}/mc^2/\sqrt{2}\,\overline{\gamma}$. The electronic efficiency is defined by the relation

$$\eta=\frac{C}{\mu}\int_{-\Psi_{\max}}^{\Psi_{\max}}\int_0^{2\pi} w_\nu F(\Psi_\nu)\,d\Theta_0\,d\Psi_\nu,$$

where $F(\Psi_\nu)=\Psi_{\max}\cos^2(2\pi/\Psi_{\max})$ is the initial distribution function of the detuning of the electrons from synchronism [compare Eq. (10)], and $\Psi_{\max}=\mu\delta/2C$.

For an electron beam with the same characteristics as in the first of the examples considered above under the assumption that the pump field is monochromatic and has a strength equal to 0.5MV/cm, the gain $C=7.6\times10^{-3}$, the bunching parameter $\mu=7.6\times10^{-2}$, and the maximum detuning from synchronism $\Psi_{\max}=3$. The electronic efficiency found by numerically integrating Eqs. (12) is plotted in Fig. 4 (dashed curve *1*). For a second example, the energy-conversion efficiency in the case of a regular undulator with a period equal to 3 cm and a magnetic field strength equal to 3 kOe ($C=7.2\times10^{-4}$, $\mu=2\times10^{-2}$, $\Psi_{\max}=3$) is depicted in the same figure by dashed curve *2*. A comparison of these curves with the results of the preceding analysis leads to the conclusion that for an electron beam with large parameter spreads (up to 30%) the use of incoherent pump radiation with a relatively small decrease in the growth rates (by a factor of about 2 in the examples considered) makes it possible to increase the energy-conversion efficiency by at least an order of magnitude.

We thank the EC-DG III/ESPRIT Commission for its support (Project No. ACTCS 9282).

[1)]At the present time, the most fully developed method for increasing the efficiency of an FEL is the use of the synchronous adiabatic deceleration regime (the inverted accelerator regime), in which electrons are captured by a synchronous combination wave, whose phase velocity then slowly (in comparison to the period of the phase oscillations of the particles) decreases as a result of the variation of the period of the undulator field.[5–7] However, sufficiently complete capture of the particles can be achieved only for beams having small (on the scale of the amplitude of the combi-

nation wave) parameter spreads. Therefore, such a method is unattractive for intense REBs.

[2)]We note that in the kinetic interaction regime under discussion the effect of the rf space-charge fields is neglected, since for beams with large parameter spreads, for which the condition $\Delta v_z/v_z \gg \omega_{p\parallel}/\omega_s$ holds (Δv_z is the translational velocity spread of the electrons, and $\omega_{p\parallel}$ is the longitudinal plasma frequency), such fields have no influence on the growth rates (see, for example, Ref. 8).

[3)]We note that there is a possibility of raising the energy density of the pump field (above the value at the outlet from the radiation source) by employing waveguide sections with a nearly critical cross section as the scattering region for the pump field. Bragg resonators in the form of waveguides whose lateral walls are corrugated with a period $\sim\lambda_i/2$ can be used for this purpose.

[1] A. G. Litvak and V. Yu. Trakhtengerts, Zh. Éksp. Teor. Fiz. **60**, 1702 (1971) [Sov. Phys. JETP **33**, 921 (1971)].

[2] A. G. Litvak and V. Yu. Trakhtengerts, Zh. Éksp. Teor. Fiz. **62**, 228 (1972) [Sov. Phys. JETP **35**, 123 (1972)].

[3] N. S. Ginzburg, Pis'ma Zh. Tekh. Fiz. **10**, 584 (1984) [Sov. Tech. Phys. Lett. **10**, 246 (1984)].

[4] V. L. Bratman, N. S. Ginzburg, and M. I. Petelin, Zh. Éksp. Teor. Fiz. **76**, 930 (1979) [Sov. Phys. JETP **49**, 469 (1979)].

[5] P. Sprangle, C. M. Tang, and W. M. Manheumer, Phys. Rev. A **21**, 302 (1980).

[6] A. T. Lin and J. M. Dawson, Phys. Rev. Lett. **42**, 1670 (1979).

[7] N. M. Kroll, Ph. L. Norton, and M. N. Rosenbluth, in *Physical Quantum Electronics*, Vol. 7, S. F. Jacobs and M. O. Scully (Eds.), Addison-Wesley, Reading, Mass. (1980), p. 113.

[8] E. Jerby and A. Gover, IEEE. J. Quantum Electron. **QE-21**, 1041 (1985).

[9] A. A. Vedenov, E. P. Velikhov, and R. Z. Sagdeev, *Nuclear Fusion, Supplement, Part 2* [in Russian] (1962).

[10] W. E. Drumond and P. Pines, Nucl. Fusion Suppl. **3**, 1049 (1962).

[11] V. N. Tsytovich, *Nonlinear Effects in Plasmas*, Plenum Press, New York (1970).

[12] N. F. Kovalev, M. I. Petelin, M. D. Raĭzer *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **18**, 232 (1973) [JETP Lett. **18**, 138 (1973)].

[13] N. S. Ginzburg, S. P. Kuznetsov, and T. N. Fedoseeva, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **21**, 1037 (1978).

[14] A. G. Luchinin and G. S. Nusinovich, *Gyrotrons* [in Russian], Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki (1989), p. 55.

[15] W. B. Colson and C. Pellegrini, *Laser Handbook*, Vol. 6, W. B. Colson, C. Pellegrini, and A. Reniory (Eds.), North-Holland, Amsterdam (1990), p. 9.

Translated by P. Shelnitz

Sasers (sound amplification by stimulated emission of radiation) in the nonlinear operating regime with various emitters

S. T. Zavtrak and I. V. Volkov

Scientific-Research Institute of Nuclear Problems at Belarus State University, 220050 Minsk, Belarus
(Submitted February 21, 1996)
Zh. Tekh. Fiz. **67**, 82–85 (July 1997)

The operation of a saser (sound amplification by stimulated emission of radiation) in the nonlinear regime is considered. A liquid with gas bubbles is chosen as the active medium. Pumping is effected by a variable electric field. The processes leading to amplification of the active mode in a flat cavity are investigated analytically for the cases of homogeneous and inhomogeneous size distributions of the bubbles. © *1997 American Institute of Physics.* [S1063-7842(97)01507-9]

INTRODUCTION

The theoretical scheme of the acoustic analog of a free-electron laser, i.e., a saser (sound amplification by stimulated emission of radiation) has recently been proposed in a series of papers.[1–4] A liquid insulator with dispersed particles homogeneously distributed in it was chosen as the active medium. These particles can be gas bubbles, produced, for example, by electrolysis. A pump wave is created in the active medium, which is enclosed in a cavity, by a variable electric field[1–3] or by mechanical vibrations of the cavity walls.[4] The spatial distribution of the particles is initially homogeneous, and their total emission is equal to zero, but they subsequently begin to group together under the action of the acoustic radiation forces. This leads to autosynchronization of the oscillating particles and amplification of the active mode.

The autophasing of the gas bubbles in the initial stage of operation of a saser was investigated, and the starting conditions for the onset of generation were investigated in Refs. 1–4. It was shown that losses of two types must be overcome for generation to begin.[2] The first is associated with the dissipation of energy within the active medium, and the second is due to the radiation losses on the end plates of the cavity. It was assumed in Refs. 1–3 that all the bubbles have the same radius and are far from the resonance region. The more realistic case of an inhomogeneous size distribution of the bubbles was considered in Ref. 4. The starting conditions found for that case turned out to be similar to the conditions previously obtained.

It was assumed in all the preceding studies that the changes in the spatial concentration of the bubbles and the active mode are small compared with the initial values of the concentration and the pump wave, i.e., the initial stage of operation of the saser was considered. In the present work we shall dispense with that assumption and consider the operation of a saser at arbitrary moments in time for the cases of homogeneous and inhomogeneous size distributions of the bubbles. For simplicity, we shall consider a saser scheme with a flat cavity and electric-field pumping.

BASIC EQUATIONS

A schematic representation of a saser is presented in Fig. 1. The active medium is confined between two planes. Emission occurs along the z axis, and L is the cavity length in that direction. The active medium is immersed in a variable electric field, which plays the role of the pump wave.[1] The volumes of the bubbles vary because of electrostriction. The resultant pressure acting on the bubbles can be represented in the form $P(\mathbf{r},t)=P_E\exp(i\omega t)+P'(\mathbf{r},t)$ (the static pressure has been omitted), where $P_E\exp(i\omega t)$ is the pressure associated with the action of the electric field and $P'(\mathbf{r},t)$ is the pressure of the active mode.

The dynamics of the gas-liquid medium are described by the following system of equations:[1–4]

$$\Delta P'-\frac{1}{c_l^2}\frac{\partial^2 P'}{\partial t^2}-(\alpha+i\beta)P'=(\alpha+i\beta)P_E\exp(i\omega t), \tag{1}$$

$$\xi\mathbf{U}=(\mathrm{Re}\,A)\nabla|P|^2-i(\mathrm{Im}\,A)(P^*\nabla P-P\nabla P^*), \tag{2}$$

$$\frac{\partial n}{\partial t}+\mathrm{div}(n\mathbf{U})=0. \tag{3}$$

Equation (1) is the wave equation for a liquid with gas bubbles. Here c_l is the velocity of sound in the pure liquid; α and β are described by the following equations

$$\alpha=\alpha(\mathbf{r},R_0,t)=-4\pi\int_0^\infty(\mathrm{Re}\,A)n(\mathbf{r},R_0,t)dR_0, \tag{4}$$

$$\beta=\beta(\mathbf{r},R_0,t)=-4\pi\int_0^\infty(\mathrm{Im}\,A)n(\mathbf{r},R_0,t)dR_0, \tag{5}$$

where

$$A=\frac{R_0}{\dfrac{\omega_0^2}{\omega^2}-1+i\delta} \tag{6}$$

is the amplitude of the scattering of sound on a bubble with a radius R_0 and a resonant frequency ω_0, and δ is the absorption constant.

The quantity $n(\mathbf{r},R_0,t)$ is the distribution function of the bubbles. It is assumed that, initially, all the bubbles are distributed homogeneously in space. Equation (2) describes the

 1063-7842/97/070793-04$10.00

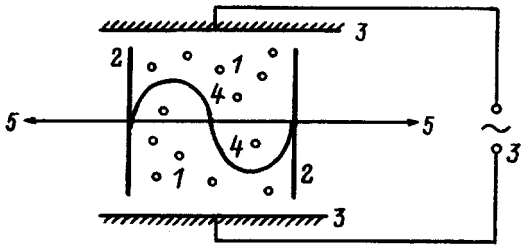

FIG. 1. Schematic representation of an electrically pumped saser: *1* — active medium; *2* — solid walls of the resonator; *3* — electromagnetic system, which creates a periodic electric field; *4* — active acoustic mode; *5* — sound radiation.

translational motion of a bubble in a sound field. Here **U** is the rate of translational motion, and the coefficient ξ has the form

$$\xi = 4\rho_l \mu_l \omega^2 R_0, \tag{7}$$

where ρ_l and μ_l are, respectively, the density and viscosity of the pure liquid.

Equation (3) is the balance equation for the number of particles in the phase volume element $d^3\mathbf{r}dR_0$ (we neglect any coagulation of the particles).

Let us investigate the one-dimensional solutions of the system of equations (1)–(3) (i.e., let us assume that all the parameters vary only along the z axis). We explicitly separate the rapidly oscillating multiplier $\exp(i\omega t)$ in P':

$$P'(\mathbf{r},t) = \overline{P(}\mathbf{r},t)\exp(i\omega t). \tag{8}$$

Substituting (8) into (1) and (2) and then (2) into (3) and applying the Bogolyubov-Mitropol'skii averaging procedure[5] to the high-frequency component $\exp(i\omega t)$, we obtain the following system of abridged equations:

$$\left[\frac{\partial^2}{\partial z^2} + \frac{\omega^2}{c_l^2} - \frac{2i\omega}{c_l^2}\frac{\partial}{\partial t}\right]\overline{P} = (\alpha + i\beta)(\overline{P} + P_E), \tag{9}$$

$$\frac{\partial n}{\partial t} = -\frac{A}{\xi}\frac{\partial}{\partial z}\left[n(\overline{P} + P_E)\frac{\partial \overline{P^*}}{\partial z}\right] + \text{c.c.} \tag{10}$$

ANALYSIS OF THE EQUATIONS IN THE FIRST STAGE OF GENERATION

Let us investigate the amplification of the active mode in a closed cavity with absolutely solid reflecting walls in the initial stage of operation of a saser. We introduce the deviation $n'(\mathbf{r},R_0,t)$ from the initial distribution function $n_0(R_0)$, i.e.,

$$n(\mathbf{r},R_0,t) = n_0(R_0) + n'(\mathbf{r},R_0,t), \tag{11}$$

and

$$\alpha = \alpha_0 + \alpha', \quad \beta = \beta_0 + \beta'.$$

Setting $|n'| \ll n_0$ and $|\overline{P}| \ll P_E$, we linearize the system of equations (9) and (10) with respect to n' and $\overline{P}$:

$$\left[\frac{\partial^2}{\partial z^2} + \frac{\omega^2}{c_l^2} - \frac{2i\omega}{c_l^2}\frac{\partial}{\partial t} - \alpha_0 - i\beta_0\right]\overline{P} = (\alpha' + i\beta')P_E, \tag{12}$$

$$\frac{\partial n'}{\partial t} = -\frac{n_0 P_E}{\xi}\left[A\frac{\partial^2 \overline{P^*}}{\partial z^2} + A^*\frac{\partial^2 \overline{P}}{\partial z^2}\right]. \tag{13}$$

In deriving these equations it was taken into account that $|\alpha_0 + i\beta_0| \ll \omega^2/c_l^2$.[1,2] We seek an expression for the active mode in the form of a standing wave

$$\overline{P}(z,t) = P_0(t)\cos(k_L z), \tag{14}$$

where $k_L = (\pi m)/L$, and $m = 1,2,\ldots$.

Then, substituting (14) into (13), we find

$$\frac{\partial n'}{\partial t} = \frac{n_0 P_E k_L^2}{\xi}(AP_0^* + A^*P_0)\cos(k_L z). \tag{15}$$

Differentiating Eq. (12) with respect to t and substituting expressions (14) and (15) into it, we obtain the equation which specifies $P_0(t)$

$$\left[\frac{\omega^2}{c_l^2} - k_L^2 - \alpha_0 - i\beta_0 - \frac{2i\omega}{c_l^2}\frac{d}{dt}\right]\frac{dP_0}{dt} = -4\pi P_E^2 k_L^2 (I_1 P_0^* + I_2 P_0), \tag{16}$$

where

$$I_1 = \int_0^\infty \frac{n_0 A^2}{\xi} dR_0, \quad I_2 = \int_0^\infty \frac{n_0 A A^*}{\xi} dR_0. \tag{17}$$

The calculation of α_0 and β_0 and integrals (17) requires knowledge of the explicit form of the distribution function $n_0(R_0)$ at the initial moment in time. Let us consider the cases of homogeneous and inhomogeneous size distributions of the bubbles.

Homogeneous distribution. In this case all the bubbles have the same radius R_0, i.e., $n_0(\overline{R_0}) = N_0\delta(\overline{R_0} - R_0)$. Then Eq. (16) takes the following form:

$$\left[\frac{\omega^2}{c_l^2} - k_L^2 - \alpha_0 - i\beta_0 - \frac{2i\omega}{c_l^2}\frac{d}{dt}\right]\frac{dP_0}{dt} = -\frac{P_E^2 k_L^2 \alpha_0^2}{4\pi N_0 \xi}(P_0^* + P_0), \tag{18}$$

where $\alpha_0 + i\beta_0 = -4\pi A N_0$.

During the derivation of Eq. (18) it was assumed that the bubbles are far from the resonance region and the dissipation of energy in the active medium is small, i.e., $|\beta_0| \ll |\alpha_0|$. We seek a solution of this equation in the form

$$P_0(t) = A^+\exp(i\omega' t) + A^-\exp(-i\omega'^* t). \tag{19}$$

Then, under the condition $|A^-| \ll |A^+|$ it is not difficult to obtain

$$\mathrm{Re}(\omega') \approx -\frac{c_l^2}{2\omega}\left(\frac{\omega^2}{c_l^2} - \alpha_0 - k_L^2\right) \equiv -\frac{\delta_L \omega}{2}, \tag{20}$$

$$\mathrm{Im}(\omega') \approx \frac{c_l^2}{2\omega}\left(\beta_0 + \frac{\alpha_0^2 k_L^2 P_E^2}{4\pi N_0 \xi \mathrm{Re}(\omega')}\right). \tag{21}$$

The real part of ω' [Eq. (20)] specifies the detuning of the angular pump frequency ω from the resonant frequency of the cavity $\omega_L = c_l\sqrt{k_L^2+\alpha_0}$, and the imaginary part of ω' [Eq. (21)] gives the gain of the active mode.[2,4] Since $\beta_0/\alpha_0 \sim \delta$, it is easy to prove from Eq. (21) that the conditions $\delta_L > 0$ and

$$P_E > P_{st} = \frac{1}{k_L}\sqrt{4\rho_l\mu_l\omega^2\delta|\ \mathrm{Re}(\omega')|} \tag{22}$$

must be satisfied for generation to begin.

Here P_{st} is the starting (threshold) pump pressure.[2,4]

Inhomogeneous distribution. As a rule, the experimentally observed size distribution of bubbles is characterized by the following relation

$$n_0(R_0) \approx aR_0^{-3}, \tag{23}$$

where the constant a specifies the local gas content in the liquid.

It was shown in the preceding study[4] that in the case of an inhomogeneous size distribution of the bubbles the main role is played by the bubbles with resonant frequencies close to the pump frequency ω. In fact, since the dissipation of energy in the medium is small, i.e., since $\delta \ll 1$, the integrands in Eqs. (4), (5), and (17) have a sharp maximum at $\omega = \omega_0$. When this is taken into account, it is not difficult to calculate the asymptotic forms of these integrals for $\delta \to 0$ (Ref. 4). As a result we have

$$\alpha_0 \approx 0, \quad \beta_0 \approx 2\pi^2/R_\omega^3 n_0(R_\omega), \tag{24}$$

$$I_1 \approx 0, \quad I_2 \approx \frac{\pi n_0(R_\omega)R_\omega^3}{2\xi(R_\omega)\delta}, \tag{25}$$

where R_ω is the radius of a bubble with the resonant frequency ω.

Substituting expressions (24) and (25) into Eq. (16) and seeking the solution of (16), as in the case of a homogeneous distribution, in form (19), we obtain expressions that are analogous to (20) and (21):

$$\mathrm{Re}(\omega') \approx -\frac{c_l^2}{2\omega}\left(\frac{\omega^2}{c_l^2} - k_L^2\right) \equiv -\frac{\delta_L\omega}{2}, \tag{26}$$

$$\mathrm{Im}(\omega') \approx \frac{c_l^2}{2\omega}\left(\beta_0 + \frac{2\pi^2 P_E^2 k_L^2 n_0(R_\omega)R_\omega^3}{\xi(R_\omega)\delta\ \mathrm{Re}(\omega')}\right). \tag{27}$$

It is easy to see that the starting conditions following from (26) and (27) coincide with those obtained in the preceding case.

NONLINEAR SASER OPERATING REGIME

Let us now consider the operation of a saser at arbitrary moments in time. We seek the dependence of the pressure on the time and the coordinates, as before, in form (14). Multiplying Eq. (9) by $\cos k_L z$ and averaging over the spatial coordinate we obtain

$$\left[\frac{\omega^2}{c_l^2} - k_L^2 - \frac{2i\omega}{c_l^2}\frac{d}{dt}\right]P_0(t) = -8\pi P_0\int_0^\infty An_2 dR_0 - 8\pi P_E\int_0^\infty An_1 dR_0, \tag{28}$$

where

$$n_1 = n_1(t) = \langle n(z,t)\cos k_L z\rangle,$$

$$n_2 = n_1(t) = \langle n(z,t)\cos^2 k_L z\rangle,$$

$$\ldots,$$

$$n_i = n_i(t) = \langle n(z,t)\cos^i k_L z\rangle.$$

Substituting Eq. (14) into (10) and multiplying Eq. (10) successively by $\cos k_L z$, $\cos^2 k_L z \ldots$, etc., after averaging over z we obtain an infinite hierarchy of equations [like the Bogolyubov–Born–Green–Yvon (BBGKY) hierarchy in statistical physics[6]]

$$\frac{\partial n_1}{\partial t} = \frac{k_L^2}{\xi}[(A^*P_0 + AP_0^*)P_E(n_0 - n_2) + (A + A^*)P_0P_0^*(n_1 - n_3)], \tag{29}$$

$$\frac{\partial n_2}{\partial t} = \frac{2k_L^2}{\xi}[(A^*P_0 + AP_0^*)P_E(n_1 - n_3) + (A + A^*)P_0P_0^*(n_2 - n_4)]. \tag{30}$$

Several assumptions must be made in order to analyze this infinite system of equations (the hierarchy is usually terminated at some level). In an approximation we replace the quantity $\cos^2 k_L z$ by 1/2 in the expressions for n_2 and n_3. Then $n_2 \approx n_0/2$, and $n_3 \approx n_1/2$. As a result, instead of Eq. (29) in an approximation we have

$$\frac{\partial n_1}{\partial t} - \frac{k_L^2(A + A^*)}{2\xi}P_0P_0^*n_1 \approx \frac{k_L^2P_En_0}{2\xi}(A^*P_0 + AP_0^*). \tag{31}$$

Separating the oscillating multiplier in $P_0(t)$, we have

$$P_0(t) = \exp(i\omega' t)\Phi_0(t), \tag{32}$$

where ω' is a frequency shift, which will be defined below.

Substituting Eq. (32) into (31) and averaging (31) with respect to the phase $\phi = \omega' t$, we find

$$n_1 = a\ \exp(i\omega' t) + a^*\exp(-i\omega' t), \tag{33}$$

where

$$a = \frac{k_L^2 n_0\Phi_0 P_E A^*}{2i\xi\omega'\left(1 + \dfrac{ik_L^2(A + A^*)}{2\xi\omega'}|\Phi_0|^2\right)}. \tag{34}$$

The function n_1 characterizes the inhomogeneity of the spatial distribution density of the bubbles. Expression (33) shows that this density oscillates with an angular frequency equal to the pump frequency shift ω'.

If we substitute (32)–(34) into Eq. (28) and average it with respect to the slow phase $\phi = \omega' t$, we obtain an equation for the frequency shift

$$\frac{\omega^2}{c_l^2}+\frac{2\omega\omega'}{c_l^2}-k_L^2-\alpha_0=\frac{2\pi P_E^2 k_L^4|\Phi_0|^2}{\omega'^2}F_1 \tag{35}$$

and an expression for the growth of the active mode

$$\frac{d\Phi_0}{dt}=-\frac{\beta_0 c_l^2}{2\omega}\left(1+\frac{4\pi P_E^2 k_L^2}{\omega'\beta_0}F_2\right)\Phi_0, \tag{36}$$

where

$$F_1=\int_0^\infty \frac{n_0|A|^2(A+A^*)}{\xi^2\left(1+\left(\frac{k_L^2(A+A^*)}{2\xi\omega'}|\Phi_0|^2\right)^2\right)}dR_0, \tag{37}$$

$$F_2=\int_0^\infty \frac{n_0|A|^2}{\xi\left(1+\left(\frac{k_L^2(A+A^*)}{2\xi\omega'}|\Phi_0|^2\right)^2\right)}dR_0. \tag{38}$$

In the initial stage of generation, in which Φ_0 is small, Eqs. (35) and (36) give ($F_2\approx I_2$)

$$\omega'=-\delta_L\omega/2, \tag{39}$$

$$\frac{d\Phi_0}{dt}=-\frac{\beta_0 c_l^2}{2\omega}\left(1-\frac{P_E^2}{P_{st}^2}\right)\Phi_0. \tag{40}$$

The expressions (39) and (40) obtained are valid in the initial stage of operation of a saser both for the case of a homogeneous size distribution of the bubbles and for the case of an inhomogeneous distribution. Equation (39) clearly demonstrates that at the beginning of generation the resultant angular frequency ($\omega+\omega'$) of the saser output is determined by the characteristic wave number of the cavity. At latter times the nonlinear terms [the right-hand side of Eq. (35)] begin to play a more significant role, which leads to an additional shift ω'.

As was noted above, if $P_E>P_{st}$ at the initial moment in time, Eq. (36) describes the increase in the amplitude of the active mode up to values at which the right-hand side of (36) vanishes. This permits estimation of the maximum amplitude of the pressure of the active mode.

In the case of a homogeneous size distribution of the bubbles we have

$$|P_{0\max}|=P_E\left(\frac{P_E^2}{P_{st}^2}-1\right)^{1/4}\left(\frac{|\alpha_0|}{\beta_0}\right)^{1/2}. \tag{41}$$

Since $\frac{|\alpha_0|}{\beta_0}\sim\frac{1}{\delta}$, Eq. (41) gives

$$|P_{0\max}|=P_E\left(\frac{P_E^2}{P_{st}^2}-1\right)^{1/4}\delta^{-1/2}. \tag{42}$$

In the case of an inhomogeneous size distribution of the bubbles, it is not difficult to calculate the asymptotes of the integrals in Eqs. (37) and (38) for $\delta\to 0$:

$$F_1\approx 0,\quad F_2\approx\frac{\pi n_0(R_\omega)R_\omega^3}{\delta\xi(R_\omega)\sqrt{\left|\frac{\Phi_0}{P_{st}}\right|^4+4}}. \tag{43}$$

When this is taken into account, we have

$$|P_{0\max}|=\sqrt{2}P_E\left(\left(\frac{P_E}{P_{st}}\right)^4-1\right)^{1/4}. \tag{44}$$

CONCLUSIONS

Thus, a saser achieves a saturation regime during its operation. This effect is similar to saturation in the theory of free-electron lasers.

[1] S. T. Zavtrak, Phys. Rev. E **51**, 2480 (1995).
[2] S. T. Zavtrak, Phys. Rev. E **51**, 3767 (1995).
[3] S. T. Zavtrak, Zh. Tekh. Fiz. **65**(6), 123 (1995) [Tech. Phys. **40**, 589 (1995)].
[4] S. T. Zavtrak and I. V. Volkov, Zh. Tekh. Fiz. **67** (4), 92 (1997) [Tech. Phys. **42**, 406 (1997)].
[5] N. N. Bogolyubov and Yu. A. Mitropol'skiĭ, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, Gordon and Breach, New York (1962).
[6] A. Isihara, *Statistical Physics*, Academic Press, New York (1971) [Russ. transl., Mir, Moscow (1973)].

Translated by P. Shelnitz

Electromagnetic waves in a round rod in the presence of an arbitrarily directed external magnetic field or anisotropy axis

Yu. F. Filippov

Institute of Radiophysics and Electronics, Ukrainian National Academy of Sciences, 310085 Kharkov, Ukraine
(Submitted January 10, 1996)
Zh. Tekh. Fiz. **67**, 86–91 (July 1997)

A rigorous theory of the propagation of electromagnetic waves in round anisotropic and semiconductor rods in the presence of an arbitrarily directed anisotropy axis or external magnetic field is developed. New types of independent waves are discovered. Exact dispersion equations are obtained for them, which define the dependence of their spectral characteristics on the parameters of the semiconductor or anisotropic crystal and on the magnitude and direction of the constant external magnetic field. The results of numerical investigations for rods made from a semiconductor or a uniaxial crystal are presented. © *1997 American Institute of Physics.* [S1063-7842(97)01607-3]

The spectral characteristics of electromagnetic waves propagating in anisotropic and semiconductor rods have been investigated only for cases in which the direction of the external magnetic field $\mathbf{H}_0$ or the anisotropy axis is parallel to their geometric axis. Establishment of the relation between the components of the dielectric tensor and the spectral characteristics of waves for arbitrary orientations is of great current interest.

THEORY

Let us consider a round homogeneous rod, which is bounded at $r=r_0$ by a vacuum and made from a material whose electrical parameters are described by the components of a dielectric tensor of the following form

$$\hat{\varepsilon}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{1}$$

Here

$$a_{11}=\varepsilon_{11}C^2+\varepsilon_{33}S_2^2+(\varepsilon_{13}+\varepsilon_{31})SC, \quad a_{22}=\varepsilon_{22},$$
$$a_{12}=\varepsilon_{12}C+\varepsilon_{32}S, \quad a_{21}=\varepsilon_{21}C+\varepsilon_{23}S,$$
$$a_{13}=(\varepsilon_{33}-\varepsilon_{11})SC+\varepsilon_{13}C^2-\varepsilon_{31}S^2,$$
$$a_{31}=(\varepsilon_{33}-\varepsilon_{11})SC+\varepsilon_{31}C^2-\varepsilon_{13}S^2,$$
$$a_{23}=\varepsilon_{23}C-\varepsilon_{21}S, \quad a_{32}=\varepsilon_{32}C-\varepsilon_{12}S,$$
$$a_{33}=\varepsilon_{11}S^2+\varepsilon_{33}C^2-(\varepsilon_{13}+\varepsilon_{31})SC,$$

$S=\sin\Theta$, $C=\cos\Theta$, Θ is the angle of inclination of the external magnetic field or the anisotropy axis ε_{33} to the geometry axis of the rod in the 1–3 plane, and the ε_{ij} are the components of the dielectric tensor when $\Theta=0$.

In particular, in a uniaxial crystal (quartz, ruby, or leucosapphire) the latter equal

$$\varepsilon_{11}=\varepsilon_{22}\neq\varepsilon_{33}, \quad \varepsilon_{12}=\varepsilon_{21}=\varepsilon_{13}=\varepsilon_{31}=\varepsilon_{23}=\varepsilon_{32}=0,$$

and in a semiconductor immersed in an external magnetic field we have

$$\varepsilon_{11}=\varepsilon_{22}=\varepsilon_L\left(1-\sum_\alpha \omega_{p\alpha}^2/\Omega_\alpha\right),$$
$$\varepsilon_{13}=\varepsilon_{31}=\varepsilon_{23}=\varepsilon_{32}=0,$$
$$\varepsilon_{12}=-\varepsilon_{21}=i\varepsilon_L\sum_\alpha \omega_{p\alpha}^2\omega_{H\alpha}(\omega\Omega_\alpha)^{-1},$$
$$\varepsilon_{33}=\varepsilon_L\left(1-\sum_\alpha \omega_{p\alpha}^2(\omega+i\nu_\alpha)^{-1}/\omega\right), \tag{2}$$

where $\Omega_\alpha=(\omega+i\nu_\alpha)^2-\omega_{H\alpha}^2$; ε_L is the lattice constant; $\omega_{P\alpha}$, $\omega_{H\alpha}$, and ν_α are the plasma frequency, the cyclotron frequency, and the collision frequency of the particles of type α; and the summation is carried out over all the particles.

It is more convenient to seek the solution of the system of Maxwell's equations ($k=\omega/c$)

$$\text{curl}\,\mathbf{E}=ik\mathbf{H}, \quad \text{div}\mathbf{H}=0,$$
$$\text{curl}\,\mathbf{H}=-ik\hat{\varepsilon}\mathbf{E}, \quad \text{div}\hat{\varepsilon}\mathbf{E}=0 \tag{3}$$

for a round rod in the cylindrical coordinate system (t,φ,z), in which the tensor $\hat{\varepsilon}$ takes the form

$$\hat{\varepsilon}=\begin{pmatrix} \sigma_+ +\sigma_- C_2+\gamma_+ S_2 & -\sigma_- S_2+\gamma_+ C_2-\gamma_- & a_{13}C_1+a_{23}S_1 \\ -\sigma_- S_2+\gamma_+ C_2+\gamma_- & \sigma_+ -\sigma_- C_2+\gamma_+ S_2 & a_{23}C_1-a_{13}S_1 \\ a_{31}C_1+a_{32}S_1 & a_{32}C_1-a_{31}S_1 & a_{33} \end{pmatrix}, \tag{4}$$

 1063-7842/97/070797-06$10.00

where

$$2\sigma_{\pm}=a_{11}\pm a_{22},\quad 2\gamma_{\pm}=a_{12}\pm a_{21},$$

$$S_m=\sin(m\varphi),\quad C_m=\cos(m\varphi).$$

The dependence of the components of this tensor on the azimuthal angle φ complicates the investigation significantly. However, an exact solution can be obtained in this case, too. For monochromatic waves which have a dependence on the time and the axial coordinate of the form

$$\Omega(z,t)=\exp[i(k_z z-\omega t)], \tag{5}$$

an investigation of system (3) is carried out to find solutions of the coupled differential equations for the axial field components of the form

$$\hat{L}_H H_z=-k\hat{\Lambda}_+ E_z,$$

$$\hat{L}_E E_z=-k^3\Lambda_- H_z, \tag{6}$$

where

$$\hat{L}_H=L+V\Delta_\perp+k^2(P_+\hat{g}_++P_-\hat{g}_-),$$

$$\hat{L}_E=k^2T+(L+k_z^2V)\Delta_\perp-k^2k_z^2(P_+\hat{g}_++P_-\hat{g}_-)$$
$$+ik_zk^2[(S_++G_+)\hat{b}_++(S_-+G_-)\hat{b}_-],$$

$\Delta_\perp$ is the transverse Laplacian operator

$$\Lambda_\pm=S_+\hat{b}_+-S_-\hat{b}_-+ik_z[P_+\hat{g}_+-P_-\hat{g}_-\pm\gamma_-\Delta_\perp],$$

$$V=\sigma_+k^2-k_z^2,\quad 2P_\pm=\sigma_-\mp\gamma_+,$$

$$T=a_{33}L-k^2[(a_{31}\delta_{213}+a_{32}\delta_{123})k^2$$
$$-(a_{13}a_{31}+a_{23}a_{32})k_z^2],$$

$$2S_\pm=(\delta_{213}\mp i\delta_{123})k^2-(a_{13}\mp ia_{23})k_z^2,$$

$$2G_\pm=(\delta_{231}\mp i\delta_{132})k^2-(a_{31}\mp ia_{32})k_z^2,$$

$$\delta_{ike}=a_{ii}a_{ke}-a_{ie}a_{ki},\quad L=L_{11}L_{22}+(\gamma_-^2-\gamma_+^2)k^4.$$

The influence of the azimuthal nonuniformity appearing in the components of $\hat{\varepsilon}$ upon passage to the cylindrical coordinate system on the characteristics of the propagating waves is described by the operators

$$\hat{b}_\pm=e^{\pm i\varphi}\left(\frac{\partial}{\partial r}\pm\frac{i}{r}\frac{\partial}{\partial\varphi}\right),$$

$$\hat{g}_\pm=e^{\pm 2i\varphi}\left[\Delta_\perp-\frac{2}{r}\left(1\mp i\frac{\partial}{\partial\varphi}\right)\left(\frac{\partial}{\partial r}\pm\frac{i}{r}\frac{\partial}{\partial\varphi}\right)\right].$$

We seek the solution of (3) in the form

$$H_z=\sum_n A_n\Psi_n(r_i\varphi)\Omega(z_it),$$

$$E_z=\sum_n B_n\Psi_n(r_i\varphi)\Omega(z_it). \tag{7}$$

Here A_n and B_n are constants $\Psi_n(z_i\varphi)=J_n(\varkappa r)e^{in\varphi}$, $\varkappa$ and k_z are the radial and axial components of the wave vector, and n is the azimuthal mode number. Using the recurrence relations between the Bessel functions $J_n(x)$ and their products (here and below, a prime denotes the derivative with respect to the argument)

$$xJ_n'(x)=nJ_n(x)-xJ_{n-1}(x)=-nJ_n(x)+xJ_{n+1}(x)$$

it can be shown that

$$\hat{b}_\pm\Psi_n(r_i\varphi)=\mp\varkappa\Psi_{n\pm1}(r_i\varphi),$$

$$\hat{g}_\pm\Psi_n(r_i\varphi)=\varkappa^2\Psi_{n\pm2}(r_i\varphi). \tag{8}$$

The influence of the operators $\hat{b}_\pm$ and $\hat{g}_\pm$ on the wave functions $\Psi_n(r_i\varphi)$ leads to alteration of only the azimuthal mode numbers.

Substituting (4) into (3), taking into account (5), and equating the coefficients in front of $\Psi_n(r_i\varphi)$, we obtain a system of algebraic equations relative to A_n and B_n:

$$(L-\varkappa^2V)A_n+\varkappa^2k^2(P_+A_{n-2}+P_-A_{n+2})$$
$$=k\varkappa[S_+B_{n+1}+S_-B_{n-1}-ik_z\varkappa$$
$$\times(B_{n-2}P_+-B_{n+2}P_--\gamma_-B_n)],$$
$$[k^2T-\varkappa^2(L+k_z^2V)]B_n-\varkappa^2k^2k_z^2(P_+B_{n-2}+P_-B_{n+2})$$
$$-ik_zk^2\varkappa[(S_++G_+)B_{n-1}-(S_-+G_-)B_{n+1}]$$
$$=\varkappa k^2[G_+A_{n-1}+G_-A_{n+1}-ik_z\varkappa$$
$$\times(P_+A_{n-2}-P_-A_{n+2}+\gamma_-A_n)]. \tag{9}$$

For the class of solutions

$$A_{n-1}=\xi A_{n+1},\quad B_{n-1}=\xi B_{n+1} \tag{10}$$

this infinite system separates, when $\xi=\pm1$, into independent blocks of the form

$$\eta_HA_n=\eta_1^+B_{n+1}+\eta_2^+B_n,$$

$$\eta_EB_n+ikk_z(\eta_1^-+\lambda_1^-)B_{n+1}=k^2(\lambda_1^+A_{n+1}-\eta_2^-A_n),$$

$$\eta_HA_{n+1}=\xi\eta_1^+B_n+\eta_2^+B_{n+1},$$

$$\eta_EB_{n+1}+i\xi kk_z(\eta_1^-+\lambda_1^-)B_n=k^2(\xi\lambda_1^+A_n-\eta_2^-A_{n+1}), \tag{11}$$

where

$$\eta_H=L-\varkappa^2\tau_-,\quad \eta_E=k^2T-\varkappa^2(L+k_z^2\tau_+),$$

$$\tau_\pm=V\pm\xi\tau_-k^2,\quad \lambda_\pm=\gamma_-\pm\xi\gamma_+,$$

$$\eta_1^+=k\varkappa(S_-\pm\xi S_+),\quad \eta_2^\pm=ikk_z\varkappa^2\lambda_\pm,$$

$$\lambda_1^\pm=k\varkappa(G_-\pm\xi G_+).$$

A nontrivial solution of (11) exists when

$$[\eta_H\eta_E+k^2(\eta_2^+\eta_2^--\xi\eta_1^+\lambda_1^+)]^2$$
$$=\xi k^2[k\lambda_1^+\eta_2^+-\eta_2^-\eta_2^++ik_z(\eta_1^-+\lambda_1^-)\eta_H]^2. \tag{12}$$

This equation defines the radial components of the wave vector. From (11) we note that coupling appears in the rod between the partial modes characterized by consecutive azi-

muthal mode numbers, i.e., n and $n+1$. Below we shall say that a nonaxial orientation of the anisotropy axis or the external magnetic field leads to an interaction between these modes. Independent waves, which we call symmetric when $\xi=1$ and asymmetric when $\xi=-1$, then appear in the structure. The radial components of their wave vectors are determined by the solutions of Eq. (12). The tangential components of the wave fields must be continuous on the surface of the rod. To satisfy this condition, two solutions must be considered. The solutions of system (5) which are finite on the axis of the rod and satisfy the Sommerfeld condition at infinity can be represented in the form of the expansions

$$H_z=\Omega(z,t)\sum_n e^{in\varphi}\begin{cases}\sum_j A_{nj}J_n(\varkappa_j r) & r\leqslant r_0,\\ G_n\mathcal{H}_n^{(1)}(\varkappa_0 r) & r\geqslant r_0,\end{cases}$$

$$F_z=\Omega(z,t)\sum_n e^{in\varphi}\begin{cases}\sum_j B_{nj}J_n(\varkappa_j r) & r\leqslant r_0,\\ R_n\mathcal{H}_n^{(1)}(\varkappa_0 r) & r\geqslant r_0,\end{cases} \tag{13}$$

where the $\varkappa_j$ are specified by the solutions of (12), the index j takes the values 1 and 2, and $\varkappa_0^2=k^2-k_z^2$.

Using (8) and (10), we can easily show that the azimuthal components E_φ and H_φ are defined by the following expressions:

$$E_\varphi=\Omega(z,t)\sum_n e^{in\varphi}\times\begin{cases}\sum_j [a_{Hj}^n A_{nj}-b_{Ej}^n B_{n+1j}-ih_{Ej}^n B_{nj}]/L,\\ [k\varkappa_0 G_n\mathcal{H}_n^{(i)'}(\varkappa_0 r)-ink_z R_n\mathcal{H}_n^{(i)}(\varkappa_0 r)]/\varkappa_0^2,\end{cases}$$

$$H_\varphi=\Omega(z,t)\sum_n e^{in\varphi}\times\begin{cases}\sum [a_{Ej}^n B_{nj}+ig_{Ej}^n B_{n+1j}+ih_{Hj}^n A_{hj}]/L,\\ [k\varkappa_0 R_n\mathcal{H}_n^{(i)'}(\varkappa_0 r)+ink_z G_n\mathcal{H}_n^{(i)}(\varkappa_0 r)]/\varkappa_0^2,\end{cases} \tag{14}$$

where

$$a_{Hj}^n=k[\tau_-\varkappa_j J_{nj}'-ink^2\lambda_- J_{nj}/r],$$
$$b_{Ej}^n=k^2[S_- J_{n+1j}-\xi S_+ J_{n-1j}],$$
$$h_{Ej}^n=k_z[n\tau_+ J_{nj}/r-i\lambda_+ k^2\varkappa_j J_{nj}'],$$
$$a_{Ej}^n=k\{[(a_{11}a_{22}-\gamma_+)k^2-(\tau_+-\xi\tau_-)k_z^2]\varkappa J_{nj}' - ink_z^2\lambda_+ J_{nj}/r\},$$
$$g_{Ej}^n=kk_z[S_- J_{n+1j}+\xi S_+ J_{n-1j}],$$
$$h_{Hj}^n=k_z[n\tau_- J_{nj}/r-ik^2\lambda_-\varkappa_j J_{nj}'].$$

The expressions obtained after substituting (13) and (14) into the boundary conditions indicated above and relations (11) form a complete system of algebraic equations with respect to A_{n+1j}, A_{nj}, B_{n+1j}, and B_{nj}. The condition for the existence of nontrivial solutions leads to the dispersion equation

$$\sum_{m=0}^{4}\varkappa_0^{2m}F_m=0, \tag{15}$$

where

$$F_0=\langle R_{12}\rangle-\alpha_2\langle T_{12},R_{12}+\beta_1 G_{12}\rangle+\beta_1\langle R_{12},G_{12}-\alpha_1 R_{21}\rangle+\xi\alpha_2\beta_2\langle R_{11},R_{22}\rangle+(\alpha_1^2+\xi\alpha_2^2)\langle T_{12}\rangle+(\beta_1^2+\xi\beta_2^2)\langle G_{12}\rangle-\langle R_{21},\alpha_1(\beta_1^2+\xi\beta_2^2)G_{12}\rangle-\langle R_{21},\alpha_1(\beta_1^2+\xi\beta_2^2)G_{12}-\beta_1(\alpha_1^2+\xi\alpha_2^2)T_{12}\rangle+(\alpha_1^2+\xi\alpha_2^2)(\beta_1^2+\xi\beta_2^2)\langle R_{21}\rangle,$$

$$\xi F_1=\alpha_2[\beta_1\langle R_{22},S_{11}\rangle+\langle S_{12},T_{12}-\beta_1 R_{21}\rangle+\beta_1\langle T_{12}P_{21}\rangle+(\beta_1^2+\xi\beta_2^2)(\langle R_{21},P_{12}\rangle-\langle R_{22},P_{11}\rangle)]+\beta_2[\langle R_{12},\alpha_1 S_{21}-P_{21}\rangle+\langle R_{11},P_{22}-\alpha_1 S_{22}\rangle+\alpha_1\langle T_{12},P_{21}\rangle-(\alpha_1^2+\xi\alpha_2^2)\langle T_{12},S_{21}\rangle],$$

$$\xi F_2=\langle S_{12}\rangle-\alpha_1[\beta_1\langle S_{11},S_{22}\rangle+(\beta_1^2+\xi\beta_2^2)(\langle P_{21},S_{21}\rangle-\langle P_{11},S_{22}\rangle)]+\beta_1[\langle P_{22},S_{11}\rangle-\langle P_{12},S_{12}\rangle]+\xi\alpha_2\beta_2[\langle T_{12},K_{12}\rangle+\langle S_{12},S_{21}\rangle]+(\alpha_1^2+\xi\alpha_2^2)(\beta_1^2+\xi\beta_2^2)\langle S_{21}\rangle+(\beta_1^2+\xi\beta_2^2)[\langle P_{12}\rangle+\langle P_{21}\rangle+\langle P_{11},P_{12}\rangle],$$

$$F_3=-\langle K_{12},\beta_2 S_{12}+\alpha_2(\beta_1^2+\xi\beta_2^2)S_{21}\rangle,$$

$$F_4=(\beta_1^2+\xi\beta_2^2)\langle K_{12}\rangle,$$

$$\langle L_{ij}\rangle=L_{ij}^n L_{ij}^{n+1},\quad \langle L_{ij}G_{km}\rangle=L_{ij}^n G_{km}^{n+1}+L_{ij}^{n+1}G_{km}^n,$$

$$R_{ij}^n=Z_{Hi}^n Z_{Ej}^n-V_{Hi}^n V_{Ej}^n,\quad T_{12}^n=Z_{H1}^n V_{H2}^n-Z_{H2}^n V_{H1}^n,$$

$$P_{ij}^n=Z_{Ei}^n b_{Ej}^n+V_{Ei}^n g_{Ej}^n,\quad K_{12}^n=b_{E1}^n g_{E2}^n-b_{E2}^n g_{E1}^n,$$

$$S_{ij}^n=Z_{Hi}^n g_{Ej}^n+V_{Hi}^n b_{Ej}^n,\quad G_{12}^n=Z_{E1}^n V_{E2}^n-Z_{E2}^n V_{E1}^n,$$

$$Z_{qi}^n=\varkappa_0^2 a_{qi}^n-k\delta_n L J_{ni},$$

$$\delta_n=\varkappa_0 r_0\mathcal{H}_n^{(1)'}(\varkappa_0 r_0)/\mathcal{H}_n^{(1)})(\varkappa_0 r_0),$$

$$V_{qi}^n=\varkappa_0^2 h_{qi}^n-nk_z LJ_{ni}/r_0,\quad \eta_0=\eta_E^2+\xi k^2 k_z^2(\eta_1^-+\lambda_1^-),$$

$$\alpha_1=\eta_2^+/\eta_H|_{\varkappa=\varkappa_2},\quad \alpha_2=-\eta_1^+/\eta_H|_{\varkappa=\varkappa_2},$$

$$\beta_1=k^2(\eta_E\eta_2^- -i\xi kk_z\lambda_1^+)(\eta_1^-+\alpha_1^-)/\eta_0|_{\varkappa=\varkappa_1},$$

$$\beta_2=[\eta_E\lambda_1^+-ikk_z\eta_2^-(\eta_1^-+\alpha_1^-)]k^2/\eta_0|_{\varkappa=\varkappa_1}.$$

Solutions (12) and (15) define the dependence of the spectral characteristics on the parameters of the problem of independent symmetric and asymmetric waves in a round homogeneous rod. The latter is made from a material whose electrical properties are specified by the components of tensor (1). One interesting feature is that these waves form as a result of the interaction of partial modes with consecutive azimuthal mode numbers.

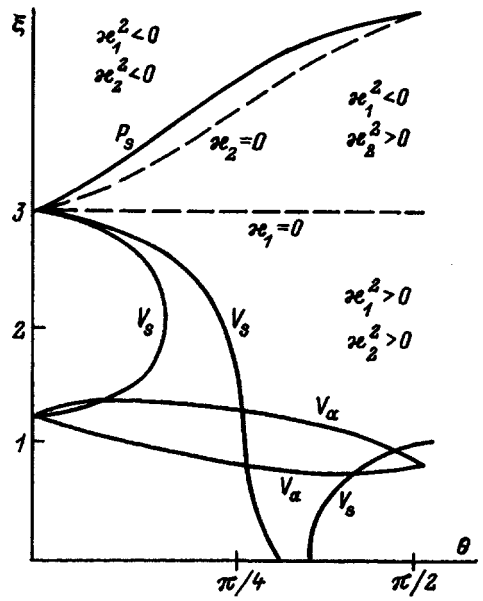

FIG. 1. Dependence of the retardation ξ on the angle of inclination Θ of the anisotropy axis for $n=5$ and $kr_0=2.5$.

The results obtained above can be generalized to the case in which the properties of the material from which the rod is made are also determined by components of the permeability tensor (a ferrite or a ferromagnetic semiconductor).

In the general case the dispersion equation can be investigated only by numerical methods. Below we present the results obtained for the propagation of waves in a rod made from a uniaxial crystal in the presence of an arbitrarily oriented anisotropy axis, as well as for axially uniform magnetoplasmon modes in the presence of an arbitrarily oriented external magnetic field.

ANISOTROPIC RODS

Figure 1 presents the results of a numerical investigation of the dependence of the retardation $\xi=k_z/k$ on the angle of inclination of the anisotropy axis Θ for $n=5$ and $kr_0=2.5$. The rod was made from a uniaxial crystal of leucosapphire ($\varepsilon_{11}=\varepsilon_{22}=9.4$, $\varepsilon_{33}=11.51$). It is seen from the figure that two asymmetric waves V_a propagate with different phase velocities in the structure when Θ is not equal to 0 or 90°. They appear as a result of the interaction of two partial modes of E polarization with consecutive azimuthal mode numbers. In fact, when $\xi=-1$, system (11) separates into independent blocks of the form

$$(L_{22}-\varkappa^2)A_n=0,$$
$$\delta B_n=-2ia_{13}\varkappa k_z B_{n+1},$$
$$\delta B_{n+1}=2ia_{13}\varkappa k_z B_n, \tag{16}$$

where $\delta=\varepsilon_{11}\varepsilon_{33}k^2-a_{33}k_z^2-a_{11}\varkappa^2$.

On the basis of (16), we note that partial modes, viz., one H ($E_z=0$) mode and two E ($H_z=0$) modes, appear in the rod. Their radial components are equal, respectively, to

$$\varkappa_1=\sqrt{L_{22}},\quad a_{11}\varkappa_2^{\pm}=[\varepsilon_{11}\varepsilon_{33}L_{22}]^{1/2}\pm a_{13}k_z. \tag{17}$$

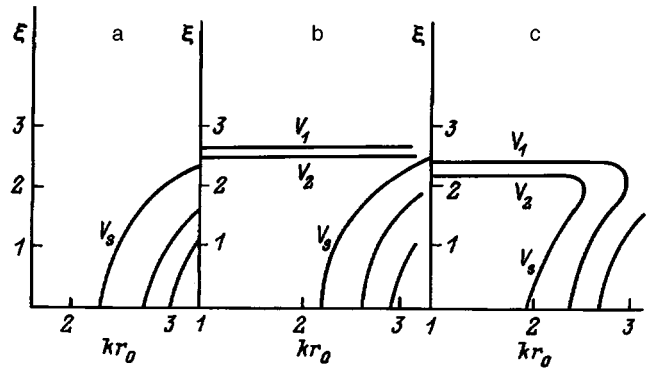

FIG. 2. Dependence of the retardation ξ of symmetric waves on kr_0 for $n=5$ and $\Theta=0$ (a), $0.3\pi/2$ (b), and $0.35\pi/2$ (c).

As was noted above, crossover of the H and E modes occurs on the surface of the rod. Either waveguide modes then appear in the structure, if $k_z>k$, or whispering gallery modes appear, if $k_z<k$. The influence of the curvature of the rod surface leads to partial emission of the latter into the vacuum. The latter decreases as the azimuthal mode index increases.

In the case of symmetric waves ($\xi=1$), system (11) separates into independent blocks of the form

$$(L_{11}-\varkappa^2)A_n=\varkappa k a_{13}B_{n+1}, \tag{18}$$
$$[L_{22}(\varepsilon_{11}\varepsilon_{33}k^2-a_{33}k_z^2)-\varepsilon_{11}L_{11}\varkappa^2]B_{n+1}=\varkappa k a_{13}A_n.$$

The nonaxial orientation of the anisotropy in this case leads to an interaction of the partial modes of E and H polarization with consecutive azimuthal mode numbers. Symmetric waves appear. The partial components of their wave vectors equal

$$\varkappa_1=\sqrt{L_{22}},\quad \varkappa_2=[\varepsilon_{11}\varepsilon_{33}k^2-a_{33}k_z^2]^{1/2}/\sqrt{\varepsilon_{11}}. \tag{19}$$

When $\varkappa_1^2$ and $\varkappa_2^2$ are positive, they are waveguide and whispering gallery modes. It is seen from Fig. 1 that they exhibit a zone of opaqueness with respect to the angle of inclination and nonuniqueness of the phase velocity at small values of kr_0. This is due to the appearance of low-frequency waveguide modes when the anisotropy axis deviates from the direction of the geometric axis. The dependence of the retardation ξ on the dimensionless parameter kr_0 at values of the angle of inclination Θ equal to 0, $0.3\pi/2$, and $0.35\pi/2$ is presented in Fig. 2a, and the dependence on the angle of inclination for kr_0 equal to 2.2, 2.3, and 2.5 is presented in Fig. 3. In an axially anisotropic rod ($\Theta=0$) the waveguide modes V_s exist only for $\omega>\omega_{cr}$ (Fig. 2a). Their critical frequency ω_{cr} is determined from the equation ($x_{cr}=\omega_{cr}r_0/c$)

$$\mathcal{H}_n^{(1)}(x_{cr})J_n'(\sqrt{\varepsilon_{11}}x_{cr})=\mathcal{H}_n^{(1)\prime}(x_{cr})J_n(\sqrt{\varepsilon_{11}}x_{cr}). \tag{20}$$

The low-frequency branches V_1 and V_2 appear for even a small deviation of the anisotropy axis from the geometric axis (Fig. 2b). Their spectral parameters at small kr_0 are determined by the solutions of the equations

$$L_{22}^3=(k\tan\Theta)^2[(\varepsilon_{11}k\varkappa_2\tan\Theta)^2+L_{22}(L_{22}+\varkappa_2^2)], \tag{21}$$

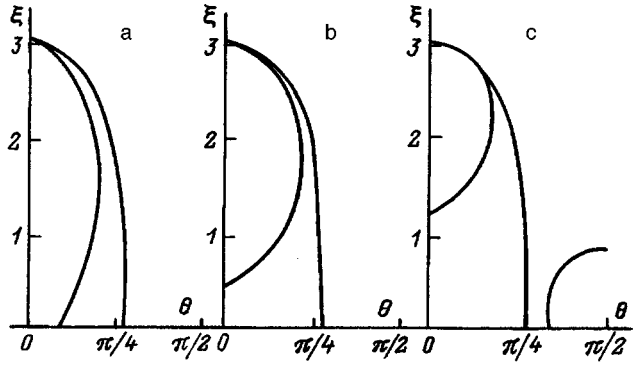

FIG. 3. Dependence of the retardation ξ of symmetric waves on the angle of inclination Θ for $n=5$ and $Kr_0=2.2$ (a), 2.3 (b), and 2.5 (c).

where $\varkappa_2$ was defined in Eq. (19).

These waves also exist below the critical frequency and depend weakly on the radius of the rod. An increase in the angle of inclination leads to an increase in the phase velocity for V_1 and V_2 and to a decrease in it for V_s. The region where the latter exist shifts to lower frequencies, and the phase velocity of the low-frequency modes depends significantly on the angle of inclination. All this leads to the appearance of the features indicated above in the behavior of the dispersion curves.

Symmetric and asymmetric whispering gallery modes were observed experimentally in Ref. 1 in a disk cavity made from leucosapphire with an anisotropy axis lying in the plane of a transverse section of the rod ($\Theta=90°$).

When $\varkappa_1^2$ and $\varkappa_2^2$ are negative in the structure and Θ is not equal to 0 or 90°, a symmetric surface wave P_s appears. Its amplitude decreases exponentially on both sides of the surface. An increase in the azimuthal mode index leads to a decrease in $\varkappa_2^2$ and the disappearance of this wave. This wave does not appear on a flat interface. Surface waves are observed on the latter in the optical range near the absorption bands.[2]

The insulating rods employed in light guides and lasers are fabricated predominantly from uniaxial crystals. The strains appearing in them during the formation of the rod texture and under the influence of mechanical stresses have a significant influence on the operation of these devices. They also lead to alteration of the values of the components of the dielectric tensor and to the appearance of off-diagonal components in it. The simultaneous performance of measurements of the spectral characteristics of the modes appearing and the investigation of the solutions obtained above make it possible not only to detect these strains, but also to investigate their parameters.

SEMICONDUCTOR RODS

The magnetoplasmons appearing in semiconductor structures are utilized in the diagnostics of the surfaces of conducting materials and the creation of solid-state devices that are controlled by an external magnetic field. The propagation of waves at an angle to the latter gives rise to effects which permit the creation of structures with new functional properties. The characteristics of waves propagating in a

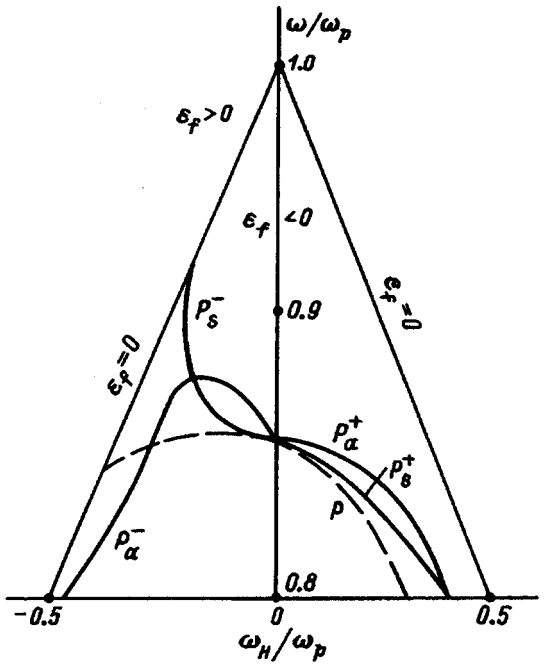

FIG. 4. Dependence of the resonant frequency ω/ω_r of surface modes on the cyclotron frequency ω_H/ω_r for $n=60$, $\Theta=\pi/4$, and $\chi=20$.

round semiconductor rod have been studied only when it is immersed in an axial magnetic field. A need has arisen to generalize the results obtained to the case of an arbitrary orientation of the external magnetic field.

For axially uniform modes ($k_z=0$) system (11) separates into

$$[\varepsilon_{11}a_f k^2-(\sigma_+-\xi\sigma_-)\varkappa^2]A_n=\varkappa k(\tau_+-\xi\tau_-)B_{n+1},$$
$$\varepsilon_{11}[\varepsilon_{33}\varepsilon_f k^2-a_f\varkappa^2]B_{n+1}=\varkappa k(\tau_+-\xi\tau_-)A_n, \qquad (22)$$

where

$$\varepsilon_f=\varepsilon_{11}+\varepsilon_{12}^2/\varepsilon_{11}, \quad a_f=\varepsilon_f C^2+\varepsilon_{33}S^2,$$

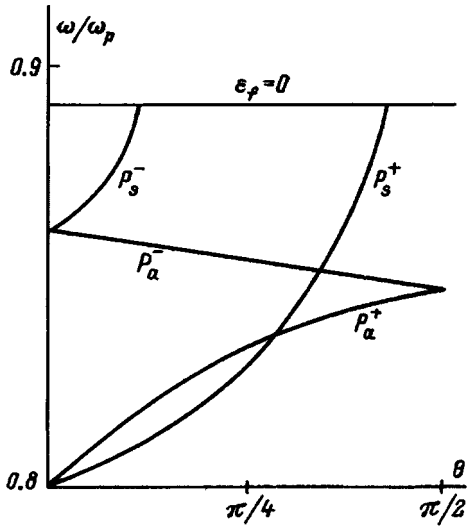

FIG. 5. Dependence of the resonant frequency of surface modes on the angle of inclination Θ for $n=60$, $\chi=20$, and $|\omega_H|=\omega_r/4$.

$$2\tau_{\pm} = -S[i\varepsilon_{33}\varepsilon_{12} \mp \varepsilon_L(\varepsilon_{33}-\varepsilon_{11})C].$$

We note from (22) that the influence of a nonaxial orientation of the magnetic field leads to the appearance of an interaction between the partial modes of E and H polarization with consecutive numbers. Symmetric ($\xi=1$) and asymmetric ($\xi=-1$) axially uniform modes appear in the structure. The radial components of the wave vectors of the symmetric modes do not depend on the angle of inclination

$$\varkappa_1 = \sqrt{\varepsilon_f}k, \quad \varkappa_2 = \sqrt{\varepsilon_{33}}k.$$

This dependence appears for asymmetric modes for which

$$2a_{11}\varkappa_{1,2}^2 = \{(\varepsilon_{33}+a_f)\varepsilon_{11} \pm [(\varepsilon_{33}+a_f)^2\varepsilon_{11}^2 - 4\varepsilon_{33}\varepsilon_{11}a_{11}a_f]^{1/2}\}k^2.$$

Figures 4 and 5 present plots of the dependence of the resonant frequencies ω/ω_r on the angle of inclination Θ and the cyclotron frequency ω_H/ω_r of the carriers for a single-component semiconductor plasma, $\varepsilon_L=17$, $n=60$, and $\chi \equiv \omega_r r_0/C=20$. In an axial magnetic field ($\Theta=0$) a series of bulk magnetoplasmons exists in the structure when $\varepsilon_f>0$. When $\varepsilon_f<0$, there is a surface magnetoplasmon (dashed curve P_s), which transforms into a bulk mode when $\varepsilon_f=0$. As the azimuthal mode index n increases, its frequency approaches the frequency of the surface magnetoplasmon existing on a planar boundary between a semiconductor and a vacuum

$$\omega = [\varepsilon_L\omega_r^2(\varepsilon_L+1)^{-1} + \omega_H^2/4]^{1/2} - \omega_H \mathrm{sgn}(n)/2.$$

A nonaxial orientation of the magnetic field leads to the appearance of the symmetric surface modes $P_s^{\pm}$ in the structure, which exist in the region corresponding to $\varepsilon_f<0$ and $a_{11}<0$, and the asymmetric modes $P_a^{\pm}$ when $\varepsilon_f<0$. Here P_j^+ is the branch of modes for $\omega_H>0$, and P_j^- is the branch of modes for $\omega_H<0$. The dependences of the resonant frequencies on the strength and angle of inclination of the external magnetic field are qualitatively different. The frequencies exhibit a significant dependence on the direction of the magnetic field (nonreciprocity). The latter is absent in a transverse magnetic field ($\omega_j^+ = \omega_j^-$).

[1] A. Ya. Kirichenko, Yu. V. Prokopenko, Yu. F. Filippov *et al.*, Radiotekh. Elektron. **33**, 300 (1989).
[2] *Surface Polaritons: Electromagnetic Waves at Surfaces and Interfaces*, V. M. Agranovich and D. L. Mills (eds.), North-Holland, Amsterdam (1982) [Russ. transl., Nauka, Moscow (1985)].

Translated by P. Shelnitz

Slowing of transverse surface waves in an isotropic plasma waveguide of noncircular cross section

V. A. Girka and I. A. Girka

Kharkov State University, Kharkov, Ukraine
(Submitted January 19, 1996)
Zh. Tekh. Fiz. **67**, 92–97 (July 1997)

[S1063-7842(97)01707-8]

INTRODUCTION

The solution of problems related to the generation of microwaves in plasma waveguides requires knowledge of the spectra of eigenfrequencies for the respective waveguide structure. This is because the excitation of waves is most efficient[1,2] near the eigenfrequencies. Slow-wave elements, viz., combs, disks, etc. are often used in the beam excitation method. They make it possible to prolong the interaction of the beam particles with the corresponding eigenwave of the structure and, as a result, to increase the efficiency of rf generators.[1–3]

The creation of multicomponent waveguide structures has provided additional possibilities for designing efficient devices for rf electronics.[4–6] The presence of various plasma and insulating inserts in a metal waveguide leads to an increase in the number of eigenwaves and, thus, to physical effects associated with their propagation and interaction. This is reflected in the design of divers radio-electronic devices for various purposes in plasma and semiconductor electronics.[1,4,6]

The possibility of using an annular electron beam to efficiently excite microwaves that propagate azimuthally near the surface of a plasma cylinder separated by a vacuum gap from a coaxial metal waveguide was demonstrated in Ref. 7. It is fairly simple to extract the energy of these azimuthal surface waves (ASW) through an insulating window of small angular dimensions in the metallic skin of the waveguide.[8] The advantages that an rf power generator employing the excitation of ASWs would have over generators that utilize longitudinal beams include, first of all, their short longitudinal dimensions and high efficiency.[7] This is the reason for the urgent interest in the problem of calculating resonant multicomponent slow-wave structures.

The slowing of electromagnetic waves propagating azimuthally near the surface of a plasma cylinder separated by an insulating layer from a metal waveguide of noncircular cross section is investigated. The distribution of the fields of these waves is studied. The correction to the eigenfrequency of an ASW caused by the small deviation of the cross section of the waveguide from a circle is determined. It is shown that the periodic variation of the curvature of the surface of the metal chamber along the direction of propagation of the wave leads to removal of the degeneracy of the frequency of the ASW with respect to the sign of the wave number, if the angular period of this variation is equal to half of the period of the wave.

FORMULATION OF THE PROBLEM

Let us investigate the propagation of a surface electromagnetic wave (SEW) near the boundary of a homogeneous plasma cylinder of radius R_0, which is separated from the metal chamber by an insulator with a dielectric constant $\varepsilon \geqslant 1$. We shall consider the case of a free cold plasma, whose dielectric constant equals

$$\varepsilon = \varepsilon_0 - \sum_{\alpha} \Omega_{\alpha}^2 \omega^{-2}, \tag{1}$$

where ε_0 is the dielectric constant of the crystal lattice of the semiconductor (in the case of a gas plasma $\varepsilon_0 = 1$), ω is the frequency of the wave, Ω_{α} is the Langmuir frequency, and α labels the type of particles: ions, electrons, or holes.

We assume that the waveguide is uniform along the axis of the cylinder: $\partial/\partial z \equiv 0$. The cross-sectional shape of the waveguide is assigned by the following dependence of the radius R_2 of the metal chamber on the azimuthal angle:

$$R_2 = R_1 \left[1 + \sum_{n=1}^{\infty} h_n \sin(n\varphi) \right]. \tag{2}$$

For the magnetic field H_z of the SEW in the plasma ($r \leqslant R_0$), from Maxwell's equations we obtain

$$\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial H_z}{\partial r} - \left(\varkappa_p^2 - \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} \right) H_z = 0, \tag{3}$$

where $\varkappa_p = k\sqrt{|\varepsilon|}$ and $kc = \omega$.

The components of the electric field of the SEW are expressed in terms of H_z in the following manner

$$E_r = \frac{i}{\varepsilon k r} \frac{d}{d\varphi} H_z, \tag{4}$$

$$E_{\varphi} = \frac{-i}{\varepsilon k} \frac{d}{dr} H_z. \tag{5}$$

Within the insulator ($R_0 < r < R_2$) the dependence of H_z, as well as of E_r and E_{φ}, on the spatial coordinates is specified by Eq. (3) and by expressions (4) and (5) after the following replacement is made in them:

$$\varepsilon \to \varepsilon_g. \tag{6}$$

On the basis of the symmetry of the problem, we seek a solution of Eq. (3) in the form

$$H_z = \sum_m H_z^{(m)}(r) \exp(im\varphi - i\omega t), \tag{7}$$

 1063-7842/97/070803-06$10.00

where m is the azimuthal mode number and φ is the azimuthal angle.

The following boundary conditions are utilized to obtain the dispersion relation: confinement of the fields to the SEW in the volume under consideration; continuity of the fields E_φ and H_z on the plasma–insulator boundary, i.e.,

$$E_\varphi(R_1+0)=E_\varphi(R_1-0), \quad H_z(R_1+0)=H_z(R_1-0); \tag{8}$$

and equality of the tangential electric field on the surface of the metal to zero, i.e.,

$$(E_\tau(R_2)\propto R_2(\varphi)E_\varphi(R_2)+E_r(R_2))\frac{dR_2}{d\varphi}$$

$$E_\tau(R_2)\equiv 0. \tag{9}$$

Assignment of the cross-sectional shape of the waveguide by Eq. (2) makes it possible to simulate any shape of a transverse section of a waveguide, for example, a rectangular shape[9] or the shape which is characteristic of the surface of the anode section of a magnetron. However, to simplify the presentation of the material, we shall confine ourselves at first to consideration of the special case

$$R_2=R_1[1+h_N\sin(N\varphi)], \tag{10}$$

and afterwards we shall generalize the results obtained to the case of an arbitrary waveguide cross section [Eq. (2)]. The choice of $R_2(\varphi)$ in the simplified form (10) is also of importance in itself. When $N=1$, this relation describes the off-center displacement (the violation of coaxiality) of the plasma column relative to the metal chamber. The case of $N=2$ describes an elliptical chamber, and $N\geqslant 3$ corresponds to a slow-wave structure with an angular period $2\pi/N$. The values of h_n, which specify the maximum deviation of the radius R_2 of the metal chamber from its mean value R_1, are assumed to be small ($h_n<R_2R_0^{-1}-1<1$). We solve the problem by successive approximations, utilizing the theory of ASWs in Ref. 10 as the zeroth approximation.

ASWs IN A WAVEGUIDE OF CIRCULAR CROSS SECTION (THE ZEROTH APPROXIMATION)

Surface electromagnetic waves propagating azimuthally in circular isotropic cylindrical plasma waveguides have been termed azimuthal surface waves (ASWs). They are E-type waves with the components E_r, E_φ, and H_z, and their properties were studied in detail in Ref. 10. Here we present only the basic information needed for the ensuing investigation. The dispersion properties of ASWs and the topology of their fields differ significantly from the properties of surface waves propagating in planar plasma waveguides. For example, ASWs are electromagnetic waves, unlike the surface potential waves that propagate along a straight plasma boundary.[11]

In the case of a free cold plasma, ASWs exist in the following frequency range

$$\varepsilon_0\omega^2<\Omega_e^2+\Omega_i^2. \tag{11}$$

The solution of Eq. (3) which satisfies the boundary condition on the axis of the waveguide for the axial component H_z of the magnetic field of an ASW within the plasma (at $r\leqslant R_0$) is expressed in terms of the modified Bessel functions $\mathbf{I}_m(\zeta)$:[12]

$$H_z^{(m)}(r)=C_1^{(m)}I_m(r\varkappa_p), \tag{12}$$

where $C_1^{(m)}$ is a normalization factor.

The solution of Eq. (3) which satisfies the boundary conditions (8) on the plasma–insulator boundary for H_z within the insulator ($R_0\leqslant r\leqslant R_2$) is expressed in terms of the Bessel functions of the first kind $J_m(\zeta)$ and the Neumann functions $N_m(\zeta)$:[12]

$$H_z^{(m)}=0.5\pi\varkappa_gR_0C_1L_m(\varkappa_gr), \tag{13}$$

where

$$L_m(\varkappa_gr)=G_1J_m(\varkappa_gr)-G_2N_m(\varkappa_gr), \tag{14}$$

$$G_1=\sqrt{\varepsilon_g/|\varepsilon|}I'_m(\varkappa_pR_0)N_m(\varkappa_gR_0) + I_m(\varkappa_pR_0)N'_m(\varkappa_gR_0), \tag{15}$$

$$G_2=\sqrt{\varepsilon_g/|\varepsilon|}I'_m(\varkappa_pR_0)J_m(\varkappa_gR_0)+I_m(\varkappa_pR_0)J'_m(\varkappa_gR_0), \tag{16}$$

a prime denotes the derivative with respect to the argument, and $\varkappa_g=k\sqrt{\varepsilon_g}$.

In the zeroth approximation ($h_N=0$) the boundary condition (9) for the tangential component of the electric field of an SEW on the insulator–metal boundary takes the form $E_\varphi(R_1)=0$. Hence we obtain the dispersion relation of an ASW[10]

$$\frac{J'_m(\varkappa_gR_0)N'_m(\varkappa_gR_1)-J'_m(\varkappa_gR_1)N'_m(\varkappa_gR_0)}{J_m(\varkappa_gR_0)N'_m(\varkappa_gR_1)-J'_m(\varkappa_gR_1)N_m(\varkappa_gR_0)} = -\left(\frac{\varepsilon_g}{|\varepsilon|}\right)^{\frac{1}{2}}\frac{I'_m(\varkappa_pR_0)}{I_m(\varkappa_pR_0)}. \tag{17}$$

Its solution in the approximation of a narrow insulating layer was studied in detail in Ref. 10; therefore, we shall not analyze it here. We note only that the dispersion of an ASW is linear (its frequency increases with the azimuthal mode number) , and, in contrast to the case of a waveguide containing a magnetically active plasma,[13] these waves are complementary (the frequencies of ASWs propagating in opposite directions are identical).

Azimuthal surface waves with different values of the azimuthal mode number m propagate independently; therefore, we shall assume that only one harmonic with the azimuthal mode number M exists in the zeroth approximation.

FIRST APPROXIMATION

In the first approximation with respect to the small parameter h_N the magnetic field of the SEW in the region $R_0<r<R_2$ is found in the form of the sum of the fundamental harmonic [which is proportional to $\exp(iM\varphi)$] and two satellite harmonics [which are proportional to $\exp(i(M\pm N)\varphi)$]:

$$H_z = 0.5\pi\varkappa_g R_0 C_1 \{L_M(\varkappa_g r)\exp(iM\varphi) + H_+ h_N L_{M+N}(\varkappa_g r)\exp[i(M+N)\varphi] + H_- h_N L_{M-N}(\varkappa_g r)\exp[i(M-N)\varphi]\}\exp(-i\omega t). \quad (18)$$

After substituting H_z in form (18) into the expressions (4) and (5) for the electric fields of the ASW [with consideration of the replacement (6)], we find that in order for H_z in form (18) to satisfy the boundary condition (9) on the noncircular metal surface (10) of a waveguide to within terms that are first-order with respect to the small parameter, the coefficients H_+ appearing in (18) must have the following values:

$$H_\pm = \pm \frac{iL_m(\varkappa_g R_1)(m^2 \pm mN - \varkappa_g^2 R_1^2)}{2\varkappa_g R_1 L'_{m\pm N}(\varkappa_g R_1)}. \quad (19)$$

Consideration of the first-order small terms does not alter the eigenfrequency of the SEW determined as the solution of dispersion relation (17).

SECOND APPROXIMATION

In the second approximation with respect to the small parameter h_N the expression for H_z contains not only the terms in Eq. (18), but also small satellite harmonics that are $\propto h_N^2 \exp(i(M\pm 2N)\varphi)$. We shall not present the expression for the amplitudes of these harmonics here, since they are cumbersome, and, moreover, these harmonics do not make a contribution to the first correction to the frequency.

After separating the terms in boundary condition (9) which are proportional to $\exp(iM\varphi)$, but are not higher than second-order with respect to the small parameter, we obtain the dispersion relation of an SEW in the form

$$D^{(0)} + D^{(2)} = 0, \quad (20)$$

where $D^{(2)} \sim h_N^2 D^{(0)}$, and the dispersion relation $D^{(0)} = 0$ in the zeroth approximation coincides with Eq. (17).

From the solution of dispersion relation (20) we determine the frequency ω of the SEW in the for of the sum $\omega = \omega_0 + \Delta\omega_N$, where ω_0 is the eigenfrequency of the SEW in the absence of corrugations $[D^{(0)}(\omega_0) = 0]$. The term

$$\Delta\omega_N = -D^{(2)}\left[\frac{\partial D^{(0)}}{\partial\omega}\right]^{-1}\Bigg|_{\omega=\omega_0} \quad (21)$$

is the second-order correction to the frequency of the SEW due to the deviation of the cross-sectional shape (10) from a circle.

The expression for $\Delta\omega_N$ has the following form:

$$\Delta\omega_N = \frac{h_N^2(m^2 + R_1^2\varkappa_g^2)}{4\varkappa_g R_1} \frac{L_m(\varkappa_g R_1)}{\dfrac{dL'_m(\varkappa_g R_1)}{d\omega_0}} \times \left[1 + \frac{L_{m+N}(\varkappa_g R_1)}{L'_{m+N}(\varkappa_g R_1)} \frac{(m^2 + mN - \varkappa_g^2 R_1^2)^2}{m^2 + \varkappa_g^2 R_1^2} \times + \frac{L_{m-N}(\varkappa_g R_1)}{L'_{m-N}(\varkappa_g R_1)} \frac{(m^2 - mN - \varkappa_g^2 R_1^2)^2}{m^2 + \varkappa_g^2 R_1^2}\right]. \quad (22)$$

Expression (22) for the correction $\Delta\omega_N$ to the frequency of the SEW does not depend on the sign of the azimuthal mode number m. This is typical of oscillations in isotropic plasma waveguides.[14] Using recurrence relations and the asymptotic expression for the Bessel functions,[12] we find approximate expressions for $\Delta\omega_N$ in the cases of broad $(\varkappa_{g,p} R_1 \gg |m|)$ and narrow $(\varkappa_{g,p} R_1 \ll 1)$ waveguides.

At short wavelengths $(kR_0 \ll 1)$ $\Delta\omega_N < 0$. We present the asymptotic expression for this correction

$$\frac{\Delta\omega_N}{\omega_0} \approx \frac{-h_N^2 \Lambda \varepsilon_g |\varepsilon|}{16\pi m!} \frac{m^2(m+N)(m+N)!}{\Lambda m^2|\varepsilon| + 2\varepsilon_g(1+|\varepsilon|)} \left(\frac{2}{\varkappa_g R_1}\right)^{2+N}, \quad (23)$$

where $\Lambda = 1 - R_0/R_1$.

This parameter, which characterizes the thickness of the insulating layer, was assumed to be small. An analysis of Eq. (22) provides evidence that reversal of the sign of $\Delta\omega_N$ can occur for sufficiently broad waveguides, i.e., when $R_1 \to \infty$, $\Delta\omega_N$ becomes positive. This occurs when the inequality

$$\Lambda \frac{R_0\Omega_e}{c} > 1 > \Lambda \frac{kR_0\varepsilon_g}{|\varepsilon_p|^{1/2}} \quad (24)$$

is satisfied.

However, it should be kept in mind that the frequency of the SEWs under investigation decreases to zero when R_1 tends to infinity and that the applicability of the analytical expression (22) for the correction to the frequency of the SEW is determined by fulfillment of the inequality $|\Delta\omega_N| \ll \omega_0$. Therefore, within the method of successive approximations used here the results obtained, primarily the expression (21) for the correction $\Delta\omega_N$ to the frequency of the SEW, have a restricted range of applicability in the case of $\Lambda R_0 \gg c/\Omega_e$. If conditions (24) are not satisfied $(1 < \Lambda k R_0 \varepsilon_g |\varepsilon_p|^{-1/2})$, i.e., if $(\varkappa_g R_0)$ is large, for $\Delta\omega_N$ we can write

$$\frac{\Delta\omega_N}{\omega_0} \approx \frac{h_N^2 \varepsilon_g}{-4}(1+\Lambda)\varkappa_g^2 R_1^2 \cdot (kR_1 \gg 1). \quad (25)$$

Comparing (23) and (25), we should note two significant differences between short-wavelength and long-wavelength perturbations. If $(\varkappa_g R_0) \ll 1$, the magnitude of $|\Delta\omega_N|/\omega_0$ decreases as ε_g increases and the number of corrugations in the waveguide, i.e., N, decreases. The rate of slowing of long-wavelength SEWs does not depend on m and N, and $|\Delta\omega_N|/\omega_0 \propto \varepsilon_g^2$. When $kR_1 \gg 1$, the frequency correction $|\Delta\omega_N|$ decreases as ε_g increases.

It is seen from the foregoing treatment that the magnitude of the correction $\Delta\omega_N$ does not depend on the presence of small terms that are $\propto \sin(n\varphi)$ $(n \neq N)$ in the equation of the metal surface (2). Therefore, in the case of the arbitrary waveguide cross section (2), the eigenfrequency can be found in the form

$$\omega = \omega_0 + \sum_{n \neq 2m} \Delta\omega_n, \quad (26)$$

where the corrections $\Delta\omega_n$ are defined by Eq. (22).

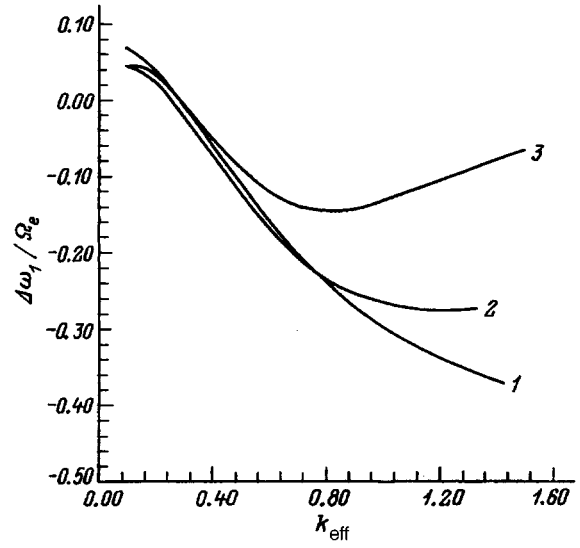

FIG. 1.

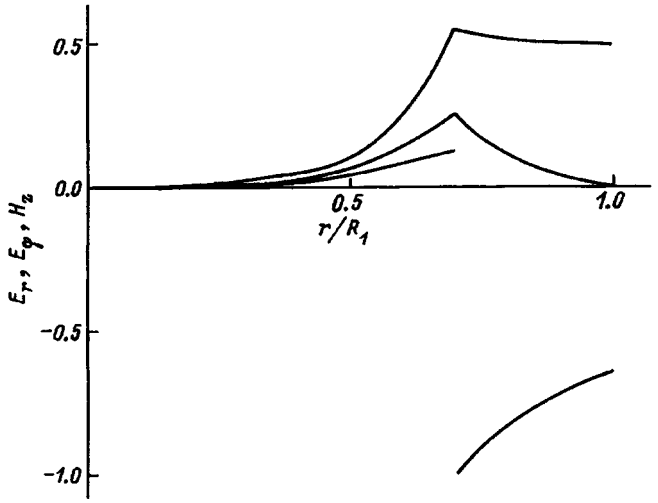

FIG. 3.

NUMERICAL ANALYSIS

The results of a numerical investigation of the magnitude of the correction $(\Delta\omega_n/\Omega_e)$ to the frequency of an SEW caused by the deviation of the waveguide shape from a cylinder as a function of the effective wave number $k_{\text{eff}} \equiv (mc)/(R_0\Omega_e)$ are presented in Figs. 1 and 2 for the cases of $N=1$ and 4, respectively. The numbers of the curves correspond to the mode number m. To make the figures clearer, the parameter of the problem was selected equal to unity: $h_n=1$ (as was noted, the h_n should, in fact, be small). The following values of the waveguide parameters were used: $\Lambda=0.3$, $\varepsilon_g=1$. The numerical calculations show that the thickness of the insulating layer, i.e., the parameter Λ, has the greatest influence on the magnitude of the correction $\Delta\omega_n(k_{\text{eff}})$. For example, an increase in ε_g from 1 to 5 (without altering the other parameters) causes an approximately twofold decrease in both ω_0 and $\Delta\omega_n$, the point $\Delta\omega_n=0$ being displaced toward larger values of the effective wave number. In general, an increase in ε_g decreases the amplitude of an SEW and its frequency, i.e., its propagation conditions worsen. A decrease in Λ from 0.3 to 0.1 (without altering the other parameters) causes a slight decrease in the eigenfrequency of an SEW, but the correction $|\Delta\omega_n(k_{\text{eff}})|$ increases by approximately seven fold for the SEW with $m=1$ and by two fold for the wave with $m=2$. This increase in the absolute value of the correction to the frequency of an SEW becomes less significant as k_{eff} and the azimuthal mode number m increase. It should be noted that the $\Delta\omega_n(k_{\text{eff}})$ curves become more gently sloping as Λ decreases and that the absolute value of the correction increases as k_{eff} increases. An increase in the number of corrugation periods (the parameter N) leads to stronger slowing of the SEWs.

The topology of the field of the SEW within a waveguide for $\Lambda=0.3$, $m=3$, $\varepsilon_g=1$, and $\varphi=0$ is shown in Fig. 3. The amplitude of the SEW is defined here in relative units, and the radial coordinate is normalized to the mean value of the waveguide radius R_1. Since the surface waves investigated are electromagnetic, the magnetic component of the field is greater than or of the order of the electric component, if the frequency of the surface wave lies in the middle of the range for the existence of ASWs. An increase in ω, of course, leads to an increase in $|H_z|$. When R_0 is increased, $|H_z|$ decreases together with the frequency of the SEW. We note that the field E_r has its greatest magnitude in the region of the insulator $(R_0<r<R_2)$ (it does not tend to zero as $r\to R_2$), while the field E_φ induced by the surface charges reaches a maximum on the boundary of the plasma cylinder with the insulating layer $(r=R_0)$.

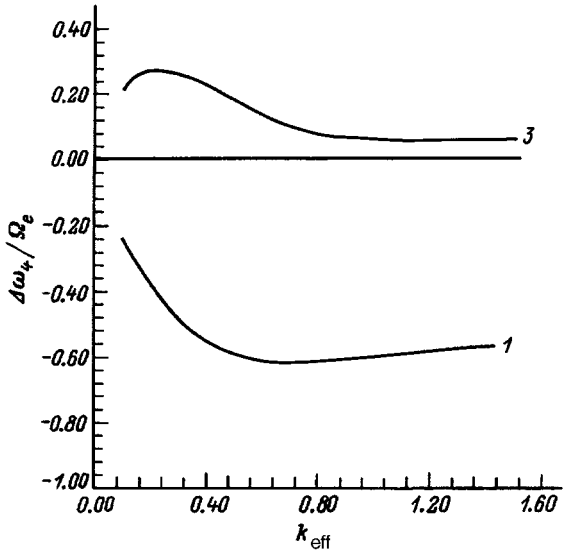

FIG. 2.

SPLITTING OF THE SPECTRA OF ASWs

The spectra of the ASWs for isotropic waveguides are degenerate relative to the sign of m. For this reason, when the relation $N=2|M|$ holds, the treatment performed is not applicable. For example, the expression (22) for the correction $\Delta\omega_N$ to the frequency becomes meaningless in this case, since its denominator contains the quantities $L'_{M\pm N}(\varkappa_g R_1)$,

which vanish by virtue of the dispersion relation (17) in the zeroth approximation. Applying perturbation theory to the case of degenerate spectra, we obtain an expression for the resonant correction $\Delta\omega_1^{(\pm)} \equiv \pm\Delta\omega_r$ to the frequency of an ASW in this case:

$$\Delta\omega_r = \frac{h_{2|M|}L_M(\varkappa_g R_1)}{2\varkappa_g R_1}(M^2+\varkappa_g^2 R_1^2)\left(\frac{dL'_M(\varkappa_g R_1)}{d\omega_0}\right)^{-1}. \tag{27}$$

This expression is simplified in the case of a plasma cylinder of large dimensions ($\varkappa_p R_0 \gg 1$):

$$\Delta\omega_r = -\frac{h_{2|m|}}{4}\left(1+\frac{\Lambda R_0}{\delta}\varepsilon_g^{\frac{1}{2}}\right)m\left(\frac{\delta}{\Lambda R_0}\right)^{\frac{1}{2}}\varepsilon_g^{-\frac{1}{4}}. \tag{28}$$

Here $\delta = c/\Omega_e$. As we see, periodic variation of the curvature of the metal surface along the direction of wave propagation results in splitting of the frequency of an ASW. The resonance correction to the frequency is proportional to the first power of the small parameter ($\Delta\omega_r \propto h_{2|m|}$). As a consequence of the resonant interaction with the nonuniqeness of the surface, the ASW harmonics that are $\propto \exp(i|m|\varphi)$ and $\propto \exp(-i|m|\varphi)$ exist in the form of two standing waves with two different frequencies $\omega = \omega_0 \pm \Delta\omega_r$. Accordingly, with consideration of the terms that are first-order with respect to the small parameter, instead of (18) the distribution of H_z in the region $R_0 < r < R_2$ takes on the form

$$\begin{aligned} H_z = {} & 0.5\pi\varkappa_g R_0 \exp(-i\omega t) \\ & \times\{C_{11}L_m(\varkappa_g r)[\exp(im\varphi) \pm i\exp(-im\varphi)] \\ & + C_r^+ L_3(\varkappa_g r)[\exp(i3m\varphi) \mp i\exp(-i3m\varphi)]\}. \end{aligned} \tag{29}$$

Here C_{11} is a normalization factor, and

$$C_r^+ = -\frac{ih_{2|m|}L_m(\varkappa_g R_1)}{2\varkappa_g R_1 L_{3m}(\varkappa_g R_1)}(3m^2-\varkappa_g^2 R_1^2)C_{11}. \tag{30}$$

Figure 4 presents plots of the dependence of the correction to the frequency of an SEW normalized to the plasma frequency ($-\Delta\omega_r/\Omega_e$) in the case of $N = 2|m|$ on the effective wave number k_{eff}. The numbers *1–3* correspond to the azimuthal mode number m, the solid curves are plots of the dependence of (ω_0/Ω_e) on k_{eff}, and the dashed curves are plots of the dependence of ($-\Delta\omega_r/\Omega_e$) on k_{eff}. The calculations were performed for the following values of the waveguide parameters: $\varepsilon_g = 1$, $h_v = 0.5$, and $\Lambda = 0.3$. The numerical analysis confirms that the resonance correction increases as the thickness of the insulating layer or the dielectric constant of the insulator decreases.

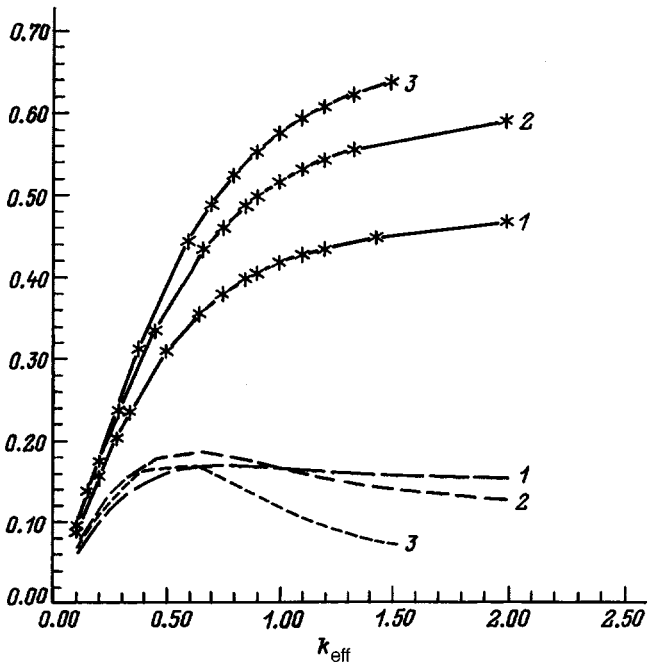

FIG. 4.

CONCLUSIONS

The influence of the deviation of the cross-sectional shape (2) of the metal chamber of an isotropic plasma waveguide from a circle on the dispersion properties of extraordinary surface electromagnetic waves propagating azimuthally has been studied theoretically in the present work. The distribution of the fields of the SEWs within waveguides has been investigated [see Eqs. (18), (19), (29), and (30) and Fig. 3]. It has been shown that the spectral composition of the SEWs propagating in the metal chamber of a waveguide can be regulated by adjusting its shape. The dependence of the correction to the frequency of an SEW due to the deviation of the shape of the metal chamber from a circle on the waveguide parameters has been analyzed [see formulas (22), (23), and (25), and Figs. 1 and 2]. It has been shown that periodic variation of the curvature of the waveguide along the direction of propagation of an SEW leads to splitting of the frequency of the SEW, after which the angular period of the fundamental harmonic is two times greater than the period of this variation [see Eqs. (27) and (28) and Fig. 4]. In this case the SEW exists in the form of the standing wave described by (29) and (30).

The electrodynamic structure investigated in the present work and its eigenwaves can be utilized in various radio-electronic devices, for example in rf emitters and generators. If an antenna in the form of a wire is placed along the symmetry axis of a waveguide and the radius of the wire is small (this condition is necessary so that its influence on the electromagnetic properties of the waveguide could be neglected), such a structure should be regarded as an antenna with a plasma covering. The use of a plasma covering in antennas has been known for a long time and is employed in practice. For example, calculations of the radiation of a spheroidal antenna with a plasma covering were presented in Ref. 15. The presence of this covering results in an increase in the radiated power. The presence of a plasma in the working volume of a standing-wave electronic tube leads to an increase in the power of the generator.[16] Thus, the results obtained in the present work are of interest for plasma electronics, radiophysics, and the development of radio-electronic devices.

[1] N. S. Erokhin, M. V. Kuzelev, and S. S. Moiseev *et al.*, *Nonequilibrium and Resonant Processes in Plasma Radiophysics* [in Russian], Nauka, Moscow (1982).
[2] A. A. Rukhadze, L. S. Bogdankevich, S. E. Rosinskiĭ, and V. G. Rukhlin, *Physics of Intense Relativistic Electron Beams* [in Russian], Atomizdat, Moscow (1980).
[3] L. S. Bogdankevich, M. V. Kuzelev, and A. A. Rukhadze, Zh. Tekh. Fiz. **50**, 233 (1980) [Sov. Phys. Tech. Phys. **25**, 143 (1980)].
[4] P. I. Markov, I. N. Onishchenko, G. V. Sotnikov, and Ya. B. Faĭnberg, Fiz. Plazmy **19**, 14 (1993) [Plasma Phys. Rep. **19**, 7 (1993)].
[5] E. P. Kurushin and E. I. Nefedov, *Electrodynamics of Anisotropic Waveguide Structures* [in Russian], Nauka, Moscow (1983).
[6] N. N. Beletskiĭ, A. A. Bulgakov, S. I. Khankina, and V. M. Yakovenko, *Plasma Instabilities and Nonlinear Phenomena in Semiconductors* [in Russian], Naukova Dumka, Kiev (1984).
[7] V. A. Girka, I. A. Girka, V. P. Olefir, and V. I. Tkachenko, Pis'ma Zh. Tekh. Fiz. **17**(1), 87 (1991) [Sov. Tech. Phys. Lett. **17**, 35 (1991)].
[8] V. A. Girka and I. A. Girka, Radiotekh. Elektron. **37**, 419 (1992).
[9] I. A. Girka and A. V. Zolotukhin, Radiotekh. Elektron. **39**, 1961 (1994).
[10] V. A. Girka, I. A. Girka, A. N. Kondratenko, and V. I. Tkachenko, Radiotekh. Elektron. **34**, 1527 (1989).
[11] N. A. Azarenko and V. V. Kostenko, Radiotekh. Elektron. **33**, 1027 (1988).
[12] E. Jahnke, F. Emde, and F. Lösch, *Tables of Higher Functions*, 6th ed., McGraw-Hill, New York (1960) [Russ. transl., Nauka, Moscow (1964)].
[13] V. A. Girka and I. A. Girka, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **34**, 386 (1994).
[14] A. N. Kondratenko, *Surface and Bulk Waves in a Confined Plasma* [in Russian], Énergoatomizdat, Moscow (1985).
[15] N. M. Roslyakov and N. A. Teplyakova, Radiotekh. Elektron. **37**, 583 (1992).
[16] G. Mishra, V. K. Tripathi, and V. K. Jain, IEEE Trans. Electron Devices **37**, 1561 (1990).

Translated by P. Shelnitz

Relativistic particles in a quadrupole waveguide

Yu. G. Pavlenko, N. D. Naumov, and A. I. Toropova

Specialized Educational-Scientific Center, M. V. Lomonosov Moscow State University, 121357 Moscow, Russia
(Submitted December 20, 1995)
Zh. Tekh. Fiz. **67**, 98–102 (July 1997)

A solution is obtained for the equations of motion of a particle in the field of a traveling electromagnetic wave propagating between two pairs of metal surfaces of hyperbolic shape. The motion of the particle on excitation of the waveguide by monochromatic and pulsed voltage sources is considered. The possibility of applying a quadrupole waveguide to beam focusing and the separation of particles according to their specific charge and longitudinal velocity is investigated. © *1997 American Institute of Physics.* [S1063-7842(97)01807-2]

INTRODUCTION

Only a few integrable problems of relativistic mechanics associated with the motion of charged particles in the field of an electromagnetic wave are presently known. A solution for a particle in the field of a plane wave was found in Ref. 1; the more complicated problem of a particle propagating along a uniform magnetic field in the field of a wave of circular polarization was solved in Ref. 2. An exact solution of the problem of the motion of a relativistic particle in the field of a plane wave in a medium with a refractive index $n>1$ was obtained in Ref. 3.

In the present work we obtain a new analytical solution for the equations of motion of a particle in the field of a transverse-electromagnetic (TEM) wave propagating in a quadrupole waveguide. The stable motion of particles is investigated both in the case of a monochromatic wave and in the case of excitation of the waveguide by pulsed voltage sources. In the former case the solution is represented by bounded aperiodic Mathieu functions, and in the latter case it is described by bounded piecewise-continuous functions. The solutions obtained enable us to determine the conditions for the passage of a particle beam through a waveguide, as well as to examine the conditions for focusing and the separation of particles according to their specific charge and longitudinal velocity. There is special interest in the motion of relativistic particles injected in the direction opposite to the direction of wave propagation. In this case it becomes possible to suppress the contribution associated with the spread of the transverse velocities of the beam particles to the amplitude of the oscillations.

ELECTROMAGNETIC WAVEGUIDE FIELD

The electrodynamic system is formed by two pairs of metal surfaces described by the equations $y^2=x^2-R^2$ and $y^2=x^2+R^2$. A fundamental TEM wave, whose critical frequency is equal to zero, can be excited in this system. This wave will propagate in the waveguide with the velocity of light at any frequency of the field excitation source. A unique feature of the open system under consideration is the fact that the configurations of the field of the fundamental wave and the field in the static case coincide.[4]

The solution of Maxwell's equations for the electric field strength and the magnetic induction is represented in terms of the potentials

$$\mathbf{E}=-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}-\nabla A_0, \quad \mathbf{B}=\operatorname{curl}\mathbf{A}. \tag{1}$$

For a transverse-electromagnetic wave these potentials are specified by the Hertz potential $G(t,\mathbf{x})$:

$$A_0=\frac{\partial G}{\partial z}, \quad \mathbf{A}=\left(0,0,-\frac{1}{c}\frac{\partial G}{\partial t}\right). \tag{2}$$

The function $G(t,\mathbf{x})$ satisfies the scalar wave equation and the boundary conditions on the waveguide surface. For the fundamental wave

$$G(t,\mathbf{x})=-\frac{\sigma}{R^2}(x^2-y^2)\int f(\xi)\,d\xi, \tag{3}$$

where $\xi=t-\sigma z/c$, $f(\xi)$ is an arbitrary function, $\sigma=1$ corresponds to a wave propagating in the positive direction of the z axis, and $\sigma=-1$ corresponds to a wave travelling in the opposite direction.

From Eqs. (1)–(3) we find

$$\mathbf{E}(t,\mathbf{x})=\frac{2}{R^2}(-x,y,0)f(\xi), \quad \mathbf{B}(t,\mathbf{x})=-\frac{2\sigma}{R^2}(y,x,0)f(\xi).$$

We next restrict ourselves to the case in which $f(\xi)$ is a periodic function. Then the intensity of the wave propagating in the waveguide is determined by the mean value of the z component of the Poynting vector

$$P_z=\sigma I, \quad I=\frac{c}{2\pi R^2}\oint dS(x^2-y^2)\langle f^2\rangle.$$

SEPARATION OF VARIABLES IN THE EQUATIONS OF MOTION

In relativistic mechanics it is convenient to assign the equation of a trajectory in the parametric form $t=t(\tau)$, $\mathbf{x}=\mathbf{x}(\tau)$, where τ is the proper time of the particle. The equations of motion of a particle of charge e and mass m have the form[5,6]

 1063-7842/97/070809-05$10.00

$$mc^2\ddot{t}=e\dot{\mathbf{x}}\cdot\mathbf{E},\quad m\ddot{\mathbf{x}}=e\dot{t}\mathbf{E}=\frac{e}{c}(\dot{\mathbf{x}}\times\mathbf{B}), \tag{4}$$

where a dot denotes differentiation with respect to the proper time.

According to the definition of four-velocity,

$$(c\dot{t})^2-\dot{\mathbf{x}}^2=c^2. \tag{5}$$

We choose the initial conditions in the form $t(0)=0$, $\mathbf{x}(0)=(x_0,y_0,0)$, $\dot{t}(0)=\gamma$, $\dot{\mathbf{x}}(0)=\gamma(v_1,v_2,v_3)$, and $v_3>0$, where $\gamma=(1-v_iv_i/c^2)^{-1/2}$. Substituting the expressions for the fields into (4), we obtain the system of equations

$$mc\ddot{t}=-(x\dot{x}-y\dot{y})F(\xi),\quad m\ddot{z}=-\sigma(x\dot{x}-y\dot{y})F(\xi), \tag{6a}$$

$$m\ddot{x}=-cx\dot{\xi}F(\xi),\quad m\ddot{y}=cy\dot{\xi}F(\xi), \tag{6b}$$

where $F=2ef/cR^2$.

A separation of variables must be performed to integrate system (6) in quadratures. Owing to the symmetry of the field, the first integral follows from Eqs. (6a):

$$\dot{t}-\frac{\sigma}{c}\dot{z}=g,\quad g=\gamma\left(1-\frac{\sigma}{c}v_3\right). \tag{7}$$

Therefore, on the trajectory of the particle

$$t-\frac{\sigma}{c}z=g\tau. \tag{8}$$

Taking into account relations (7) and (8), from (6b) we obtain the equations

$$\ddot{x}+\varkappa(\tau)x=0, \tag{9}$$

$$\ddot{y}-\varkappa(\tau)y=0, \tag{10}$$

where $\varkappa(\tau)=gF(g\tau)$.

From the first integrals (5) and (7) we find

$$\dot{t}=\frac{1}{2}\left(\frac{1}{g}+g\right)+\frac{1}{2gc^2}(\dot{x}^2+\dot{y}^2), \tag{11}$$

$$\frac{\sigma}{c}\dot{z}=\frac{1}{2}\left(\frac{1}{g}-g\right)+\frac{1}{2gc^2}(\dot{x}^2+\dot{y}^2). \tag{12}$$

Thus, after integration of linear equations (9) and (10), the functions $t(\tau)$ and $z(\tau)$ are determined from Eqs. (11) and (12). Transforming (11) and (12), we obtain the equivalent relations

$$\dot{t}=\gamma+\frac{1}{2gc^2}[\dot{x}^2+\dot{y}^2-\gamma^2(v_1^2+v_2^2)],\quad \dot{z}=\gamma v_3+\sigma c(\dot{t}-\gamma).$$

It follows from (11) and (12) that for $|\dot{x}/c|$, $|\dot{y}/c|\gg\sqrt{2\gamma g}$ the kinetic energy of the particle $T=mc^2(\dot{t}-1)\gg\gamma mc^2$. In this case the components of the velocity $\mathbf{v}=\dot{\mathbf{x}}/\dot{t}$ take the values $|v_x/c|$, $|v_y/c|\ll\sqrt{g/2\gamma}$, and $v_z\sim\sigma c$. When $|\dot{x}/c|$, $|\dot{y}/c|\ll\sqrt{2\gamma g}$, the kinetic energy $T\sim\gamma mc^2$.

A PARTICLE IN THE FIELD OF A MONOCHROMATIC WAVE

Let the function $f(\xi)$ have the form $f=V\cos\omega\xi$, where V specifies the intensity of the wave. Going over to the dimensionless quantities $s=\omega g\tau/2$ and $b=4eV/mg\omega^2R^2$ in (9) and (10), we obtain the equations

$$\frac{d^2x}{ds^2}+2b\cos 2sx=0, \tag{13}$$

$$\frac{d^2y}{ds^2}-2b\cos 2sy=0, \tag{14}$$

which are special cases of the Mathieu equation

$$\frac{d^2u}{ds^2}+(\mu+2q\cos 2s)u=0.$$

It is known from the theory of Mathieu functions that the parameter plane (μ,q) contains regions corresponding to bounded and unbounded solutions.[7–9] In the stability region the solution is an aperiodic function. When $\mu=0$, the solutions are stable in the ranges $0<b\leqslant 0.92$, $7.51\leqslant b\leqslant 7.58$, $21.309\leqslant b\leqslant 21.312,\ldots$. These conditions hold when the parameters of the wave are appropriately selected. Representing b in the form $b=(eV/mgc^2)(\lambda/\pi R)^2$, we find that in the case of an electron beam the wavelength equals $\lambda\simeq 2.22\times 10^3R\sqrt{bg/V}$ and that for a beam of ions with the mass number A we have $\lambda\simeq 10^5R\sqrt{bgA/V}$, where λ and V are numerical values in SI units.

a) Let us first consider the region $0<b\leqslant b_0$, $b_0=0.92$. The first stability region in the (μ,q) plane lies to the right of the curve $\mu_{c0}=-q^2/2+7q^4/128+\ldots$ and is bounded by the curve $\mu_{c1}(q)=1-q-q^2/8$ for $q>0$ and by the curve $\mu_{s1}(q)=\mu_{c1}(-q)$ for $q<0$. A uniformly convergent expansion of the solution of the Mathieu equation is obtained using averaging in the Hamiltonian form.[6] In the vicinity of the transitional $\mu_{c0}(q)$ curve we obtain

$$u(s)=U\left(1+\frac{q}{2}\cos 2s\right)-\frac{q}{2}W\sin 2s+\ldots,$$

$$\frac{du}{ds}=W\left(1-\frac{q}{2}\cos 2s\right)-Uq\sin 2s+\ldots. \tag{15}$$

Here the evolution of the variables U and W is specified by the averaged motion

$$U=A\cos\Theta s+\frac{B}{\Theta}\sin\Theta s,$$

$$W=\frac{dU}{ds},$$

where $\Theta=\sqrt{\mu-\mu_{c0}}$, and A and B are constants.

In accordance with Ref. 9, one of the linearly independent solutions on the transitional $\mu_{c0}(q)$ curve contains the multiplier s. We note that the solution of the Mathieu equation obtained within the Gaponov-Miller gradient potential model[10,11] or by the Kapitsa method[12] contains only the first term in (15) $u(s)=U(s)$, and $\mu_{c0}=-q^2/2$. Setting $\mu=0$ and $q=b$ in (15), we obtain a solution of Eq. (13); if $\mu=0$

and $q=-b$, we have a solution of Eq. (14) with the coefficients A' and B'. Taking into account the initial conditions, we find

$$A=x_0\left(1+\frac{b}{2}\right)^{-1}, \quad B=\frac{2\gamma v_1}{\omega g}\left(1-\frac{b}{2}\right)^{-1},$$

$$A'=y_0\left(1-\frac{b}{2}\right)^{-1}, \quad B'=\frac{2\gamma v_2}{\omega g}\left(1+\frac{b}{2}\right)^{-1}.$$

The realization of stable motion requires that all the beam particles satisfy the inequalities $|x(\tau)|<R$ and $|y(\tau)|<R$. In the utlrarelativistic case, i.e., when $v_3\sim c$ and $|v_1|$, $|v_2|\ll v_3$, g is strongly dependent on σ: $g(\sigma=1)\simeq 1/2\gamma$, and $g(\sigma=-1)\simeq 2\gamma$. Therefore, in the case of a wave propagating toward a beam, there is a possibility for suppressing the contribution caused by the spread of the transverse velocities to the amplitude of the oscillations. From the inequalities $|B|/\Theta_0<|A|<R$ and $|B'|/\Theta_0<|A'|<R$ ($\Theta_0\simeq b\sqrt{2}$) we can obtain estimates for the upper bound on the initial values of the transverse velocity components

$$\left|\frac{v_1}{c}\right|<2\pi b(2-b)|x_0|[\lambda(2+b)\sqrt{2}]^{-1},$$

$$\left|\frac{v_2}{c}\right|<2\pi b(2+b)|y_0|[\lambda(2-b)\sqrt{2}]^{-1},$$

where $|x_0|$, $|y_0|<R$.

Setting $R=10^{-3}$ m, $V=10^3$ V, and $b=0.125$, for elections in the Stanford Linear Collider with an energy of 50 GeV we obtain $\lambda\simeq 11$ m and $|v_1/c|$, $|v_2/c|\leqslant 5\times 10^{-5}$.

b) Let us now find the solutions of Mathieu's equation in the vicinity of the $\mu_{s1}(q)$ and $\mu_{c1}(q)$ curves for $b\sim b_0$ by the averaging method. The uniformly convergent expansion of the general solution has the form

$$u(s)=\frac{1}{2}\,\mathrm{Re}[(Ar^{1/2}+iBr^{-1/2})(\alpha_1 F+\alpha_2 F^*)\exp(ins)]. \tag{16}$$

Here A and B are constants; $\delta=\mu-1$; $r=n/(l+k)$; $n=\sqrt{l^2-k^2}$ is the characteristic exponent;

$$2l=-\delta+\frac{\delta^2}{4}+\frac{q^2}{8}+\ldots, \quad 2k=q\left(1-\frac{\delta}{2}+\ldots\right),$$

$$\alpha_{1,2}=r^{-1/2}\pm r^{1/2},$$

$$F=\left(1-\frac{\delta}{4}\right)\exp(-is)-\frac{q}{4}\exp(is)+\frac{q}{8}\exp(-3is)+\ldots.$$

The solution in the instability region follows from (16) after the replacement $n\to -i\sqrt{k^2-l^2}$. In this approximation the derivative equals

$$\frac{du}{ds}=\frac{1}{2}\,\mathrm{Im}[(Ar^{1/2}+iBr^{-1/2})(\alpha_1 G+\alpha_2 G^*)\exp(ins)],$$

$$G=\left(1+\frac{\delta}{4}\right)\exp(-is)-\frac{q}{4}\exp(is)+\frac{3q}{8}\times\exp(-3is)+\ldots.$$

On the transitional $\mu=\mu_{s1}$ curve we have $l+k=0$, and on the curve $\mu=\mu_{c1}$ we have $l-k=0$. Setting $\mu=0$ and $q=b$ in (16), we obtain the solution of Eq. (13)

$$x(\tau)=\left(A\cos ns-\frac{B}{r}\sin ns\right)F_1-(Ar\sin ns+B\cos ns)F_2, \tag{17}$$

where $F_1=\mathrm{Re}\,F$ and $F_2=\mathrm{Im}\,F$ are periodic functions with the period π.

If we set $\mu=0$ and $q=-b$ and perform the replacement $r\Rightarrow r'=n/(l-k)$ in (16), it becomes the solution of Eq. (14). It is noteworthy that the extremum values of $x(\tau)$ and $y(\tau)$ increase in the vicinity of $b\sim b_0$, since in this case $n\sim r\sim\sqrt{b-b_0}$ and $r'\sim 1/\sqrt{b-b_0}$. The constants are found from the initial conditions:

$$A=4x_0\left(5-\frac{b}{2}\right)^{-1}, \quad B=\frac{8\gamma v_1}{\omega g}\left(3+\frac{b}{2}\right)^{-1},$$

$$A'=4y_0\left(5+\frac{b}{2}\right)^{-1}, \quad B'=\frac{8\gamma v_2}{\omega g}\left(3-\frac{b}{2}\right)^{-1}.$$

The realization of stable motion requires fulfillment of the conditions $|B|/r<|A|<R$ and $|A'|r'<|B'|<R$.

A regime of velocity confinement of the beam is possible in the stability region under consideration. If $|v_1/c|$, $|v_2/c|\ll 1$, the beam of ions of charge Ze_0 passing through the waveguide in the case of $\sigma=1$ will consist of particles whose initial longitudinal velocities satisfy the condition $v_3\leqslant c\beta(\omega)$ with $\beta=(\omega^4-\Omega^4)(\omega^4+\Omega_0^4)^{-1}$, where $\Omega^2=4Ze_0V/mbR^2$, $\Omega_0^2=4Ze_0V/mb_0R^2$, and $\Omega>\Omega_0$. As the frequency increases in the range $\Omega_0\leqslant\omega\leqslant\Omega_0(2/\varepsilon-1)^{1/4}$, β takes values in the range $0\leqslant\beta\leqslant 1-\varepsilon$. If this beam passes through a second waveguide, where $\omega'\geqslant\omega$ and $\sigma=-1$, ions for which $v_3\geqslant c\beta(\omega')$ appear at the exit.

c) In the stability region $b_L\leqslant b\leqslant b_R$, where $b_L=7.51$ and $b_R=7.58$, the initial values of the longitudinal velocities of the particles passing through a waveguide with a fixed frequency in the case of $\sigma=1$ satisfy the condition

$$\beta_1(\omega)\leqslant\frac{v_3}{c}\leqslant\beta_2(\omega),$$

where $\beta_i=(\omega^4-\Omega_i^4)/(\omega^4+\Omega_i^4)$, $\Omega_{1,2}^2=4Ze_0V/mR^2b_{L,R}$, and $\omega\geqslant\Omega_1$.

The parameter β_2 can be represented in the form $\beta_2=\beta_1+\Delta\beta$, and

$$\Delta\beta=\frac{1}{2}\left(1-\frac{b_L^2}{b_R^2}\right)(1-\beta_1^2)\simeq 5\times 10^{-3}(1-\beta_1^2).$$

We note that for $\omega=\Omega_1(2/\varepsilon-1)^{1/4}$ we have $\beta_1=1-\varepsilon$.

MOTION WITH PULSED EXCITATION OF THE WAVEGUIDE

Let us now consider the motion of a particle in a wave whose structure is shaped by a periodic sequence of pulses of opposite polarity with the amplitudes V_1 and V_2 and the durations t_1 and t_2. Let t_1, $t_2\ll T$, where $T=2\pi/\omega$ is the

period of the wave. In this very simple case the function $f(\xi)$ can be represented in terms of a Dirac δ function:

$$f(\xi)=f_1\sum_{n=1}^{N}\delta[\xi-(n-1)T]-f_2\sum_{n=1}^{N}\delta\left[\xi-\left(n-\frac{1}{2}\right)T\right],$$

where $f_i=t_iV_i$.

Then in Eqs. (9) and (10) we have

$$\varkappa(\tau)=\frac{2e}{mR^2}\left(f_1\sum_{n=1}^{N}\delta[\tau-(n-1)T_0]-f_2\sum_{n=1}^{N}\delta\left[\tau-\left(n-\frac{1}{2}\right)T_0\right]\right),$$

where $T_0=T/g$.

The solution of Eq. (9) is a piecewise-continuous function. The magnitude of the break in the derivative is determined by integrating the equation in the vicinity of the points $(n-1)T_0$ and $(n-1/2)T_0$. Taking into account the boundary conditions, for $n\geqslant 1$ we find

$$x(\tau)=A_i^{(n-1)}u_i(\tau,n-1),\quad \left(n-\frac{3}{2}\right)T_0\leqslant\tau\leqslant(n-1)T_0,$$

$$x(\tau)=\alpha_{ik}A_k^{(n-1)}u_i(\tau,n),\quad (n-1)T_0\leqslant\tau\leqslant\left(n-\frac{1}{2}\right)T_0,$$

$$x(\tau)=A_i^{(n)}u_i(\tau,n),\quad \left(n-\frac{1}{2}\right)T_0\leqslant\tau\leqslant nT_0. \tag{18}$$

Here $u_1(\tau,n)=1$, $u_2(\tau,n)=\tau/T_0-n+1$, $A_i^{(n)}=M_{ik}A_k^{(n-1)}$, the matrix $M_{ik}=\beta_{ij}\alpha_{jk}$, $\alpha_{11}=\alpha_{12}=1$, $\alpha_{21}=-2b_1$, $\alpha_{22}=1-2b_1$, $\beta_{11}=1-b_2$, $\beta_{12}=-b_2/2$, $\beta_{21}=2b_2$, $\beta_{22}=1+b_2$, and the $b_i=4e\pi^2V_i/mgs_i\omega^2R^2$ are constants, where $s_i=T/t_i$ is the off-on time ratio. Because of the initial conditions $A_1^{(0)}=x_0$ and $A_2^{(0)}=\gamma v_1/T_0$.

It is clear that $A_k^{(n)}=C_{kj}A_j^{(0)}$, where the matrix $C=M^n$ and $\det M=1$. To calculate C and investigate the behavior of the solution, we find the eigenvalues λ of the matrix M:

$$M_{11}=1+b_1b_2-b_2,\quad M_{12}=1+b_1b_2-\frac{3}{2}b_2,$$

$$M_{21}=-2(b_2-b_1-b_1b_2),$$

$$M_{22}=1-2b_1+3b_2-2b_1b_2.$$

A similar problem arises in finding the regions for the transmission of optical radiation by multilayer media.[13] From the equality $\det(M-\lambda I)=0$ we obtain

$$\lambda_{1,2}=\frac{1}{2}\mathrm{Tr}M\pm i\left[1-\left(\frac{1}{2}\mathrm{Tr}M\right)\right]^{1/2}.$$

The condition for the bounds on the function $x(\tau)$ has the form $|\mathrm{Tr}M|\leqslant 2$. Setting $\mathrm{Tr}M=2\cos\delta$, we have $\lambda_{1,2}=\exp(\pm i\delta)$. We next introduce the projection operators

$$P_1=(M-\lambda_2 I)(\lambda_1-\lambda_2)^{-1},$$

$$P_2=(M-\lambda_1 I)(\lambda_1-\lambda_2)^{-1},$$

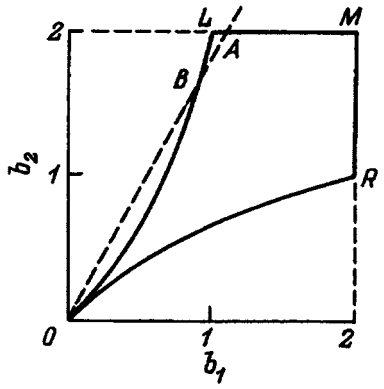

FIG. 1. Stability region.

which satisfy the conditions $P_i^2=P_i$, $P_1P_2=P_2P_1=0$, and $P_1+P_2=I$.[14] Since $M=\lambda_iP_i$, we have $C=M^n=\lambda_1^nP_1+\lambda_2^nP_2$ or

$$C_{11}=\frac{1}{\sin\delta}[M_{11}\sin n\delta-\sin(n-1)\delta],$$

$$C_{12}=M_{12}\frac{\sin n\delta}{\sin\delta},\quad C_{21}=M_{21}\frac{\sin n\delta}{\sin\delta},$$

$$C_{22}=\frac{1}{\sin\delta}[M_{22}\sin n\delta-\sin(n-1)\delta]. \tag{19}$$

The solution of Eq. (10) follows from expressions (18) after the replacements $A_k^{(n)}\Rightarrow B_k^{(n)}=C'_{kj}A_j^{(0)}$, $B_1^{(0)}=y_0$, and $B_2^{(0)}=\gamma v_2/T_0$. The elements of $C'=(M')^n$ are specified by Eqs. (19), where M' is the matrix M in which the replacements $b_1\Rightarrow -b_1$ and $b_2\Rightarrow -b_2$ have been made; accordingly, $\delta\Rightarrow\delta'$, where $2\cos\delta'=\mathrm{Tr}M$.

The common stability region of the solutions of Eqs. (9) and (10), which is specified by the system of inequalities

$$\left|1-b_1+b_2-\frac{1}{2}b_1b_2\right|\leqslant 1,\quad \left|1+b_1-b_2-\frac{1}{2}b_1b_2\right|\leqslant 1,$$

has the form of the curvilinear quadrangle $OLMR$ in the (b_1,b_2) plane (Fig. 1), which is bounded by the straight lines $b_1=2$ and $b_2=2$ and by the curves $b_2=h(b_1)$ and $b_2=2b_1/(2+b_1)$, where $h(b_1)=2b_1/(2-b_1)$. It is noteworthy that $\delta=\pi/2$ at $b_1=b_2=\sqrt{2}$ and that the solutions of Eqs. (9) and (10) are periodic functions with the period $4T/g$.

A pulsed regime in radio-frequency mass spectrometry was considered in Ref. 15 for nonrelativistic ions. For the straight line $b_2=kb_1$ we have $k=f_2/f_1$, i.e., the angle of inclination of the working line is determined only by the parameters of the rectangular pulses. The optimum operating regime is realized in the vicinity of points L and R. When $k=2-\varepsilon$, the straight line cuts off the region ABL from the stability diagram in the vicinity of point L. From the equations $h(b_1)=kb_1$ and $2=kb_1$ we find the coordinates of points A and B: $b_{1A}=2/k\simeq 1+\varepsilon/2$ and $b_{1B}=2(k-1)/k\simeq 1-\varepsilon/2$, which specify the range of permissible values of b_1. The experimental values x_m and y_m of the functions $x(\tau)$ and $y(\tau)$ depend on the values of δ and

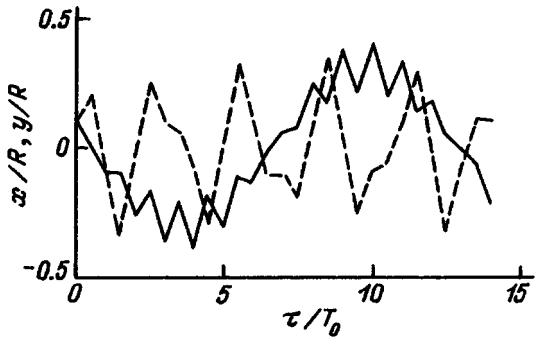

FIG. 2. Plots of $x(\tau)$ (solid line) and $y(\tau)$ (dashed line) for $b_1=1$ and $b_2=1.75$.

δ'. When $b_1=b_{1A}$, we have $\delta\simeq\sqrt{2\varepsilon}$ and $\delta'=\pi$. Therefore, for the same initial conditions, in the case of $b_1\sim b_{1A}$ we obtain $|y_m|\geqslant|x_m|$. On the other boundary of the stability region at $b_1=b_{1B}$ we find $\delta=0$ and $\delta'\simeq\pi+\sqrt{6\varepsilon}$. Here we have $|y_m|\leqslant|x_m|$ for $b_1\sim b_{1B}$. Figure 2 presents plots of $x(\tau)/R$ (solid line) and $y(\tau)/R$ (dashed line) for the midpoint $b_1=1$ when $\varepsilon=0.25$, $x_0=y_0=0.1R$, and $v_1=v_2=0$.

Let the values of b_{1A} and b_{1B} for an assigned frequency of the wave correspond to $\eta_A=4\pi^2V_1/s_1b_{1A}\omega^2R^2$ and $\eta_B=\eta_A+\Delta\eta$. Then all the ions which satisfy the condition $\Delta\eta/\eta_A\leqslant(2-k)/(k-1)\simeq\varepsilon$ enter the collector. All the ions with different values of η must successively enter the stability region in order to separate them according to the parameter $\eta=mg/e$. The intersection of this region along the working straight line can be achieved by varying the frequency at a fixed off-on ratio. We choose the frequency of the wave $\omega=\omega_A$ such that $b_1=b_{1A}$, i.e., $\omega_A^2=4\pi^2V_1/\eta s_1b_{1A}R^2$. Then, as the frequency of the wave is varied in the range $\omega_A\leqslant\omega\leqslant\omega_B$, where $\omega_B^2=4\pi^2V_1/\eta s_1b_{1B}R^2$, the ions with the corresponding value of η intersect the stability region.

This work was supported by the Russian Fund for Fundamental Research.

[1] D. M. Volkov, Zh. Éksp. Teor. Fiz. **7**, 1286 (1937).
[2] R. Redmond, J. Math. Phys. **6**, 1163 (1965).
[3] A. S. Dement'ev and Yu. G. Pavlenko, Zh. Éksp. Teor. Fiz. **62**, 161 (1972) [Sov. Phys. JETP **35**, 86 (1972)].
[4] L. D. Gol'dshteĭn and N. V. Zernov, *Electromagnetic Fields and Waves* [in Russian], Sov. Radio, Moscow (1956), p. 517.
[5] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon, New York (1975).
[6] Yu. G. Pavlenko, [in Russian], Izd. Mosk. Univ., Moscow (1991).
[7] N. W. McLachlan, *Theory and Application of Mathieu Functions*, Clarendon Press, Oxford (1947) [Russ. transl., IL, Moscow (1953)].
[8] M. J. O. Strutt, *Lamesche-, Mathieusche- und Verwandte Funktionen in Physik und Technik*, Springer, Berlin (1932) [Russ. transl., GNTI Ukrainy, Kharkov–Kiev (1935)].
[9] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed., Cambridge University Press, Cambridge (1927) [Russ. transl., Fizmatgiz, Moscow (1963)].
[10] M. A. Miller, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **1**, 110 (1958).
[11] A. V. Gaponov and M. A. Miller, Zh. Éksp. Teor. Fiz. **34**, 242 (1958) [Sov. Phys. JETP **7**, 168 (1958)].
[12] L. D. Landau and E. M. Lifshitz, *Mechanics*, 1st ed., Pergamon Press, Oxford–New York (1960).
[13] S. Solimeno, B. Crosignani, and P. DiPorto, *Guiding, Diffraction, and Confinement of Optical Radiation*, Academic Press, Orlando (1986) [Russ. transl., Mir, Moscow (1989)].
[14] V. M. Galitskiĭ, B. M. Karnakov, and V. I. Kogan, *Problems in Quantum Mechanics* [in Russian], Nauka, Moscow (1973), p. 13.
[15] É. P. Sheretov and V. I. Terent'ev, Zh. Tekh. Fiz. **42**, 953 (1972) [Sov. Phys. Tech. Phys. **17**, 755 (1972)].

Translated by P. Shelnitz

Method for the self-consistent description of the dynamics of electron beams in cyclic systems

N. D. Naumov

(Submitted January 31, 1996)
Zh. Tekh. Fiz. **67**, 103–107 (July 1997)

Equations describing the envelopes of an electron beam in a modified betatron and an $l=2$ stellatron are obtained on the basis of a self-similar solution of the dynamic equations of a charged fluid. It is shown that the poloidal motion of a beam caused by a toroidal magnetic field consists of rotation of the beam as a whole and internal movement of the fluid with elliptical current lines. © *1997 American Institute of Physics.* [S1063-7850(97)02007-7]

INTRODUCTION

The need to develop self-consistent models of the motion of flows of charged particles in external fields stems from practical problems in the formation, acceleration, and transport of beams and electron rings.[1–5] Definite information on the behavior of beams of charged particles is contained in the envelope equations, which are usually obtained on the basis of a solution of the kinetic equation[2,6] or by the method of moments.[7,8] However, the application of these methods to a theoretical analysis of the behavior of electron beams in cyclic systems with a longitudinal magnetic field (for example, in a modified betatron or a stellatron) is formidable, and for this reason the simpler problem of the motion of individual electrons in a combination of external and self-fields has been considered for these systems.[9]

This paper proposes a method for the approximate solution of the self-consistent equations of motion of a thin annular beam in an external field. The method is applicable to low-current beams, where the influence of the self-field on the longitudinal motion of the particles can be neglected. A macroscopic description of the transverse motion of the beam is achieved on the basis of a self-similar solution of the dynamic equations of the charged fluid. This makes it possible to obtain the equations describing the envelopes of an annular beam of elliptical cross section, as well as to establish the relationship between the angular rate of poloidal rotation of the beam as a whole and its internal transverse motion. This relationship is a consequence of the conservation of the sum of the longitudinal components of the vorticity and the magnetic field.

LONGITUDINAL MOTION OF A BEAM

The main difficulty in the approximate solution of the dynamic equations of a charged fluid is associated with the choice of the approximation of the collective effects, since the self-field of the fluid is not known *a priori* and since, on the one hand, it influences the motion of the fluid and, on the other hand, it itself depends on that motion. In this section we consider the conditions under which the influence of the self-field on the longitudinal motion can be neglected.

If the beam current I is much smaller than the limiting Alfvén current I_A, the transverse motion of the beam has a nonrelativistic character, i.e., the ratio of the transverse velocity of each particle to the longitudinal velocity is of the order of ε, where the characteristic small parameter $\varepsilon^2 \sim I/I_A$ (Ref. 7). In this case the momentum of the fluid $\mathbf{P} \approx m\gamma \mathbf{V}$, where γ is the relativistic factor of the longitudinal motion of the particles. Then the equations of motion of a charged fluid[3] in a cylindrical coordinate system (r,Θ,z) take the form

$$\frac{dV_r}{dt} \equiv \left[\frac{\partial}{\partial t} + V_r\frac{\partial}{\partial r} + \frac{V_\Theta}{r}\frac{\partial}{\partial \Theta} + V_z\frac{\partial}{\partial z}\right]V_r = \frac{V_\Theta^2}{r} + \frac{1}{m\gamma}\left[eE_r + \frac{e}{c}(V_\Theta B_z - V_z B_\Theta) - \frac{1}{n}\frac{\partial p}{\partial r}\right], \quad (1)$$

$$\frac{dV_\Theta}{dt} + \frac{1}{r}V_r V_\Theta = \frac{1}{m\gamma}\left[eE_\Theta + \frac{e}{c}(V_z B_r - V_r B_z)\right], \quad (2)$$

$$\frac{dV_z}{dt} = \frac{1}{m\gamma}\left[eE_z + \frac{e}{c}(V_r B_\Theta - V_\Theta B_r) - \frac{1}{n}\frac{\partial p}{\partial z}\right], \quad (3)$$

where $\mathbf{E}$ and $\mathbf{B}$ are the intensities of the sum of the external field and the self-field, and p is the pressure caused by the emittance of the beam.

Setting for a linear beam

$$E_z = -\frac{1}{c}\frac{\partial A_{ez}}{\partial t}, \quad B_r = \frac{1}{r}\frac{\partial A_z}{\partial \Theta}, \quad B_\Theta = -\frac{\partial A_z}{\partial r}, \frac{\partial p}{\partial z} = 0,$$

for the function $\Pi = (V_z + eA_z/\gamma mc)/u$ from Eq. (3) we obtain

$$\frac{d\Pi}{dt} = 0, \quad (4)$$

i.e., the function Π is conserved as the beam moves. Here u is the characteristic value of the longitudinal velocity of the particles, and $A_z = A_{ez} + A_{sz}$, where A_{ez} and A_{sz} are the components of the vector potentials of the external field and the self-field, respectively.

If the function Π is initially identical at all points, it remains everywhere identical and invariant with time as the fluid moves. In this case, for V_z we find

$$V_z = u - \frac{e}{\gamma mc}A_z. \quad (5)$$

Thus, in the present case the longitudinal velocity depends on the coordinates. Therefore, the model of a homogeneous distribution of the longitudinal velocity over the

beam cross section is applicable only under certain assumptions. For example, for a uniform cylindrical beam $eA_{sz}/\gamma mcu = -Ir^2/a^2 I_A$ at $r \leqslant a$, where a is the radius of the beam. It then follows from Eq. (5) that the uniform-beam approximation and, accordingly, the neglect of the influence of the self-field on the longitudinal motion of the particles is justified only to first-order terms with respect to ε (provided $eA_{ez}/\gamma mcu \sim \varepsilon$). These estimates are consistent with the exact stationary solution of the self-consistent equations for a beam with a sharp boundary.[10]

In the case of an annular beam

$$E_\Theta = -\frac{1}{c}\frac{\partial A_\Theta}{\partial t}, \quad B_r = -\frac{\partial A_\Theta}{\partial z}, \quad B_z = \frac{1}{r}\frac{\partial r A_\Theta}{\partial r}.$$

Then for the function

$$\Pi = \left(V_\Theta + \frac{e}{\gamma mc} A_\Theta \right) \frac{r}{uR}, \tag{6}$$

where u and R are the characteristic values of the rotation rate of the particles and the radius of the ring, Eq. (2) leads to conservation law (4) and thus to an expression analogous to (5) for the azimuthal component of the hydrodynamic velocity

$$V_\Theta = \frac{C}{r} - \frac{e}{\gamma mc} A_\Theta, \tag{7}$$

where the constant C is determined from the initial conditions.

Let us consider an annular ring of elliptical cross section

$$n(x,t) = \frac{N}{2\pi^2 abr} H\left(1 - \frac{q_1^2}{d^2} - \frac{q_2^2}{b^2} \right), \tag{8}$$

where N is the number of particles in the ring, a and b are the semiaxes of its cross section, $H(\tau)$ is a Heaviside step function, and q_i is the coordinate system associated with the beam axis and the symmetry axes of the transverse section of the beam.

For a thin beam the scalar potential and the azimuthal component of the vector potential of the self-field have the form[5]

$$\Phi = 2\frac{I}{u}\left[L - \Phi_0 - \frac{Lq_1}{\pi R} + \frac{L}{16R^2}(5q_1^2 - q_2^2) \right],$$

$$A_{s\Theta} = \beta\left[\Phi + \frac{IL}{2uR^2}(q_1^2 + q_2^2) \right].$$

Here $L = \ln[16R/(a+b)]$, $\beta = u/c$, and Φ_0 characterizes the potential of an infinitely long charged cylinder of elliptical cross section with a constant charge density. Hence it is seen that, as in the case of a linear beam, the contribution of the self-field to the azimuthal motion for a low-current annular ring is quadratic with respect to the small parameter. Neglecting the terms of order ε^2, we find that the distribution of the longitudinal velocity of the fluid in the ring is inhomogeneous over the cross section

$$V_\Theta \simeq \frac{C}{r} - \frac{e}{\gamma mc} A_{e\Theta}. \tag{9}$$

Thus, for low-current beams the influence of the self-field on the longitudinal motion of the particles can be neglected in the linear approximation with respect to the small parameter. We also note that Eqs. (5) and (7) can be regarded as a consequence of the choice of a distribution function whose dependence on the integral of motion, i.e., the longitudinal component of the generalized momentum of the particle, is assigned in the form of a delta function within the kinetic description. The averaging of that component of the generalized momentum on passage to the hydrodynamic description leads to the results (5) and (7) obtained above for the longitudinal velocity.

TRANSVERSE DYNAMICS OF AN ELECTRON RING

As a concrete problem, let us consider the motion of an annular beam in a modified betatron, in which along with the usual betatron field having the potential

$$A_{e\Theta} = \frac{1}{r}\int_0^r B_\beta(x) x\, dx - \frac{z^2}{2}\frac{\partial B_\beta}{\partial r},$$

there is a toroidal magnetic field $B_{e\Theta} = B_2 R/r$. Here $B_\beta(r) = B_1 (R/r)^n$; and n, R, B_1, and B_2 are constants. The characteristic value of the longitudinal velocity of the particles in the ring is clearly proportional to its radius. For the class of self-similar motions under consideration, the transverse velocity components are proportional to the distance to the beam axis.[11] Therefore, here the small value of ε signifies a small ratio of the transverse dimensions of the beam to its radius. Approximate solutions of the self-similar type can be constructed for a thin beam by expanding the terms with the azimuthal velocity in Eqs. (1) and (3)[12] and restricting ourselves to the terms that are linear with respect to the small parameter. Thus, the external magnetic field for performing such a procedure should be weakly nonuniform, while the components of the external electric field can be linearly dependent on r and z.

It is convenient to consider the transverse motion of a beam of elliptical cross section in the coordinate system q_i associated with the symmetry axes of the cross section. We go over to this coordinate system in two stages. First, in Eqs. (1) and (3) we make the replacements

$$r = s + \rho, \quad z = \sigma + \zeta, \quad V_r = \dot{s} + V_\rho, \quad V_z = \dot{\sigma} + V_\zeta,$$

where a dot denotes differentiation with respect to time; the functions of time, s and σ characterize the position of the beam axis; and V_ρ and V_ζ are the components of the velocity of the fluid relative to the coordinate system that is associated with the beam axis and has the same unit vectors as the original coordinate system.

Since the axial oscillations of the beam take place near the $z = 0$ plane, $z/s \sim \varepsilon$. Then, under the condition that the azimuthal velocity of the fluid on the beam axis is initially equal to u, we find the following expression for V_Θ from (7) to within the accuracy adopted:

$$V_\Theta(r,t) = U\left(1 - \frac{\rho}{s} \right) - \omega_1 \rho \left(\frac{R}{s} \right)^n.$$

The following notation is used here and in the following: $\omega_i = eB_i/\gamma mc$, $U = us_0/s + \omega_1[(s_0^2/s)(R/s_0)^n - s(R/s)^n]/(2-n)$. After linearizing Eqs. (1) and (3), we obtain the following equations for the coordinates of the beam axis:

$$\ddot{s} + \omega_2 \sigma \frac{\dot{R}}{s} - \omega_1 U\left(\frac{R}{s}\right)^n - \frac{U^2}{s} = hL\left(\frac{1+\beta^2}{2s} - \frac{\ddot{s}}{c^2}\right), \quad (10)$$

$$\ddot{\sigma} - \omega_2 \dot{s}\frac{R}{s} - n\omega_1 U\frac{\sigma}{s}\left(\frac{R}{s}\right)^n = \frac{hL}{c^2}\ddot{\sigma}, \quad (11)$$

where $h = 2c^2 I/I_A$.

The first term on the right-hand side of Eq. (10) is due to elongation of the ring along the larger radius, and the second term, like the first term on the right-hand side of Eq. (11), reflects the increase in the electron effective mass in the ring.[6] If these terms are omitted, Eqs. (10) and (11) become the equations of motion of a single charged particle near the $z = 0$ plane.

The terms which depend on the coordinates ρ and ζ define the equations for the velocity components

$$\left(\frac{\partial}{\partial t} + V_\rho \frac{\partial}{\partial \rho} + V_\zeta \frac{\partial}{\partial \zeta}\right) V_{\rho,\zeta} = F_{\rho,\zeta} + G_{\rho,\zeta},$$

$$G_\rho = \rho\Gamma_1 - \omega_2 V_\zeta \frac{R}{s}, \quad G_\zeta = \omega_2 V_\rho \frac{R}{s} - 2\rho\Lambda + \zeta\Gamma_2, \quad (12)$$

where F_ρ and F_ζ are determined by the emittance and the self-field of the beam,

$$\Gamma_1 = \omega_2 \dot{\sigma} R/s^2 - 3U^2/s^2 - \omega_1 (R/s)^n \times [(3+n)U/s + \omega_1 (R/s)^n],$$

$$\Gamma_2 = n\omega_1 (U/s)(R/s)^n, \quad \Lambda = \omega_2 \dot{s} R/2s^2.$$

Since the orientation of the beam cross section can vary as a consequence of the poloidal motion of the fluid, the introduction of the coordinates q_i associated with the symmetry axes of the cross section corresponds to passage to a coordinate system, whose unit vectors are turned through a certain angle Ψ relative to the unit vectors of the original coordinate system:

$$q_1 = \rho \cos\Psi + \zeta \sin\Psi, \quad q_2 = \zeta \cos\Psi - \rho \sin\Psi.$$

Accordingly, V_ρ and V_ζ should be represented in terms of the velocity components V_i of the fluid in the new coordinate system

$$V_\rho = V_1 \cos\Psi - V_2 \sin\Psi - \zeta\Omega,$$

$$V_\zeta = V_2 \cos\Psi + V_1 \sin\Psi + \rho\Omega,$$

where $\Omega = \dot{\psi}$ is the angular velocity of the poloidal rotation of the beam as a whole. Then from Eqs. (12) for the V_i we obtain

$$\left(\frac{\partial}{\partial t} + V_i \frac{\partial}{\partial q_i}\right) V_1 - 2\Omega V_2 - q_1\Omega^2 - q_2\dot{\Omega} = F_1 + G_\rho \cos\Psi + G_\zeta \sin\Psi, \quad (13)$$

$$\left(\frac{\partial}{\partial t} + V_i \frac{\partial}{\partial q_i}\right) V_2 + 2\Omega V_1 - q_2\Omega^2 - q_1\dot{\Omega} = F_2 + G_\zeta \cos\Psi - G_\rho \sin\Psi. \quad (14)$$

In the new coordinate system q_1/a and q_2/b are self-similar variables. The transverse motion of the fluid in this coordinate system also includes movement with elliptical current lines; therefore, the expressions for the V_i have the form

$$V_1 = \dot{a}\frac{q_1}{a} - \omega a \frac{q_2}{b}, \quad V_2 = \dot{b}\frac{q_2}{b} + \omega b \frac{q_1}{a}, \quad (15)$$

where ω is a certain function of time. We note that in this case the particle density (8) satisfies the continuity equation. Plugging the expressions (15) into Eqs. (13) and (14), we ultimately obtain a system of equations for the functions of time a, b, ω, and Ω introduced:

$$\ddot{a} - (\omega^2 + \Omega^2 + \Gamma_1 \cos^2\psi - \Lambda \sin 2\psi + \Gamma_2 \sin^2\psi)a = 2\omega\Omega b - \omega_2(\omega b - \Omega a) + \frac{2h}{(a+b)\gamma^2} + \frac{H}{a^3} - \frac{hL}{s^2}, \quad (16)$$

$$\ddot{b} - (\omega^2 + \Omega^2 + \Gamma_2 \cos^2\psi + \Lambda \sin 2\psi + \Gamma_1 \sin^2\psi)b = 2\omega\Omega a - \omega_2(\omega a + \Omega b) + \frac{2h}{(a+b)\gamma^2} + \frac{H}{b^3} + \frac{hL}{2s^2}, \quad (17)$$

$$\dot{\omega}(a^2 - b^2) + 2\omega(a\dot{a} - b\dot{b}) = \left(\omega_2 \frac{R}{S} - 2\Omega\right)(a\dot{b} - b\dot{a}) + ab[2\Lambda \cos 2\psi + (\Gamma_1 - \Gamma_2)\sin 2\psi], \quad (18)$$

$$\dot{\Omega} + \left(\Omega - \frac{R\omega_2}{2s}\right)\left(\frac{\dot{a}}{a} + \frac{\dot{b}}{b}\right) = \frac{1}{2ab}\frac{d}{dt}\left[\omega_2 \frac{R}{s} - \omega(a^2 + b^2)\right],$$

where $H = (uE/\pi)^2$, E is the emittance of the beam, and the terms containing L are due to the curvature of the beam.

The equation for Ω can be integrated in a simple manner:

$$\Omega = \frac{R\omega_2}{2s} + \left[\omega_0(a_0^2 + b_0^2) + 2a_0 b_0\left(\Omega_0 - \frac{R\omega_2}{2s_0}\right) - \omega(a^2 + b^2)\right]\frac{1}{2ab}. \quad (19)$$

As we have already noted, in the framework of the self-similar approach the influence of the longitudinal magnetic field on the rotation of the beam as a whole and the internal motion of the fluid can be determined from a general relation, which is a consequence of the dynamic equations of a charged fluid:

$$\frac{d}{dt}\frac{\mathbf{R}\nabla S}{n}=0.$$

Here $\mathbf{R}=\mathbf{W}+e\mathbf{B}/\gamma mc$; $\mathbf{W}=\mathrm{curl}\mathbf{V}$ is the vorticity; n is the density of the fluid; and S is an arbitrary function, which is conserved as the fluid moves.

In particular, the function for an annular beam is

$$S=\delta(r-s)\,\delta\left(\Theta-\int_0^t V_\Theta(s,\tau)\frac{d\tau}{s(\tau)}\right)\delta(z-\sigma).$$

Then, from the condition of the conservation of $\mathbf{R}\nabla S/n$ during motion of the fluid we obtain Eq. (19).

Equations (16) and (17) simplify in the case of an ordinary betatron. As can easily be seen, if $\omega_2=0$ and there is initially no poloidal motion of the fluid, such motion does not appear later on. Omitting the terms caused by the curvature of the beam for simplicity, in this case we have

$$\ddot{a}+\varkappa a=\frac{2h}{(a+b)\gamma^2}+\frac{H}{a^3},\quad \ddot{b}-\Gamma_2 b=\frac{2h}{(a+b)\gamma^2}+\frac{H}{b^3},\quad (20)$$

where $\varkappa=3U^2/s^2+\omega_1(R/s)^n[(3+n)U/s+\omega_1(R/s)^n]$.

A further simplification is possible under the condition that the radial oscillations of the beam take place near the axis of the external field, i.e., $(s-R)/R\sim\varepsilon$, and if the particles initially rotate on the axis of the field with the cyclotron frequency, i.e., if $u=-\omega_1 R$. Then, setting s_0, $s\simeq R$ in (20) in accordance with the accuracy adopted, we find

$$\ddot{a}+\omega_1^2(1-n)a=\frac{2h}{(a+b)\gamma^2}+\frac{H}{a^3},$$

$$\ddot{b}+n\omega_1^2 b=\frac{2h}{(a+b)\gamma^2}+\frac{H}{b^3}.$$

RIPPLED ANNULAR BEAM

If the system under consideration does not have axial symmetry, relation (4) does not hold. In this case, including the axisymmetric part of the external field in function (6) we obtain

$$\frac{d\Pi}{dt}=e(\gamma mcuR)^{-1}(V_z B_r'-V_r B_z'),\quad (21)$$

where the terms of the field that depend on the azimuthal angle are marked with a prime.

Hence it is seen that the expression (9) for the azimuthal velocity can also be applied to systems that do not have axial symmetry, if the right-hand side of Eq. (21) is proportional to a small parameter raised to the second power.

For an $l=2$ stellatron, whose magnetic field can be represented in the form[12]

$$B_{er}'=\frac{B_3}{R}(\xi\sin k\Theta-z\cos k\Theta),$$

$$B_{e\Theta}=B_2\frac{R}{r},\quad B_{ez}=B_1+B_{ez}',$$

where $\xi=r-R$, $B_{ez}'=-B_3(\xi\cos k\Theta+z\sin k\Theta)/R$, and k is an integer, the contribution of the part of the external field which depends on the azimuthal angle to the right-hand side of (12) will clearly be of the order of ε^2, if the motion of the beam near the axis of the field is considered, i.e., if $(s-R)/R,\ \sigma/R\sim\varepsilon$.

In this case a model of a rippled annular beam can be constructed for an electron beam in an $l=2$ stellatron. As in the case of linear beams in periodic focusing systems,[2,4] neglect of the influence of the self-field on the longitudinal motion of the particles is possible, if the beam current is small in comparison with the Alfvén current and significant variation of the transverse dimensions occur at distances considerably greater than the lengths of the semiaxes. Then from (9) we find

$$V_\Theta=U\left(1-\frac{\rho}{s}\right)-\omega_1\rho,$$

where $U=us_0/s+\omega_1(s_0^2/s-s)/2$.

If we introduce the variable $\eta=\Theta-ut/R$ instead of Θ, then, to within the accuracy adopted the dynamic equations do not contain terms with a derivative with respect to η, i.e., the dependence of the characteristics of the transverse motion of the beam on this parameter has a parametric character. This dependence is defined for each cross section of the beam by the initial conditions. Following the scheme described above, for the coordinates of the beam axis we ultimately find equations analogous to (10) and (11)

$$\ddot{s}+\omega_2\dot{\sigma}\frac{R}{s}+\omega_2\frac{u}{R}[(s-R)\cos\varphi+\sigma\sin\varphi]-\frac{U^2}{s}=0,$$

$$\ddot{\sigma}-\omega_2\dot{\sigma}\frac{R}{s}+\omega_2\frac{u}{R}[(s-R)\sin\varphi-\sigma\cos\varphi]=0,$$

where $\varphi=k(\eta+ut/R)$ and the terms due to the self-field of the beam have been omitted for brevity.

The angular rate of poloidal rotation of the beam as a whole is specified by Eq. (19), and the variation of the characteristics of the transverse motion of the beam is given by Eqs. (16)–(18), where

$$\Gamma_1=\omega_2\frac{\dot{\sigma}}{R}-\omega_1^2-3(u+\omega_1R)\frac{u}{R^2}-\Gamma_2,$$

$$\Lambda=\frac{\omega_2\dot{s}}{2R}+\omega_3\frac{u}{R}\sin\varphi,\quad \Gamma_2=\omega_3\frac{u}{R}\cos\varphi.$$

Thus, the proposed method makes it possible to obtain a system of ordinary differential equations that describe the transverse oscillations of an annular beam of charged particles in an external field. The solution of this system of equations is a separate problem.

This work was performed with support from the Russian Fund for Fundamental Research.

[1] A. A. Kolomensky and A. N. Lebedev, *Theory of Cyclic Accelerators*, North-Holland, Amsterdam (1966).
[2] I. M. Kapchinskiĭ, *Dynamics of Particles in Resonant Linear Accelerators* [in Russian], Atomizdat, Moscow (1966).
[3] R. C. Davidson, *Theory of Nonneutral Plasmas*, Benjamin, New York (1974) [Russ. transl., Mir, Moscow (1978)].

[4] S. I. Moloḳovskiĭ and A. D. Sushkov, *Intense Electron and Ion Beams* [in Russian], Énergoatomizdat, Moscow (1991).
[5] V. P. Sarantsev and É. A. Perel'shteĭn, *Collective Acceleration of Ions by Electron Rings* [in Russian], Atomizdat, Moscow (1979).
[6] O. I. Yarkovoĭ, Zh. Tekh. Fiz. **36**, 988 (1966) [Sov. Phys. Tech. Phys. **11**, 731 (1966)].
[7] E. P. Lee and R. K. Cooper, Part. Accel. **7**, 83 (1976).
[8] N. Yu. Kazarinov and É. A. Perel'shteĭn, Zh. Tekh. Fiz. **50**, 101 (1980) [Sov. Phys. Tech. Phys. **25**, 58 (1980)].
[9] V. V. Dolgopolov, Yu. V. Kirichenko, Ya. F. Leleko *et al.*, Zh. Tekh. Fiz. **65** (6), 141 (1995) [Tech. Phys. **40**, 599 (1995)].
[10] D. A. Hammer and N. Rostoker, Phys. Fluids **13**, 1831 (1970).
[11] N. D. Naumov, Fiz. Plazmy **19**, 1406 (1993) [Plasma Phys. Rep. **19**, 738 (1993)].
[12] C. W. Robertson, A. Mondelli, and D. Chernin, Phys. Rev. Lett. **50**, 507 (1983).

Translated by P. Shelnitz

Determination of the final equilibrium radius and the increase in the emittance of a nonequilibrium relativistic electron beam during transport along an external magnetic field in a gas-plasma scattering medium

E. K. Kolesnikov and A. S. Manuĭlov

St. Petersburg State University; V. I. Smirnov Scientific-Research Institute of Mathematics and Mechanics, 198904 St. Petersburg, Russia
(Submitted December 26, 1995)
Pis'ma Zh. Tekh. Fiz. **67**, 108–111 (July 1997)

An equation describing the evolution of the transverse energy of a segment of a paraxial axisymmetric relativistic electron beam (REB) propagating in a gas-plasma scattering medium along an external magnetic field is used to find the equation relating the final equilibrium radius of the beam to its initial nonequilibrium value. An analytical expression for the increase in the mean-square emittance of an REB during transport up to achievement of the equilibrium state is found for the case considered. The dependence of the final equilibrium radius and the corresponding increase in the mean-square emittance on the density of the scattering medium and the induction of the external magnetic field is investigated. © *1997 American Institute of Physics.* [S1063-7842(97)01907-7]

In the recent period researchers have shown increasing interest in the dynamics of the transport of relativistic electron beams (REBs) in gas-plasma media in a self-pinching regime.[1–10] An initial deviation of the beam parameters from their equilibrium values has frequently been observed in experimental and numerical investigations of the propagation of REBs in such media. In particular, a small deviation of the radius of an REB from the equilibrium condition on injection can lead to the development of resistive sausage instability of the beam, which is characterized by rapidly growing radial pulsations.[2,4,6,10] However, as numerical simulations and the analytical theory have shown, the sausage mode is suppressed over a broad range of parameters of the beam and the background gas-plasma medium, and the beam achieves a certain equilibrium state.[1,6,10] For this reason, there would be definite interest in theoretical predictions of the final equilibrium radius of an REB, as well as the increase in the mean-square emittance of such a beam during its evolution toward the final equilibrium state. This problem was previously treated in Refs. 4 and 5. However, the influence of the external longitudinal magnetic field, as well as the repeated Coulomb scattering of beam particles on background gas atoms (which must often be taken into account in practice), on the final equilibrium radius of the beam and the increase in its emittance was not studied in Refs. 4 and 5.

The present work is devoted to the derivation of generalized equations for determining the final equilibrium radius and the corresponding increase in the mean-square emittance with allowance for the influence of the external magnetic field and the scattering in the background gas-plasma medium.

Let us consider a monovelocity, azimuthally symmetric paraxial REB propagating along a constant external magnetic field with the induction B_0 in a gas-plasma scattering medium. We choose the cylindrical coordinate system (r,Θ,z) such that the vector B_0 is parallel to the z axis. It is assumed that partial charge neutralization takes place with a neutralization fraction f_c and that partial magnetic (current) neutralization occurs with a factor f_m. In addition, for simplicity, we assume that there are no energy losses from the beam particles over the REB transport distances considered and that the neutralization fractions f_c and f_m, the total beam current I_b, and the induction of the external magnetic field do not vary time, i.e.,

$$\frac{d\gamma}{dt}=0, \quad \frac{d(\varkappa T_B)}{dt}=0, \quad \frac{d\Omega_c}{dt}=0, \tag{1}$$

where $\varkappa=(1-f_m)-(1-f_c)/\beta^2$; $\beta=v_z/c$ (v_z is the longitudinal velocity of the electron beam, and c is the velocity of light); $T_B=e\beta I_b/(2c)$ is the effective Bennett temperature, which characterizes the self-pinching of the beam by its own electromagnetic field; $\Omega_c=eB_0/(\gamma mc)$ is the cyclotron frequency of the beam particles in the external magnetic field B_0; and m and e are the electron mass and charge.

We shall henceforth assume that the doubled root-mean-square beam radius R_i on injection differs from the equilibrium value R_0 by the small amount δR_0

$$R_i=R_0+\delta R_0, \tag{2}$$

where $|\delta R_0|/R_0 \ll 1$.

We note that, if conditions (1) are satisfied, the integral of motion of the transverse beam segment characterized by the injection time τ for the beams under consideration is the transverse energy Ψ of the beam segment, which is defined in the following manner:[1]

$$\Psi=\varepsilon_\perp+A_\beta, \tag{3}$$

where

$$\varepsilon_\perp=\int d\mathbf{r}_\perp \chi \frac{\widetilde{p^2}}{2m\gamma}, \tag{4}$$

$$A_\beta=-\frac{1}{2}\int d\mathbf{r}_\perp \chi e\beta \frac{\varkappa}{(1-f_m)}A_z \tag{5}$$

 1063-7842/97/070819-04$10.00

are, respectively, the mean transverse kinetic energy and mean potential energy of the particles in the segment in the collective electromagnetic field of the plasma-beam system. Here $\chi = J_b(r)/I_b$, where $J_b(r)$ is the radial current density profile of the beam, I_b is the total current of the REB, $\widetilde{p^2}$ is the mean-square momentum of the beam particles in the segment, and A_z is the z component of the vector potential of the collective electromagnetic field.

Generalizing Lee's kinetic theory[1] for paraxial REBs to the case of the influence of the external longitudinal magnetic field, we can obtain the equation for the evolution of the transverse energy of a beam segment in the form

$$\frac{d\Psi}{dt} = -\frac{\varepsilon_\perp}{\gamma}\frac{d\gamma}{dt} + A_\beta \frac{d\ \ln(\varkappa T_B)}{dt} - \frac{1}{2\gamma}\frac{d(\Omega_c \gamma)}{dt} m\gamma L + \int d\mathbf{r}_\perp \chi S, \tag{6}$$

where S characterizes the rate of the conversion of energy into transverse motion of the beam particles as a result of multiple Coulomb scattering of the REB electrons on atoms of the background gas-plasma medium, and L is the mean angular momentum of the particles in the segment under consideration.

When simplifying assumptions (1) hold, Eq. (6) reduces to the simple form

$$\Psi_f - \Psi_i = \gamma m(\beta c)^3 \sigma_s n_g t_f, \tag{7}$$

where Ψ_f and Ψ_i are, respectively, the values of the transverse energy of the beam segment in the final equilibrium state and on injection of the REB, σ_s is the transport cross section for small-angle multiple Coulomb scattering of the beam particles on atoms of the background gas, n_g is the density of the atoms of the background medium, and t_f is the time of achievement of the equilibrium state.

It is not difficult to see that in the absence of scattering ($\sigma_s = 0$) Ψ is the integral of motion for the beam segment under consideration. This situation was considered for $B_0 = 0$ in Refs. 4 and 5. We next introduce the parameter

$$\Gamma = \frac{A_\beta}{\varkappa T_B} + \ln\left(\frac{\sqrt{2}R_c}{R}\right)^2, \tag{8}$$

where R_c is the radius for which $A_z|_{r=R_c} = 0$ It is not difficult to show that Γ can be represented in the following form:

$$\Gamma = -2\int_0^{R_c} \frac{dr}{r}\left(\frac{I_b(r)}{I_b}\right)^2 + \ln\left(\frac{\sqrt{2}R_c}{R}\right)^2, \tag{9}$$

where $I_b(r) = \int_0^r dr' 2\pi r' J_b(r')$ is the beam current in a tube of radius r.

The parameter Γ is a constant for different forms of the radial current density profile $J_b(r)$ of the beam. In particular, $\Gamma = 0.193$ for a uniform profile, $\Gamma = 0.116$ for a Gaussian profile, and $\Gamma = 0.195$ for a Bennett profile cut off at $r = 3R_b$. Thus, Γ varies very slightly as the radial current density profile of the REB evolves. In addition, it can be shown that $d\Gamma/dt$ characterizes the influence of the phase mixing of the beam particle trajectories on the variation of the mean-square emittance of the beam.[11]

Let us turn to the derivation of the equation which relates the final equilibrium doubled root-mean-square beam radius R_f to the initial doubled root-mean-square radius R_i, which is assumed to be known. For this purpose we use the equation of the beam envelope for the doubled root-mean-square radius R. In the case under consideration we have[2]

$$\frac{d^2R}{dt^2} + \frac{dR}{dt}\frac{1}{\gamma}\frac{d\gamma}{dt} + \frac{4\varkappa T_B}{\gamma m R} + \frac{\Omega_c^2}{4}R = \frac{4(E^2 + P_\Theta^2/m^2)}{\gamma^2 R^3}, \tag{10}$$

where

$$P_\Theta = m\gamma\left(L + \frac{\Omega_c R^2}{4}\right) \tag{11}$$

is the mean canonical angular momentum of the particles in the beam segment, and

$$E^2 = \frac{\gamma^2 R^2}{4}\left[\frac{4\varepsilon_\perp}{m\gamma} - \left(\frac{dR}{dt}\right)^2 - \left(\frac{2L}{R}\right)^2\right] \tag{12}$$

is the mean-square emittance of the REB.

In addition, from Eqs. (3), (7), and (8) we obtain the following equation:

$$\varepsilon_{\perp f} + \varkappa T_B\left[\Gamma_f + \ln\left(\frac{R_f}{\sqrt{2}R_c}\right)^2\right] - \varepsilon_{\perp i} - \varkappa T_B\left[\Gamma_i + \ln\left(\frac{R_i}{\sqrt{2}R_c}\right)^2\right] = t_f \gamma m(\beta c)^3 \sigma_s n_g, \tag{13}$$

where the subscript f refers to the final values of the parameters, and i refers to the initial values.

Since the final state is an equilibrium state, we have

$$\left(\frac{dR}{dt}\right)_f = \left(\frac{d^2R}{dt^2}\right)_f = 0. \tag{14}$$

Using (10)–(12) and (14) we can easily obtain the condition for dynamic equilibrium in the final stage:

$$\frac{(\varepsilon_\perp)_f}{m\gamma} = -\frac{L_f \Omega_c}{2} + \frac{\varkappa T_B}{m\gamma}. \tag{15}$$

It can be shown under the condition $d\gamma/dt = 0$ that P_Θ is the integral of motion of the beam segment. We then have

$$L_f = L_i + \frac{\Omega_c}{4}(R_i^2 - R_f^2). \tag{16}$$

Substituting (16) into (15), we obtain

$$\frac{(\varepsilon_\perp)_f}{m\gamma} = -\frac{L_i \Omega_c}{2} - \frac{\Omega_2^2}{8}R_i^2\left(1 + \frac{R_f^2}{R_i^2}\right) + \frac{\varkappa T_B}{m\gamma}. \tag{17}$$

Using (13) and (17), we can write the following equation

$$(1 - \Gamma_f - \Gamma_i) + \ln\left(\frac{R_f}{R_i}\right)^2 - \frac{L_i k_c}{(\beta c)U} - \frac{k_c^2}{4U}R_i^2\left[1 - \left(\frac{R_f}{R_i}\right)^2\right] = \frac{2\varepsilon_{\perp i}}{m\gamma(\beta c)^2 U} + \frac{2L_r \sigma_s n_g}{U}, \tag{18}$$

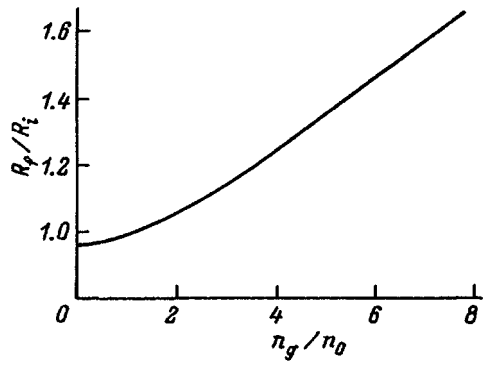

FIG. 1.

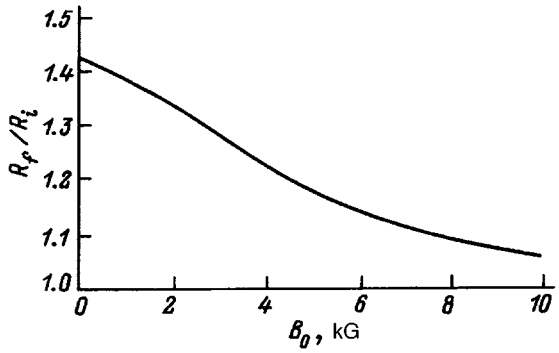

FIG. 3.

where $U=\varkappa I_b/I_A$ is the generalized perveance of the beam, $I_A=\beta\gamma mc^3/e$ is the limiting Alfvén current, $k_c=\Omega_c/\beta c$ is the wave number of the cyclotron oscillations of the beam particles in a longitudinal external magnetic field with an induction B_0.

In the absence of a scattering process ($\sigma_s=0$) and an external magnetic field ($k_c=0$), Eq. (18) leads to the following equation, which was obtained in Ref. 4, for an initially cold beam ($\varepsilon_{\perp i}=0$):

$$\ln\left(\frac{R_i}{R_f}\right)^2=1+(\Gamma_f-\Gamma_0). \tag{19}$$

Since Γ varies weakly as a function of the form of the radial current density profile of the beam ($\Gamma_f-\Gamma_0\ll 1$), we have

$$R_i\simeq 1.649R_f, \tag{20}$$

$$E_f^2\simeq UR_f^2=0.368UR_0^2. \tag{21}$$

We next express the value of $\varepsilon_{\perp i}$ in Eq. (18) in terms of R_i and R_0, where R_0 is the equilibrium radius corresponding to the initial emittance of the beam. When $(dR/dt)_i=0$ and $L_i=L_0$, from (11)–(13) we obtain

$$\frac{2\varepsilon_\perp}{m\gamma(\beta c)^2U}=\frac{R_0^2}{R_i^2}-\frac{2L_0k_c}{\beta_cU}\frac{R_0^2}{R_i^2}. \tag{22}$$

Then, substituting this expression into Eq. (18), we ultimately obtain

$$(1+\Gamma_f-\Gamma_i)+\ln\left(\frac{R_f}{R_i}\right)^2+\frac{k_c^2}{4U}R_i^2\left[\left(\frac{R_f}{R_i}\right)^2-1\right]$$
$$=\frac{R_0^2}{R_i^2}+\frac{L_0k_c}{\beta cU}\left[1-2\left(\frac{R_0}{R_i}\right)^2\right]+\frac{2L_r\sigma_sn_g}{U}. \tag{23}$$

We next determine the change in the mean-square emittance of the beam during its transport to the final equilibrium state:

$$\Delta E^2=E_f^2-E_i^2, \tag{24}$$

where E_i^2 and E_f^2 are, respectively, the initial and final values of the mean-square emittance of the beam segment.

In accordance with Eqs. (11), (12), and (14), we have

$$\frac{\Delta E^2}{(\gamma\beta c)^2}=\frac{U}{2}(R_f^2-R_0^2)+\frac{k_c^2}{16}(R_f^4-R_0^4)-\left(\frac{L_i}{\beta c}+\frac{k_cR_i^2}{4}\right)^2$$
$$+\left(\frac{L_0}{\beta c}+\frac{k_cR_0^2}{4}\right)^2, \tag{25}$$

where R_0 is the equilibrium root-mean-square beam radius corresponding to the initial emittance E_i.

Let us use the equations (23) and (25) derived to examine several simplified situations. We first consider the case when $k_c=0$ and $\sigma_s=0$ ($B_0=0$, and there is no scattering). Then, using (2), from (23) we have

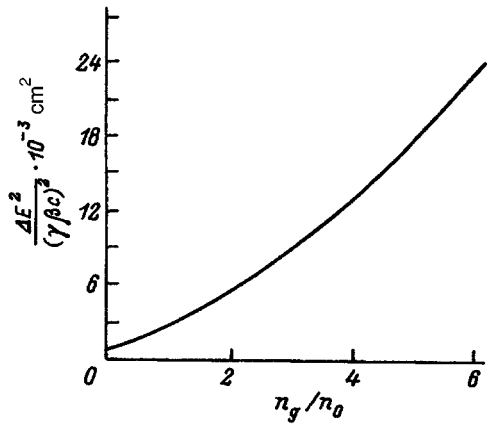

FIG. 2.

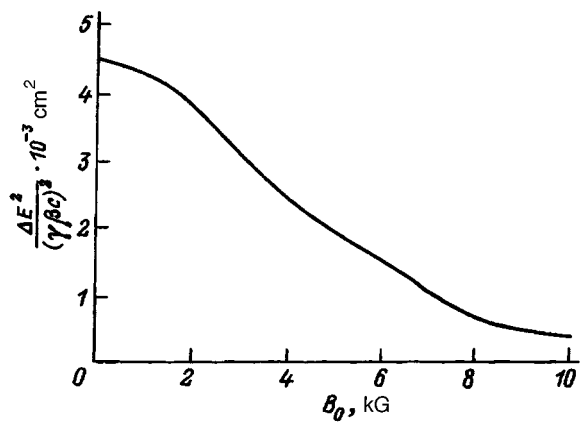

FIG. 4.

$$\left(\frac{R_f}{R_0}\right)^2 \simeq 1+\frac{1}{2}\frac{R_i^2}{R_0^2}\left(\frac{R_0^2}{R_i^2}-1\right)^2, \tag{26}$$

which, under the condition $|\delta R_0|/R_0 \ll 1$, reduces to the expression

$$\left(\frac{R_f}{R_0}\right)^2 \simeq 1+\frac{2(\delta R_0)^2}{R_0^2}. \tag{27}$$

In this case from (25) we can easily obtain the result in Ref. 5

$$\Delta E^2=(\gamma\beta c)^2 U(\delta R_0)^2. \tag{28}$$

In the absence of an external magnetic field ($k_c=0$) and the presence of a scattering process, with consideration of the weak variation of the parameter as the radial current profile of the beam evolves ($|\Gamma_f-\Gamma_0| \ll 1$) we have

$$\left(\frac{R_f}{R_0}\right)^2 \simeq \exp(\Psi)\left[1+2\left(\frac{\delta R_0}{R_0}\right)^2\right], \tag{29}$$

where $\Psi=2L_r\sigma_s n_g/U$.

The increase in the mean-square emittance in this case has the form

$$\Delta E^2=(\gamma\beta c)^2\frac{U}{2}[R_0^2(\exp(\Psi)-1)+2\exp(\Psi)(\delta R_o)^2]. \tag{30}$$

As an example, Figs. 1 and 2 present plots of the dependence of R_f/R_i and $\Delta E^2/(\gamma\beta c)^2$ on the density n_g of the background gas atoms ($n_0)\cdot 10^{19}$ cm^{-3}) for the case of an REB with an electron energy $E=5$ MeV ($\gamma=10$), $I_b=10$ kA, $\varkappa=0.5$, $R_0=1.3$ cm, $R_i=1$ cm, and $L_r \approx 3L_\beta \approx \sqrt{U}/R_b$. It is not difficult to see that the scattering process, as expected, can have a significant influence on the final equilibrium radius R_f, as well as on the increase in the emittance of the beam.

For the case with no scattering ($\sigma_s=0$), but with an external longitudinal magnetic field B_0, we can obtain the following expressions for the initial nonrotating beam ($L_i=L_0=0$) from (23) and (25):

$$\ln\left(\frac{R_f}{R_i}\right)^2+\frac{k_c^2}{4U}R_i^2\left[\left(\frac{R_f}{R_i}\right)^2-1\right]=\left(\frac{R_0}{R_i}\right)^2-1, \tag{31}$$

$$\frac{\Delta E^2}{(\gamma\beta c)^2}\simeq\frac{U}{2}(R_f^2-R_0^2)+\frac{k_c^2}{16}(R_f^4-R_0^4). \tag{32}$$

Figures 3 and 4 present the dependence of R_f/R_i and $\Delta E^2/(\gamma\beta c)^2$ on the induction B_0 of the external magnetic field for the REB parameters presented above. As is seen from the figures, the corresponding induction values can have an appreciable influence on the final equilibrium radius and the change in the mean-square emittance.

[1] E. P. Lee, Phys. Fluids **19**, 60 (1976).
[2] E. P. Lee and R. K. Cooper, Part. Accel. **7**, 83 (1976).
[3] E. R. Nadezhdin and G. A. Sorokin, Fiz. Plazmy **9**, 989 (1983) [Sov. J. Plasma Phys. **9**, 576 (1983)].
[4] E. P. Lee, Lawrence Livermore National Laboratory Report UCID-18495 (1980), p. 15.
[5] E. P. Lee and S. S. Yu, Lawrence Livermore National Laboratory Report UCID-18330 (1979), p. 23.
[6] M. Lampe and C. Joyce, Phys. Fluids **26**, 3371 (1983).
[7] E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **60**(3), 40 (1990) [Sov. Phys. Tech. Phys. **35**, 298 (1990)].
[8] E. K. Kolesnikov and A. S. Manuĭlov, Zh. Tekh. Fiz. **61**(12), 43 (1991) [Sov. Phys. Tech. Phys. **36**, 1351 (1991)].
[9] E. K. Kolesnikov, A. S. Manuĭlov, and I. V. Abashkina, Zh. Tekh. Fiz. **64**(11), 136 (1994) [Tech. Phys. **39**, 1160 (1994)].
[10] E. K. Kolesnikov and A. S. Manuĭlov, Pis'ma Zh. Tekh. Fiz. **17**(3), 46 (1991) [Sov. Tech. Phys. Lett. **17**, 96 (1991)].
[11] W. A. Barletta, E. P. Lee, and S. S. Yu, Nucl. Fusion **21**, 961 (1981).
[12] R. F. Fernsler, R. F. Hubbard, and M. Lampe, J. Appl. Phys. **75**, 3278 (1994).

Translated by P. Shelnitz

Amorphous gadolinium–cobalt films as sensitive media for the topographic mapping of nonuniform temperature fields

V. E. Ivanov, G. S. Kandaurova, and A. V. Svalov

A. M. Gor'kiĭ Ural State University, 620083 Ekaterinburg, Russia
(Submitted February 28, 1996)
Zh. Tekh. Fiz. **67**, 112–116 (July 1997)

The results of an investigation of the response of the domain structure of Gd–Co films to stationary, spatially nonuniform planar thermal fields created by one or two nominal point heat sources are presented. Several simple techniques for the topographic mapping of these fields by observing and measuring the parameters of the domain structure of amorphous Gd–Co films are formulated. © *1997 American Institute of Physics.* [S1063-7842(97)02107-7]

INTRODUCTION

The temperature effects of the domain structure of thin magnetic films open up possibilities to use them in the topographic mapping of spatially nonuniform thermal fields of various nature in cases in which other methods, which employ, in particular, paint films and liquid-crystal coatings, do not provide the desired result. The temperature dependence of the stripe-domain structure in permalloy films permits the topographic mapping of the thermal fields of fast processes[1] and stationary thermal fields.[2] The inconvenience of the method of observing the domain structure in these materials using powdered deposits has placed transparent iron garnet films[3] and amorphous Gd–Co and Gd–Fe films,[4] which are less expensive and simpler to prepare, among the promising materials. This paper is devoted to expanding the possibilities for the topographic mapping of spatially nonuniform (planar) stationary temperature fields.

SAMPLES AND EXPERIMENTAL METHOD

Amorphous Gd–Co films of thickness 0.5-1 μm with perpendicular anisotropy were obtained by hf ion-plasma sputtering on water-cooled glass substrates, and their free surface was protected by a glass coating of thickness $\sim$500 Å. The magnetic moment compensation temperature T_{com} was somewhat above room temperature and amounted to $T_{\text{com}}=60\pm3$ °C. Typical plots of the temperature dependence of the coercive force $H_c(T)$, its reciprocal $(H_c)^{-1}$, and the domain width d_0 in the demagnetized (equilibrium) state for the films investigated are presented in Fig. 1. These results, on the one hand, characterize the objects studied and, on the other hand, are needed to reconstruct a picture of the nonuniform temperature field $T(x,y)$ in the plane of a film from the parameters of the domain structure and the coercivity of the domain walls. The temperature gradient was created by the temperature difference between a relatively massive brass ''thermostat'' and a heat source, which was the junction of a copper–Constantan thermocouple. The ''source'' was heated by passing an electric current along Constantan conductors, and the thermopower was measured by a copper–Constantan thermocouple. The thermal contact surface between the source and the test amorphous film can be assumed to be a circle of radius $\sim$40 μm in a first approximation. The temperature T of the films investigated in this contact region was assumed to be constant and approximately equal to the source temperature T_s, if the film being investigated was in direct contact with the source. After the current through the source was established, the achievement of a stationary state by the system took approximately 20 s. The domain structure was observed by utilizing the polar magnetooptical Kerr effect either on the substrate side (in this case the film was in contact with the heat source) or on the film deposition side (in this case a glass substrate of thickness 0.24 mm was in thermal contact with the source). The effect of an external quasistatically varying magnetic field H_0 directed along a normal to the film was observed at a fixed value of the source temperature T_s; the film temperature T remained left constant. The state obtained after ''magnetic shaking'' by a sign-alternating magnetic field with a frequency of 50 Hz and an amplitude that decreases from $H_m>H_s$, where H_s is the static saturation field, to zero was assumed to be close to equilibrium. The domain width in the demagnetized state was measured on photomicrographs of the domain structures.

TOPOGRAPHIC MAPPING OF A CENTROSYMMETRIC TEMPERATURE FIELD FROM THE PARAMETERS OF THE INDUCED DOMAIN STRUCTURE OF AMORPHOUS Gd–Co FILMS

Variation of the temperature of amorphous Gd–Co films near the compensation temperature T_{com} leads to strong qualitative and quantitative changes in the magnetic properties and the domain structure.[5] Therefore, a noticeable response of the domain structure to the temperature gradient should be expected in a nonuniform temperature field that includes T_{com}. Figure 2a shows the simplest domain pattern of an amorphous Gd–Co film, which appears under the action of a nonuniform temperature field created by a nominal point heat source in the presence of a constant external field $H_0=5$ kOe. The domain structure of the film was observed from the substrate side, i.e., in this case the film was in contact with the heat source. In the original state, in which T_s is equal to room temperature T_0, i.e., $T_s=T_0<T_{\text{com}}$, the sample is single-domain. A quasistatic increase in T_s causes the abrupt appearance of a ''white'' domain of circular shape at a certain threshold temperature $T_s=T_{\text{thr}}=108$ °C, which corresponds to a radius of the circular domain $R_{\text{thr}}=210$

 1063-7842/97/070823-05$10.00

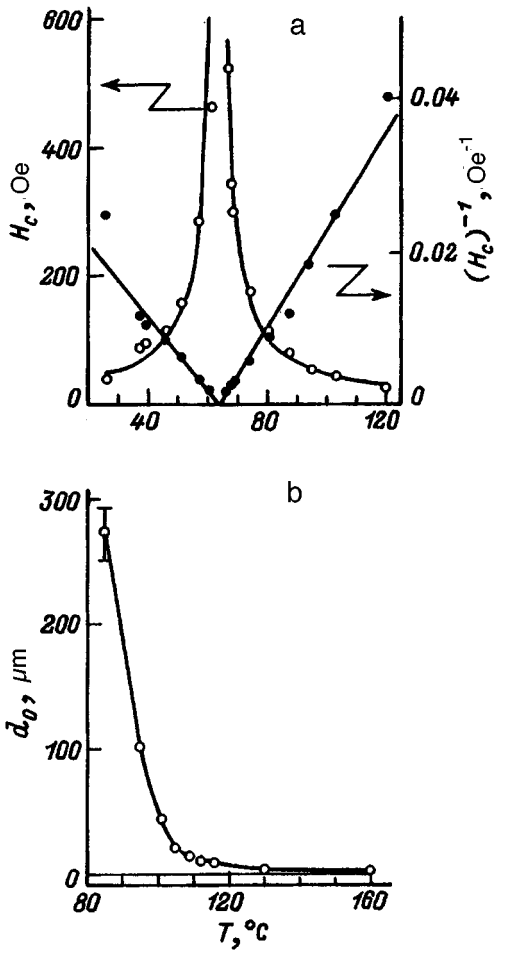

FIG. 1. Temperature dependences for the series of amorphous Gd–Co films investigated: a – coercive force and its reciprocal, b – domain width.

μm. A further increase in T_s leads to an increase in the radius R of the domain, the dependence of R on T_s being linear in a first approximation:

$$R(T)=R_0+\alpha\Delta T,$$

where $\alpha=8$ μm/°C is the temperature coefficient for variation of the radius.

At a fixed value of T_s, lowering the magnetic field strength to zero does not alter the picture of the domain structure, and the application of a field of opposite orientation leads to a domain pattern consisting of annular domains (Fig. 2b). As the field of opposite sign increases to its maximum value (-5 kOe) the two annular domain walls ''merge'' with the stationary domain wall of the original circular domain with the resultant formation of a ''black'' domain, which duplicates the shape of the original circular domain exactly; n magnetization-reversal cycles with successively decreasing values of the maximum magnetic field in each cycle (H_n) leads to the formation of $2n+1$ annular domain walls (Fig. 2c). The experiment shows that a domain pattern that is stable in a given field H_n remains unchanged when the field decreases to zero. It should be noted that in the temperature range $T_{\rm com}-40$ °C$<T<T_{\rm com}+40$ °C the state which is close to the equilibrium state (after magnetic ''shaking'') is a single-domain state or a random domain pattern resulting from halting of the domain walls at defects. Therefore, we call the domain structure described here, which forms only in a constant magnetic field, an induced domain structure. Figure 3 shows the distribution of the resultant magnetization M_s in a cross section of the film passing through the symmetry center. The drawing illustrates the mechanism of the formation of the domain pattern in Fig. 2b. For simplicity we assume that there is a linear decrease in the temperature from the source to the periphery, i.e., $dT/dx=$const, and that $M_s=0$ on the line where $T=T_{\rm com}$. As a consequence of the small values of M_s, the magnetostatic self-energy can be neglected. In this case the formation of the induced domain structure is determined by the competition between two forces: the external field H_0, which causes nucleation of the reverse magnetic phase and pushes the domain walls forward, on the one hand, and the coercive force $H_{cw}\sim C/M_s$, which hinders the nucleation and advancement of the domain walls, on the other hand. Thus, the domain pattern in Fig. 2 is determined by the distribution $H_{cw}(x,y)$, which, in turn, is determined by the distribution of the temperature, since $H_{cw}\sim C/M_s\sim c/(T-T_{\rm com})$. Here C and c are constants, which depend on the temperature in a first approximation. In Fig. 3 CDW denotes the compensation domain wall formed in the zeroth cycle, which is located on the line where $T=T_{\rm com}$ and H_{cw} reaches its maximum value. This wall subsequently remains stationary under any change in the field. The two mobile Bloch domain walls (BDWs), which appear after the first cycle of variation of H_0 (Fig. 3), are located on lines where $H_0=H_{cw}$. Since a specific value of H_{cw} and a specific value of T correspond to a given value of H_0, it is clear that the positions of the domain walls correspond to isotherms. The proposed techniques for the topographic mapping of a temperature field on the basis of the domain pattern in Fig. 2c are based on the following assumptions. The temperature at the domain wall appearing in the first cycle coincides with the compensation temperature of the test film, i.e., $T=T_{\rm com}$, and the maximum field in the nth cycle (H_n), which shapes the nth domain wall, coincides with the coercive force at the site of that domain wall.

Figure 4 shows a topogram of the temperature field $T(x,y)$ in the plane of a film, which was constructed in the following manner: a calibrated reference grid was preliminarily drawn on the photograph (Fig. 2b), the origin of coordinates being juxtaposed with the lower left-hand corner of the photograph. Using the foregoing arguments, the coercive force H_{cn} in the nth cycle was mapped to each pair of coordinates (x,y) of the domain wall. The distribution of the temperature in the plane of the film was found from the plot of $H_c(T)$ (Fig. 1a). It is seen from Fig. 4 that the topogram of $T(x,y)$ in the plane of the film is centrosymmetric.

Let us mention several features of the formation of an induced domain structure, which can lead to errors in the topographic mapping of temperature fields. These features include the hysteresis in the dependence of $R=R(T_s)$ and the hysteresis in the appearance and disappearance of the domain. The latter occurs because collapse of the domain occurs at a lower temperature ($T_{\rm col}\approx 80$ °C) than its appearance ($T_{\rm thr}=108$ °C) and corresponds to a radius $R_{\rm col}<R_{\rm thr}$.

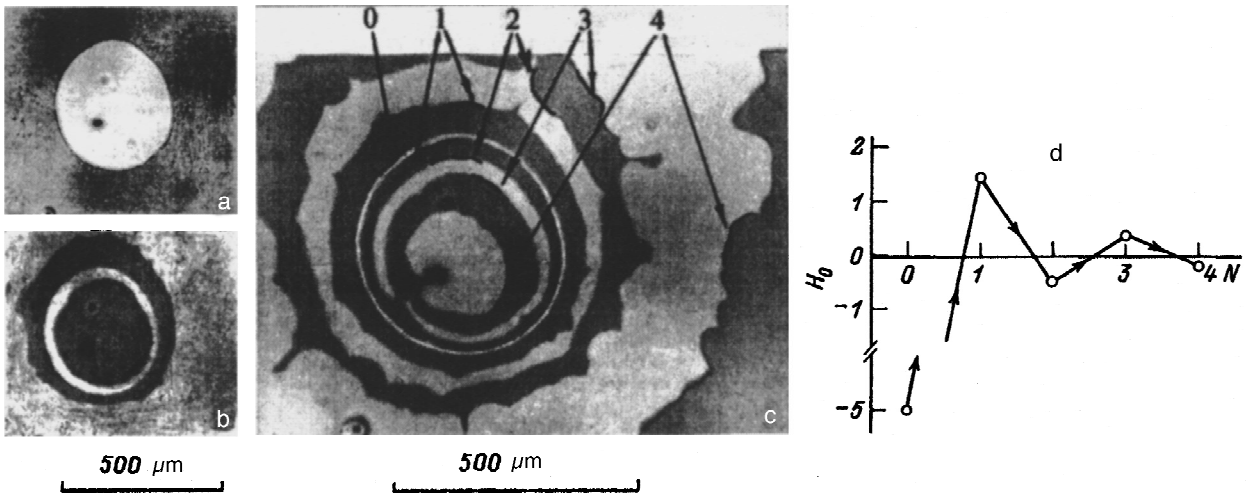

FIG. 2. Domain pattern formed in a nonuniform temperature field after the threshold temperature $T_s=108$ °C is achieved in a field $H_0=5$ kOe and the field is lowered to zero (a), in a magnetic field of opposite sign $H_0=-0.6$ kOe (b), and after four cycles of varying the direction and amplitude of the external magnetic field (c), as well as the field variation scheme (d), $T_s=80$ °C.

The temperature hysteresis of the appearance and collapse of the circular domain is most probably due to the delay in the formation of the magnetization-reversal nucleus in the region where T exceeds T_{com}. The duality of the $R(T)$ dependence can be attributed to the fact that a) the field H_0 is insufficient, i.e., $H_0<H_{c\max}$, and the domain wall starts only when H_{cw} decreases and becomes comparable to H_0, and b) there is a delay in the variation of the temperature in response to the variation of T_s at sites that are distant from the source.

TOPOGRAPHIC MAPPING OF A TEMPERATURE FIELD FROM THE PARAMETERS OF A SERPENTINE DOMAIN STRUCTURE OF AN AMORPHOUS Gd–Co FILM

Figure 5 shows the response of the domain structure of an amorphous Gd–Co film in the demagnetized state to a nonuniform temperature field in a regime in which a stationary state is achieved. The domain structure was observed from the film deposition side, i.e., in this case a substrate with a thickness of 0.24 mm was in thermal contact with the source. The initial state (Fig. 5a) was obtained by demagnetization by a variable magnetic field with a decaying amplitude at room temperature. After the source temperature $T_s=232$ °C was established, the stationary state of the system was achieved within about 20 s. It should be noted that during this time a small change in the source temperature T_s (from 232 to 235 °C) leads to significant alteration of the form of the domain structure (Figs. 5b–d). Quantitatively, this is manifested in significant alteration (Fig. 6a, curves *1–3*) of the coordinate dependence of the domain width d_0 (the x axis was oriented parallel to the long side of the frame in Fig. 5, and the origin of coordinates was juxtaposed with the left edge of the photograph). As the film temperature (at the center of the photograph) increases, the $d_0(x)$ curve becomes less steep. Plots of $T(x)$ (Fig. 6b) were constructed from the $d_0(x)$ and $d_0(T)$ (Fig. 1b) curves. The temperature on the surface of the test film varies quite considerably (by about 15 °C) during the achievement of the stationary (steady-state) regime, the slope of the $T(x)$ curves [i.e., the gradient of $T(x)$] remains approximately unchanged. A comparison of the temperature T_s of the thermocouple and the temperature in the plane of the film determined from the domain width (in the demagnetized state) reveals that the temperature difference amounts to 100 °C. Such a difference between the source temperature and the temperature determined from the parameters of the domain structure at the center of the photograph is most probably caused by the presence of the substrate between the source and the test

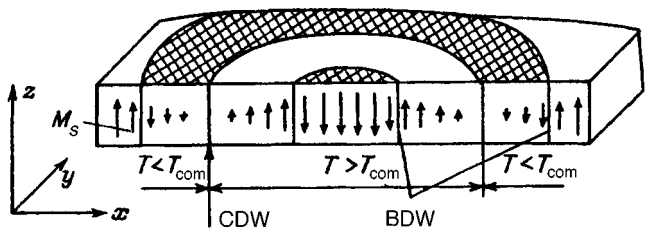

FIG. 3. Diagram of the distribution (along the coordinate x) of the resultant magnetization $\mathbf{M}_s$ on magnetization reversal of a circular domain.

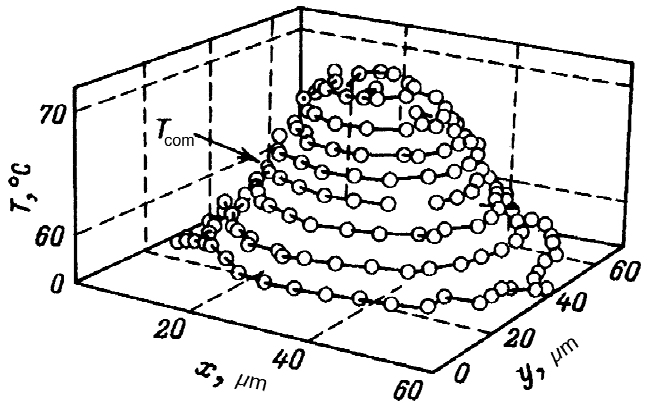

FIG. 4. Topogram of a nonuniform temperature field created by a nominal point heat source in the plane of a test amorphous Gd–Co film.

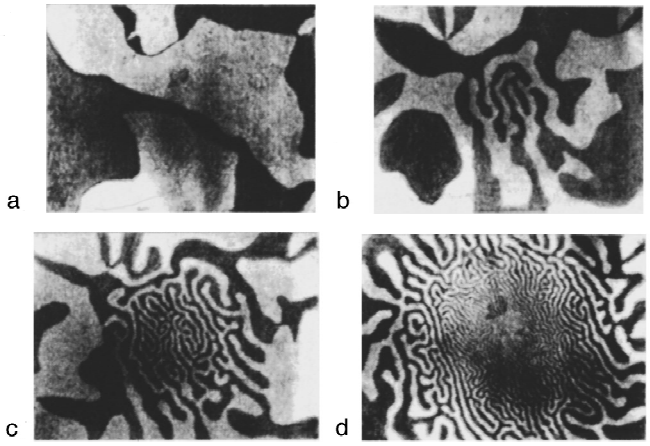

FIG. 5. Response of the domain structure of an amorphous Gd–Co film. Source temperature T_s, °C: a — 30, b — 232, c — 233, d —235.

film. All the experiments were henceforth carried out in the stationary regime, and the time for the establishment of that regime was equal to 20 s.

Nonuniform temperature fields of more complex form, created by several sources of different power, can be encountered in practice. For this reason, we shall next present some results of the influence of a temperature field created by two sources (which will be called S_1 and S_2).

Figure 7a presents a photograph of the equilibrium domain structure in a nonuniform temperature field created by two sources with different temperatures. The photograph was placed in a calibrated XY reference grid in order to measure the dependence of $d_0(x,y)$ and to reconstruct a topogram of the temperature field $T(x,y)$ from it. Knowledge of the dependence of d_0 on T (Fig. 1b) (in a uniform temperature field) makes it possible to qualitatively estimate the variation of a nonuniform temperature field from the form of the ser-

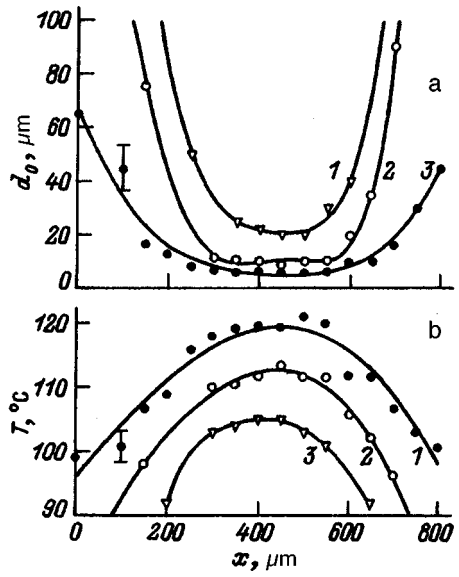

FIG. 6. Temporal variation of the one-dimensional dependence of the domain width (a) (in the demagnetized state) and of the temperature (b). T_s, °C: *1* — 232, *2* — 233, *3* — 235.

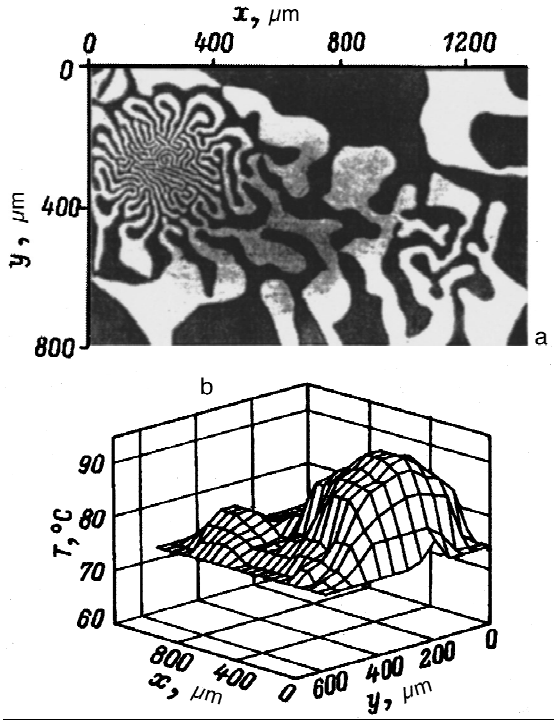

FIG. 7. Domain structure of an amorphous Gd–Co film (in the demagnetized state), on which a temperature gradient was created by two point heat sources: T_{s1}=227 °C (on the left), T_{s2}=177 °C (on the right) (a) and topogram of the nonuniform field (b).

pentine domain structure in the vicinity of S_1 and S_2. Several qualitative conclusions can be drawn from Fig. 7a. In particular, the two sources create domain structures of circular symmetry with some violation of the symmetry between them. In other words, the domain patterns created by S_1 and S_2 exhibit an interaction with one another, which can be interpreted as a manifestation of the overlap of the temperature fields of S_1 and S_2. Construction of a topogram of the nonuniform temperature field requires measurement of the dependence of the domain width on the two coordinates in the plane of the film, i.e., $d_0(x,y)$, which reflects the properties of the temperature field and the dependence of d_0 on T (Fig. 1b), the latter being a characteristic of the test film. An example of a topogram of a planar temperature field $T(x,y)$ created by two point sources is presented in Fig. 7b. Temperature hysteresis[6] and defects of various kinds, including stabilization of the domain walls,[7] distort the picture of the domain structure and introduce errors into the determination of the temperature field. To avoid such distortions of the domain structure, first, the details of the technology of depositing defect-free films, particularly the measures taken to prevent stabilization of the domain walls, must be carefully observed,[6] and, second, the test films must be thoroughly demagnetized by a variable field of sufficient amplitude.

Thus, it has been established as a result of this research that the response of domain structures of amorphous Gd–Co films with perpendicular anisotropy to nonuniform stationary thermal fields points out a strong correlation between quantitative parameters of the domain structure and characteristics of the temperature field. This provides some basis to regard amorphous films of the Gd–Co type as potentially

suitable for the topographic mapping of spatially nonuniform stationary temperature fields of complex pattern.

[1] B. M. Abakumov, I. A. Pan'shin, E. A. Podpalyĭ *et al.*, *Recording Optical Information on Thin Magnetic Films* [in Russian], Atomizdat, Moscow (1976), p. 120.
[2] A. G. Ravlik, V. N. Samofalov, Yu. V. Zolotnitskiĭ *et al.*, Prib. Tekh. Éksp. (3), 241 (1988).
[3] V. V. Randoshkin and A. Ya. Chervonenkis, *Applied Magnetooptics* [in Russian], Énergoatomizdat, Moscow (1990).
[4] A. E. Gafner, I. A. Pan'shin, E. A. Podpalyĭ *et al.*, *Reports to the 8th All-Union Seminar ''New Magnetic Microelectronics Materials''* [in Russian], Donetsk, 1982, pp. 356–357.
[5] G. S. Kandaurova and V. E. Ivanov, in *Amorphous Film Alloys of Transition and Rare-Earth Metals* [in Russian], Krasnoyarsk (1988), pp. 150–163.
[6] V. E. Ivanov, G. S. Kandaurova, and E. F. Ledovskaya, Izv. Vyssh. Uchebn. Zaved., Fizika (4), 37 (1988).
[7] V. E. Ivanov, G. S. Kandaurova, M. F. Karimov *et al.*, Fiz. Met. Metalloved. **79** (3), 59 (1995).

Translated by P. Shelnitz

Potential possibilities of a laser plasma as a negative-ion source

S. V. Latyshev

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia
(Submitted February 20, 1996)
Zh. Tekh. Fiz. **67**, 117–120 (July 1997)

The potential possibilities of a plasma heated by laser radiation as a negative-ion source are analyzed theoretically. It is shown that the efficiency of negative-ion formation in a laser plasma in the heating phase reaches $10^{15}-10^{16}$ ions/J when the parameters of the laser radiation are optimally adjusted. © *1997 American Institute of Physics.* [S1063-7842(97)02207-1]

Definite interest has recently been displayed in the study of the plasma produced by focused laser radiation on the surface of a solid as a negative-ion source. This interest is due mainly to the need to create an intense negative-ion source for accelerator technology.[1] The research is presently confined to two areas. In one of them the expansion products of the laser plasma are investigated at large distances from the target (of the order of several meters) using time-of-flight mass spectrometry.[2,3] It was shown in these studies that a significant negative-ion yield (on the level of several percent in comparison with the positive, singly charged ions) appears when the plasma is heated by very intense laser radiation. For example, when a CO_2 laser ($\lambda=10.6$ μm) was employed, the maximum negative-ion yield was observed for flux densities $q\sim(3-5)\times10^9$ W/cm^2, and when a neodymium laser ($\lambda=1.06$ μm) was employed, the corresponding flux densities significantly exceeded 10^{10} W/cm^2. At such flux densities the temperature of the laser plasma in the heating phase has a value of $10-30$ eV, which significantly exceeds the temperature at which negative ions are efficiently formed ($T\sim1$ eV). The researchers reporting these results believe that the formation of negative ions in a laser plasma takes place mainly in the plasma expansion phase, in which the temperature drops to appropriate values.

In other studies negative ions were extracted from a laser plasma by a strong electrostatic field directly in the heating phase or immediately after removal of the heating radiation.[4–6] Neodymium lasers were used in these studies, and the maximum negative-ion yield was observed at flux densities $q\sim6\times10^7-6\times10^8$ W/cm^2.

From the standpoint of creating a laser negative-ion source, the second route is preferable, since it does not require large flux densities of laser radiation and allows us to hope to create a continuous negative-ion source with presently existing laser technologies.

This paper is devoted to a theoretical analysis of the second group of experiments. The present analysis is based on the assumption that thermodynamic ionization equilibrium is achieved in a laser plasma in the heating phase at flux densities of the heating radiation equal to 10^7-10^9 W/cm^2. In fact, the following criteria are satisfied for characteristic laser plasma densities $n_e\sim10^{19}-10^{21}$ cm^{-3}, a temperature $T\sim1$ eV, and a time for existence of the plasma of the order of the hydrodynamic expansion time $\tau\cong d/V\sim10$ ns (the diameter of the focal spot $d\sim10$ μm, and the velocity of sound $V\sim10^5$ cm/s): the criterion of classicality

$$\varepsilon_f=\frac{2\pi^2\hbar^2}{m_e}\left(\frac{3n_e}{8\pi}\right)^{2/3}\ll\frac{3}{2}T,$$

the criterion of ideality

$$\frac{e^2}{\langle r_{ee}\rangle}=e^2n_e^{1/3}\ll\frac{3}{2}T,$$

and the criterion of the attainability of equilibrium

$$\tau_{eq}=\frac{1}{\sigma_{ea}n_ev_e}\ll\tau,$$

where the cross section for the interaction of an electron with an atom $\sigma_{ea}\sim10^{-16}$ cm^2, the thermal electron velocity $v_e\sim10^7$ cm/s, ε_f is the Fermi energy, and $\langle r_{ee}\rangle$ is the mean distance between electrons.

The influence of the photodissociation of negative ions under the action of laser radiation on the ionic composition is more complicated. When multiphoton processes are neglected, photodissociation of the negative ions does not occur, if $\hbar\omega<U_-$, where $\hbar\omega$ and U_- are the photon energy and the binding energy of the electron in the negative ion. This condition generally holds for radiation from a CO_2 laser ($\hbar\omega\approx0.1$ eV), but far from always for radiation from a neodymium laser ($\hbar\omega\approx1$ eV). The corresponding estimates of the photodissociation cross section $\sigma_{\rm ph}\sim10^{-17}$ cm^2 (Ref. 7) show that when the radiation flux density $q\sim10^8$ W/cm^2, the time of the photodissociation process $\tau_{\rm ph}=\hbar\omega/q\sigma_{\rm ph}$ for radiation from a neodymium laser equals 10^{-10} s. This time is comparable to the time for the establishment of ionization equilibrium $\tau_{\rm eq}$; therefore, short-wavelength laser radiation can significantly reduce the number of negative ions in the plasma through photodissociation.

When the above criteria are satisfied, the ionization equilibrium in the plasma is described by a system of Saha equations. In situations where negative ions are formed efficiently in the plasma, the existence of positively charged ions with an extent of ionization greater than $+1$ can be neglected. With the additional neglect of the existence of molecular formations, which is valid only for a restricted list of substances, we obtain a system of Saha equations in the following form:

$$\begin{aligned}
&C_eC_+/C_0=a,\quad a=(T/E)^{3/2}\exp(-U_0/T),\\
&C_eC_0/C_-=A,\quad A=(T/E)^{3/2}\exp(-U_-/T),\\
&C_e=C_+-C_-,\quad C_0+C_++C_-=1,
\end{aligned}\tag{1}$$

 1063-7842/97/070828-04$10.00

where $C_e=n_e/N$, $C_+=n_+/N$, $C_-=n_-/N$, and $C_0=n_0/N$ are the relative concentrations of electrons, positive ions, negative ions, and neutral atoms; N is the total concentration of atomic particles; and U_0 and U_- are the ionization potential of the neutral atom and the binding energy of the electron in the negative ion.

The quantity $E=2^{1/3}\pi\hbar^2N^{2/3}/m_e$ has the dimensions of energy and is highly reminiscent of the Fermi energy for an electron gas, being distinguished only by the fact that it contains the total concentration of atomic particles instead of the concentration of electrons and by a slight difference in the numerical multiplier (in the Fermi energy the numerical factor is $3^{2/3}\pi^{4/3}/2\cong 4.78$, while in E it is $2^{1/3}\pi\cong 3.96$). The values of E for several concentrations of atomic particles are presented below.

N, cm^{-3}	E, eV
10^{19}	0.014
10^{20}	0.065
10^{21}	0.30
10^{22}	1.4

In Ref. 6 the system of Saha equations presented above was investigated numerically for several simple substances in the temperature range $T\sim 0.3-4.5$ eV and the range of concentrations of atomic particles $N=10^{19}-10^{22}$ cm^{-3}. An analysis of the results obtained revealed the following laws: the relative concentration of the negative ions increases with increasing plasma density and reaches 10% at the plasma density $N=10^{22}$ cm^{-3} and the corresponding optimum temperature; as the plasma density increases, the optimum temperature, at which the maximum relative concentration of the negative ions is achieved, increases. The typical values of the optimum temperatures are $T\sim 1-4$ eV.

This paper describes an analytical investigation of the system of Saha equations presented above and proposes some simple analytical estimates of the optimum parameters of a laser plasma as a negative-ion source.

Solving the first, second, and fourth equations of system (1) with respect to the electron concentration C_e, we obtain the following equations for the concentrations:

$$C_0=1/(1+a/C_e+C_e/A),$$

$$C_+=1/(1+C_e/a+C_e^2/(aA)),$$

$$C_-=1/(1+A/C_e+aA/C_e^2). \qquad (2)$$

Substituting the expressions for the concentrations (2) into the third equation of system (1), we obtain a cubic equation for finding C_e

$$C_e^3+(A+1)C_e^2+aAC_e=aA. \qquad (3)$$

Using Cardan's formulas to solve Eq. (3) is not convenient from the standpoint of practical calculations. For this reason, attempts were made to obtain an approximate solution which would achive sufficient accuracy while remaining rather simple.

As a result of the investigations performed, it was found that the solution of the quadratic equation obtained from Eq. (3) by discarding the cubic term differs from the exact solution by no more than 10%, while it is fairly simple:

$$C_e=2/[1+(1+4/b)^{1/2}], \quad b=aA/(1+A). \qquad (4)$$

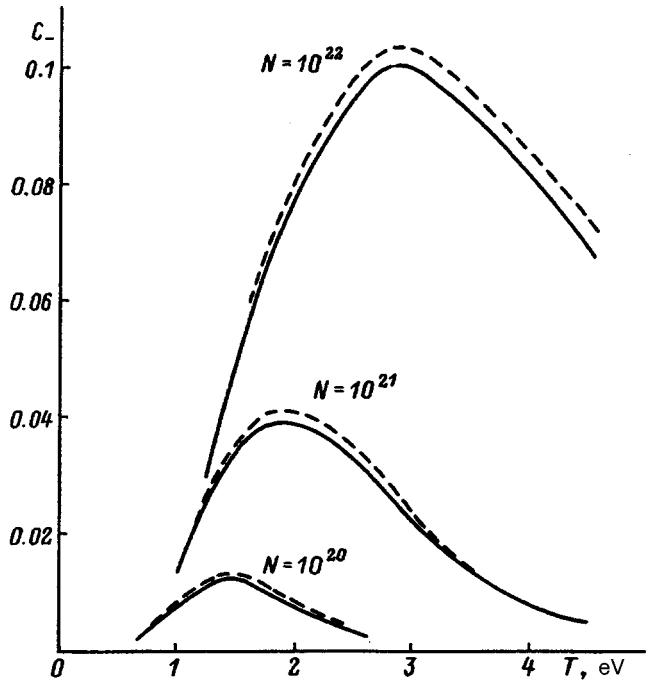

FIG. 1. Dependence of the relative concentration of Br^- ions on the temperature for various plasma densities (solid curves — exact calculation, dashed curves — approximate calculation).

The very high accuracy of Eq. (4) is attributed to the fact that the cubic term is always smaller, and in real systems much smaller, than the quadratic term, since $C_e\leqslant 1$, and $A\to\infty$ as $C_e\to 1$. Figure 1 compares the exact solution of system (1) from Ref. 6 and the approximate solution based on Eqs. (2) and (4) in the case of the negative ion of bromine Br^-.

Knowledge of complete solutions of the Saha equations, either exact or approximate, is actually not necessary for many practical problems. It is very often quite sufficient to know the optimum parameters of the plasma: the temperature at which ions charged to the extent of interest are present in the plasma in the largest possible number and the corresponding value of their concentration.

The equation which specifies the optimum plasma temperature T^*, at which the relative concentration of the negative ions in the plasma is maximal (C_-^*), is obtained in the following manner. We differentiate all the equations in system (1) with respect to the temperature and substitute $T=T^*$ into the relations obtained. Then, taking into account the condition for the maximum of the concentration of the negative ions, $C_-^{*\prime}=0$, we obtain a system of four linear equations for the three derivatives $C_e^{*\prime}$, $C_+^{*\prime}$, and $C_0^{*\prime}$. The condition under which this system is solvable implicitly assigns the optimum plasma temperature T^*

$$C_+^*(C_0^*-C_e^*)/(C_e^*C_0^*+C_e^*C_+^*+C_+^*C_0^*)=\gamma,$$

$$\gamma=(1.5T^*+U_-)/(1.5T^*+U_0). \qquad (5)$$

We note that the coefficient γ depends weakly on the temperature and falls within the range from U_-/U_0 (for $T^*=0$) to 1 (for $T^*=\infty$). In real systems $\gamma\approx 0.2-0.3$, since the optimum temperature always satisfies the inequality $T^*\ll U_0$, and $U_-/U_0\approx 0.1-0.2$.

In order to determine the optimum temperature from Eq. (5), substitution of the exact solutions of system (1) into Eq. (5) is formally necessary, but such a course of action is unrealistic, since an extremely cumbersome equation is obtained even when the approximate solutions of system (1) in the form of Eqs. (2) and (4) are plugged into it. Therefore, an attempt was made to replace the concentrations by the solutions of the Saha equations in which only two particle states, viz., neutral atoms and positively charged ions, are taken into account:

$$C_eC_+/C_0=a, \quad C_e=C_+, \quad C_++C_0=1. \tag{6}$$

The combined solution of this system and Eq. (5) leads to the following expressions for the optimum temperature and the concentrations:

$$a^*=(1-2\gamma)^2/(2+\gamma-\gamma^2),$$

$$C_+^*=(1-2\gamma)/(2-\gamma),$$

$$C_0^*=(1+\gamma)/(2-\gamma). \tag{7}$$

As we have already noted above, the parameter γ depends weakly on the temperature and generally takes the values $\gamma\approx 0.2-0.3$. In this case the concentrations of the positively charged ions and the neutral atoms at which the relative concentration of the negative ions is maximal are

$$C_+^*\approx C_e^*\approx 0.2-0.3; \quad C_0^*\approx 0.7-0.8. \tag{8}$$

The results of about 70% of the 76 exact calculations presented in Ref. 6 fall in the ranges of values (8), and the remaining results are close to these ranges. As for the equation for determining the optimum temperature, it takes on the form $a^*\sim 0.05-0.15$. A comparison with exact calculations reveals that the best agreement with the exact calculations is exhibited by the equation

$$a^*=(T^*/E)^{3/2}\exp(-U_0/T^*)=0.1. \tag{9}$$

Figure 2 presents a comparison of the exact values of the optimum temperatures for several substances from Ref. 6 with approximate equation (9). This comparison shows that the disparity does not exceed 10%. Approximately the same accuracy is displayed by the simple approximate formula

$$T^*\cong U_0/[\ln(U_0/E)+1.5]. \tag{10}$$

The satisfactory agreement with the exact calculations of the optimum temperature T^*, as well as the corresponding values of the concentration of the charged particles C_+^* and the concentration of the neutral atoms C_0^*, allows us to give a fairly simple formula for calculating the maximum concentration of the negative ions. For this purpose, we substitute the corresponding concentration values from (8) and the approximate expression (10) for the optimum temperature into the expression for C_- obtained from the first and second equations in system (1). We then have

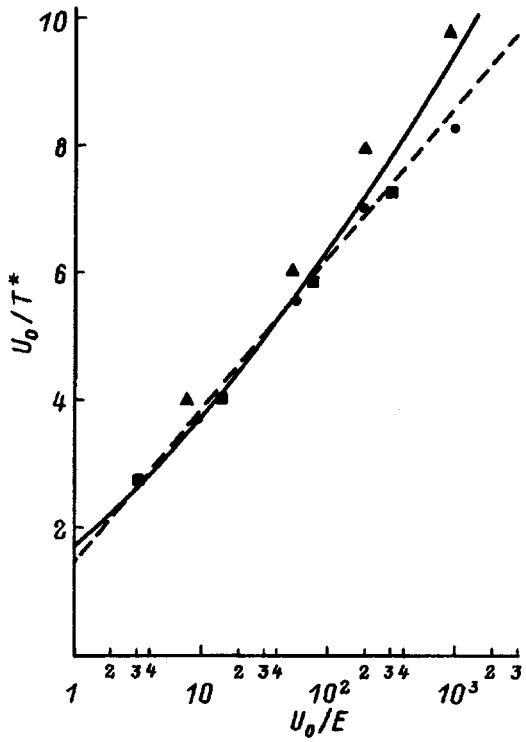

FIG. 2. Plots for determining T^*: solid curve — analytical dependence, dashed curve — approximate formula: ● — H^-, ■ — K^-, ▲ — Br^-.

$$C_-^*=\frac{C_o^{*2}}{C_+^*}\exp\left(-\frac{U_0-U_-}{T^*}\right)\cong\left(\frac{E}{3U_0}\right)^{(1-U_-/U_0)}. \tag{11}$$

The accuracy of formula (11) is of the order of 1.5, which is sufficient for design calculations of ion sources.

One consequence of Eq. (9) for the optimum temperature and approximate formulas (10) and (11) is that the optimum temperature and the concentration of the negative ions depend mainly on the value of U_0/E, which contains the plasma concentration N and the ionization potential of the neutral atom U_0 and depends weakly on the binding energy of the electron in the negative ion U_-. Formulas (9)–(11) also confirm the correlations noted in Ref. 6, viz., the increases in the optimum temperature and the relative concentration of the negative ions with increasing plasma concentrations.

The energy efficiency of the generation of negative ions equals

$$Ef=C_-^*/W_p\cong 10^{18}C_-^* \ [\text{ions/J}], \tag{12}$$

where $W_p=3/2T^*(C_0^*+C_+^*+C_-^*+C_e^*)+U_0C_+^*-U_-C_-^*$ is the energy of the plasma.

Taking into account that for a laser plasma $C_-^*\sim 10^{-2}-10^{-1}$, we obtain a limiting estimate of the energy efficiency of the generation of negative ions in a laser plasma: $Ef\sim 10^{16}-10^{17}$ ions/J. The quasi-two-dimensional hydrodynamic calculations performed in Ref. 6 give the more modest estimate $Ef\sim 10^{15}-10^{16}$ ions/J. This is because the density and temperature profiles in a laser plasma

vary in opposite directions, and the optimum conditions for the generation of negative ions are consequently realized only in a small region.

Thus, the results of this work show that a laser plasma is a highly efficient source of negative ions. For example, when a 100 W continuous-wave laser is used, the creation of a continuous negative-ion source with a current equal to 10−100 mA is perfectly realistic. The diameter of the optimum focal spot of the laser radiation should then be of the order of 10 μm. It is preferable to use a short-wavelength laser, but the condition $U_- > \hbar\omega$ must then be satisfied. The greatest difficulty in developing a laser negative-ion source lies in the system for extracting the ions from the hot plasma. The experience that has been gained in creating pulsed laser sources of negative ions,[3–5] as well as highly charged positive ions (see, for example, Ref. 8), can be useful in this area.

[1] D. G. Koshkarev, Nuovo Cimento A **106**, 1567 (1993).
[2] Yu. A. Bykovskiĭ, V. I. Romanyuk, and S. M. Sil'nov, Pis'ma Zh. Tekh. Fiz. **14**, 927 (1988) [Sov. Tech. Phys. Lett. **14**, 410 (1988)].
[3] Yu. A. Bykovskiĭ, V. I. Romanyuk, and S. M. Sil'nov, in *13th International Conference on Coherent and Nonlinear Optics* [in Russian], 1988, Part 1, p. 365.
[4] G. Korschinek and T. Henkelmann, Nucl. Instrum. Methods Phys. Res. A **302**, 376 (1991).
[5] T. Henkelmann, J. Sellmair, and G. Korschinek, Nucl. Instrum. Methods Phys. Res. B **56/57**, 1152 (1991).
[6] S. V. Latyshev, G. Korschinek, and T. Henkelmann, Preprint No 26 [in Russian], Institute of Theoretical and Experimental Physics, Moscow (1992).
[7] B. M. Smirnov, *Negative Ions*, McGraw-Hill, New York–London (1982).
[8] Y. A. Bykovskii, S. V. Latyshev, and B. Y. Sharkov, Laser Part. Beams **2**, 49 (1984).

Translated by P. Shelnitz

Thermally stimulated emission of electrons and photons in nonlinear LiB_3O_5 crystals

I. N. Ogorodnikov, V. I. Kirpa, A. V. Kruzhalov, and A. V. Porotnikov

Ural State Technical University, 620002 Ekaterinburg, Russia
(Submitted February 21, 1996)
Zh. Tekh. Fiz. **67**, 121–125 (July 1997)

The results of an investigation of the thermally stimulated emission of electrons and photons in lithium triborate single crystals are presented. The investigation is performed mainly by simultaneously measuring the thermally stimulated emission of electrons and photons after various radiation treatments. In particular, the thermally stimulated exoelectron emission of LiB_3O_5 crystals is discovered for the first time and studied in detail. An analysis of the experimental results reveals two new trapping centers. It is established that the thermal annealing of these centers at 450 and 515 K makes a contribution only to the exoelectron emission. It is postulated that a field fluctuation process, which lowers the potential barriers of the trapping centers, takes place in LiB_3O_5 at $100-140$ K. This process is attributed to thermally stimulated relaxation of the radiation-induced charge. A significant dependence of the parameters of the thermally stimulated processes in LiB_3O_5 on the parameters of the radiation treatment is discovered.
 [S1063-7842(97)02307-6]

INTRODUCTION

In Refs. 1–3 we began a study of the suprathermal emission of electrons and photons in thermally stimulated processes in the condensed state of matter using modulation thermal activation spectroscopy and a model of the field fluctuation processs (FFPs). The results obtained in Refs. 2 and 3 demonstrated the promising nature of the approach selected for analyzing a long list of phenomena. Suprathermal emission is, in particular, a factor which limits the possibility of utilizing solid-state dosimeters at small doses of ionizing radiation, where the contributions of the spontaneous emission and radiation-induced processes to the thermally stimulated luminescence (TSL) or thermally stimulated exoelectron emission (TSEE) become comparable. In nonlinear- and integrated-optical elements suprathermal emission can be one of the factors producing electron avalanches, which lead to lowering of the threshold for optical breakdown of the insulator and, consequently, to an increase in the probability of optical breakdown in intense optical fields of laser radiation.

Nonlinear crystals of lithium triborate LiB_3O_5 (space group $Pna2_1$, transparency band at 159–3500 nm) represent a promising material for nonlinear and integrated optics owing to the fortunate combination of comparatively high nonlinear characteristics and an elevated threshold for optical breakdown. In Refs. 4 and 5 we performed initial investigations of the point defects, luminescence, and thermally stimulated recombination processes in lithium triborate crystals following exposure to comparatively large doses of ionizing radiation. The experience gained from studying spontaneous emission in other materials [for example, in BeO (Refs. 1 and 2)] indicates that effective suppression of the spontaneous emission occurs as the dose of ionizing radiation increases. For this reason, reliable experimental data on the spontaneous emission can be obtained only for unirradiated materials or for samples irradiated in the range of so-called small doses of ionizing radiation. The exact limits of this range depend not only on the properties of the material, but also on its history, the experimental conditions, the type of ionizing radiation, etc. A special investigation of the conditions for the appearance of spontaneous emission is required for each specific material.

The purpose of the present work was to investigate the thermally stimulated emission of electrons and photons from LiB_3O_5 crystals by modulation thermal activation spectroscopy after exposure to ''small doses'' of various types of ionizing radiations.

MATERIALS AND METHODS

Lithium triborate single crystals of high optical quality were grown under the guidance of V. A. Maslov by a modified fluxed melt method with drawing of the seed in platinum crucibles within single- and two-zone furnaces.[6] The principal crystallographic characteristics of the crystals obtained are in good agreement with the data in Refs. 7 and 8. Samples in the form of plane-parallel polished plates with a normal perpendicular to the Z axis of the crystals were prepared for the investigation.

The thermally stimulated emission of electrons and photons from LiB_3O_5 crystals was investigated by performing simultaneous TSEE and TSL measurements in an ultrahigh vacuum (with a residual pressure of $\sim 10^{-6}$ Pa) in a linear or linear–oscillatory mode of temperature variation at 77–600 K.[9] The basic feature of the linear–oscillatory temperature variation is modulation of the linear heating law of the sample by relatively rapid heating–cooling temperature oscillations with an amplitude up to 10 K. The period of one oscillation and the mean increase in temperature between oscillations were equal to 100 s and 2.5 K, respectively. The mean heating rate is considerably lower in this case than in the case of strictly linear heating. Modulation of the sample temperature variation law by heating–cooling temperature oscillations leads to corresponding oscillations of the luminescence and emission response, whose analysis yields

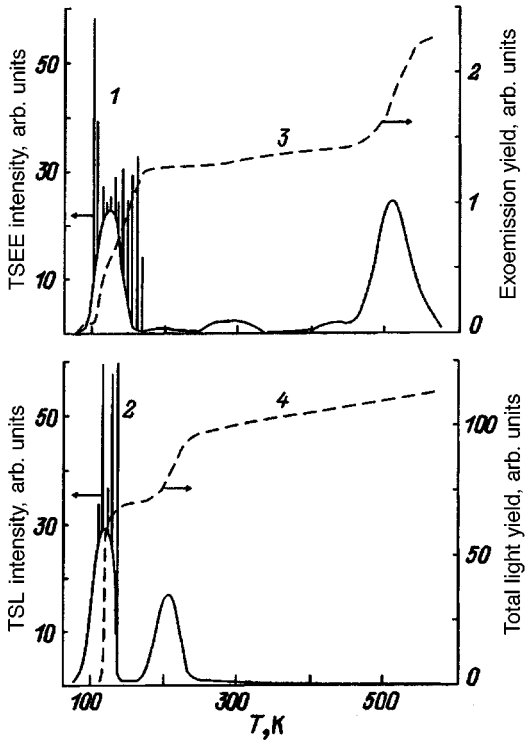

FIG. 1. TSEE (*1*), TSL (*2*), total exoemission yield (*3*), and total light yield (*4*) of LiB_3O_5 crystals during linear heating at the rate 10 K/min following electron bombardment at 77 K.

qualitative information on the parameters of the thermally stimulated processes. The principle parameter determined in this case is the ''experimental mean activation energy'' $\langle E \rangle$

$$\langle E \rangle = \frac{\Sigma_{i=1}^{N} E_i \cdot I_i}{\Sigma_{i=1}^{N} I_i}, \tag{1}$$

where E_i and I_i are the activation energy and the contribution of the ith elementary relaxation process.

A calculation of $\langle E \rangle$ is performed for each temperature oscillation by the initial raise method using the formula

$$\langle E \rangle = -K_B \frac{d \ln(I)}{d(1/T)}, \tag{2}$$

where K_B is Boltzmann's constant, and $I = I(T)$ is the experimentally recorded intensity of the exoelectron emission (a VÉU-6 secondary electron multiplier) or luminescence (an FÉU-106 photomultiplier in the photon counting mode).

The LiB_3O_5 crystals were irradiated by x radiation from a BSV2 x-ray tube (25 kV, 10 mA, copper anticathode) or by electron bombardment (10 keV, 500 μA) at various temperatures.

EXPERIMENTAL RESULTS

The results of the simultaneous measurements of the TSL and TSEE curves of LiB_3O_5 crystals during linear heating at the rate of 10 K/min following electron bombardment at 77 K are presented in Fig. 1. It follows from it that the TSL is concentrated mainly in the range 100–140 K, where

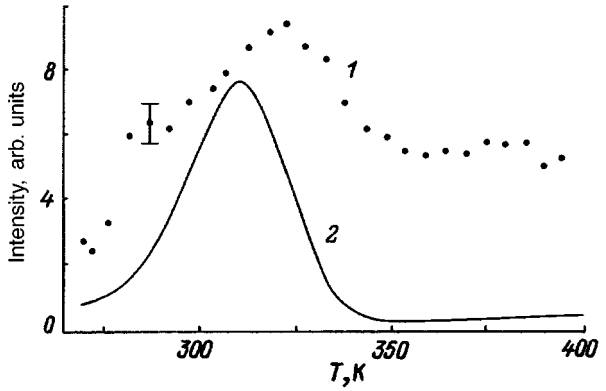

FIG. 2. TSEE (*1*) and TSL (*2*) of LiB_3O_5 crystals during linear heating at the rate 10 K/min following irradiation by x-ray photons at 250 K.

more than half of the stored total light yield is emitted, and in a peak at 210 K. Weakly expressed shoulders are observed on the TSL curve at 90–95, 150–180, and 230–300 K. No TSL is detected above 350 K in LiB_3O_5.

In the case of TSEE the release of a considerable portion of the stored total exoemission sum (about half) also begins at 100 K, but covers a broader temperature range from 100 to 180 K than in the case of TSL and is characterized by the absence of the intense peak at 210 K. The TSEE curve contains shoulders at 90–95, 180–230, and 250–330 K. In the high-temperature region the TSEE pattern differs significantly from that for TSL: the TSEE curve displays an intense peak at 551 K and a pronounced shoulder at 400–480 K.

Special mention should be made of the appearance of so-called spontaneous emission, i.e., short intense spikes of electron (100–180 K) and photon (100–140 K) emission, in LiB_3O_5 during heating following electron bombardment. The irradiation of LiB_3O_5 with x-ray photons at 250 K enabled us to create a larger concentration of trapping centers whose thermal decay produces TSL and TSEE peaks at 300–350 K (Fig. 2), and we obtained comparable values for the TSL and TSEE intensities in this temperature range. The corresponding TSEE curve has a gently sloping maximum at 330 K, and the TSL curve contains a clearly distinguished peak at 310 K.

The results of the simultaneous measurements of the TSL and TSEE curves of LiB_3O_5 crystals in the linear–oscillatory temperature mode following irradiation by x-ray photons at 77 K are presented in Fig. 3. As follows from Fig. 3, the temperature positions of these spikes correlate with the temperature oscillations over the entire range investigated: an increase in temperature causes an increase in the emission of electrons and photons, the emission maximum corresponds to the temperature maximum of the oscillation, and the drop in temperature during an oscillation causes a corresponding drop in emission. In other words, the oscillations of the temperature, luminescence, and electron emission are strictly correlated (symbatic).

It follows from Fig. 3 that the envelope of the luminescence response is represented by two peaks at 95 and 200 K, on which the luminescence oscillations induced by the corresponding temperature oscillations are superimposed. This

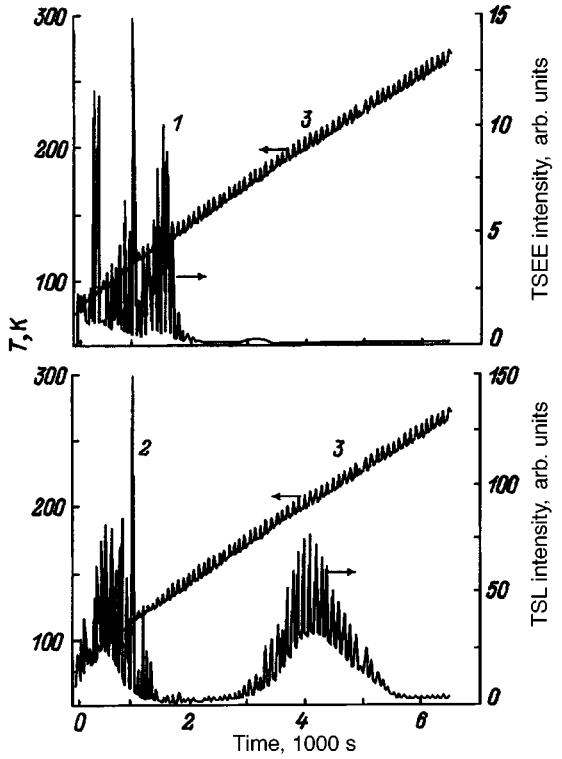

FIG. 3. Thermally stimulated emission of electrons (*1*) and photons (*2*) from LiB_3O_5 crystals during linear–oscillatory heating (*3*) following electron bombardment at 77 K.

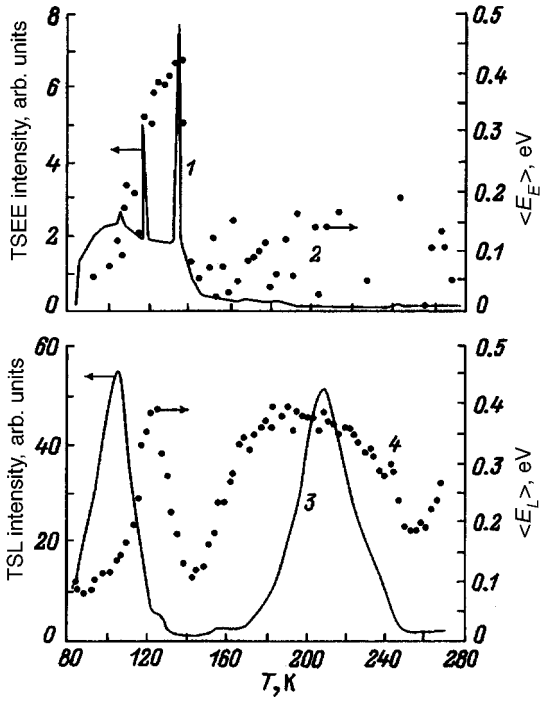

FIG. 4. Temperature dependence of the mean thermal activation energy (*2*, *4*) and of the intensity (*1*, *3*) of the TSEE (*1*, *2*) and TSL (*3*, *4*) of LiB_3O_5 crystals during linear–oscillatory heating following irradiation by x-ray photons at 77 K.

envelope can be matched to the temperature positions of the TSL maxima in Fig. 1. The exoemission following x irradiation at 77 K (Fig. 3) is less pronounced: the envelope of the emission response is represented by a damped, weakly structured curve, on which emission oscillations are superimposed. At 100–120 (luminescence) and 100–140 K (exoelectron emission) the amplitude of the electron and photon emission responses varies randomly: intense spontaneous emission is superimposed on the regular thermally stimulated process. However, in all cases the oscillations of the luminescence and the exoelectron emission mimic the temperature oscillations (Fig. 3).

Figure 4 presents plots of the temperature dependence of the mean activation energies for TSEE ($\langle E_E\rangle(T)$) and TSL ($\langle E_L\rangle(T)$) in LiB_3O_5 crystals. The data for each temperature point were obtained by treatment of the experimental data presented in Fig. 3 by the initial raise method (curve *2*). At 77–100 K $\langle E_L\rangle(T)\approx 0.08-0.09$ eV, and as the temperature rises further, $\langle E_L\rangle(T)$ at first increases to a maximum value $E_{\max}\approx 0.39-0.40$ eV (120–125 K) and then drops to the 0.10 eV level (140 K). A similar peak is observed for TSEE on the $\langle E_E\rangle(T)$ curve ($E_{\max}\approx 0.42$ eV at 140 K). The numerical value of the mean activation energy $E_{\max}$ is elevated at this temperature and cannot be explained within the simple model of the thermally stimulated delocalization of charge carriers from trapping centers. Above 140 K the $\langle E_E\rangle(T)$ and $\langle E_L\rangle(T)$ curves do not have such features. The considerable spread of the values of $\langle E_E\rangle(T)$ at 140–280 K is a consequence of the low TSEE intensity. The increase in the numerical values of $\langle E_L\rangle(T)$ at 140–180 K is stipulated, accordingly to Eq. (1), by the increase in the contribution of the charge carriers delocalized in this temperature range to the overall recombination process. In the region of the TSL peak (180–250 K) the $\langle E_L\rangle(T)$ curve apparently reflects the temperature dependence of the thermal activation energy of the recombination process.

The mean activation energies for TSEE during linear–oscillatory heating at 400–600 K are presented in Fig. 5. Our attention is drawn by the two fairly intense peaks at 425 and 508 K with values of the activation energy $\langle E\rangle(T)$ equal to about 0.75 and 1.00 eV, respectively. These TSEE peaks (Fig. 5) are comparable to those for the case of linear heating (Fig. 1) and are not displayed on the TSL curve. The exo-

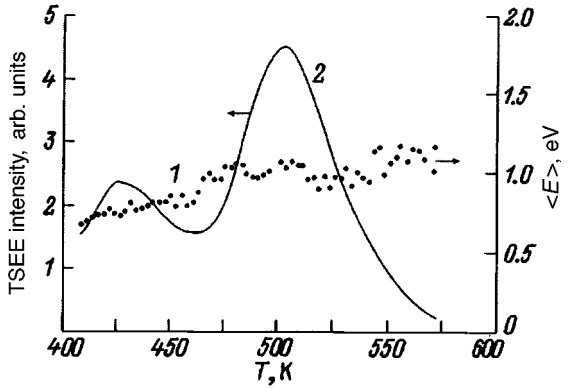

FIG. 5. Temperature dependence of the mean thermal activation energy (*1*) and TSEE intensity (*2*) of LiB_3O_5 crystals during linear–oscillatory heating following electron bombardment at 300 K.

emission response during linear–oscillatory heating at 400–600 K also mimics the temperature oscillations.

DISCUSSION

In Refs. 2 and 3 we developed a model of field fluctuation processes (FFPs). The FFPs in this model include various thermally stimulated physical processes (ionic processes, adsorption, etc.), which influence the delocalization, migration, and relaxation of electron excitations. The features observed for LiB_3O_5 at 100–140 K (the elevated value of the mean activation energy and the nonisothermally elevated, electron and photon emission response, which is in phase with the temperature oscillations) are qualitatively consistent with those found in Refs. 2 and 3 for a FFP that lowers the potential barrier of the trapping centers. At the same time, the features in the kinetics of the thermally stimulated processes cited can be attributed only to surface processes (for example, adsorption phenomena, electric discharges through molecules of the residual atmosphere adsorbed on the crystal surface, etc.). In fact, the spontaneous emission at 100–140 K is displayed on both the TSL and TSEE curves. It is seen from Fig. 1 that about half of the partial exoemission and light yields, whose temperature dependence corresponds with the annealing of paramagnetic B^{2+} centers and thermobleaching in the band at 306 nm,[4] is emitted in this region. In Ref. 10 we advanced a model of the B^{2+} center in crystalline LiB_3O_5 in the form of an interstitial boron ion, which traps an electron upon irradiation, on the basis of quantum-chemical cluster calculations. The results of the calculation in Ref. 10 provide evidence that several interstitial sites for boron ions, which are almost equivalent with respect to the Madelung potential and the energy position of the local level of the additional electron, can exist in the LiB_3O_5 lattice. Such an energy distribution of the B^{2+} trapping centers can be one of the reasons why the plots of the temperature dependence for the annealing of the paramagnetic B^{2+} centers, the thermobleaching in the band at 306 nm, and the buildup of the exoelectron and light sums (Fig. 1) are not elementary curves. In addition, the results in Ref. 11 indicate that the thermally stimulated processes in LiB_3O_5 in the temperature range 100–180 K are accompanied by relaxation of the radiation-induced electric charge. This thermally stimulated process can cause significant variation of the local electric field and is perfectly suitable for the role of the postulated fluctuation-induced process that lowers the height of the potential barrier of the trapping centers.

The TSL peak at 210 K, which we previously attributed to the thermal decay of hole O^- trapping centers,[4] is characterized by a smooth drop in the value of $\langle E\rangle(T)$ from 0.40 ($T=180$ K) to 0.29 eV ($T=240$ K) and is not displayed on the TSEE curve (Fig. 4). Several plausible reasons for such behavior of $\langle E\rangle(T)$ can be cited. First, thermal quenching of the intrinsic luminescence of LiB_3O_5, which is excited both by photons with energies exceeding E_g and in recombination processes (for example, TSL), begins in the temperature range corresponding to this TSL peak.[4] According to the theory of thermally stimulated processes in the condensed state of matter,[12] the decrease in the recombination luminescence yield in the region of the thermally stimulated delocalization of charge carriers leads to an apparent decrease in the observed mean thermal activation energy. Second, in Ref. 13 we showed that the TSL peak of LiB_3O_5 crystals at 180–240 K is the result of the superposition of at least two overlapping elementary peaks. The contour and the temperature position of this TSL peak are determined by the relative positions and the intensity ratio of its elementary components. According to Eq. (1), the value of the mean thermal activation energy $\langle E\rangle(T)$ is determined not only by the activation energies of the elementary components, but also by the ratio of their intensities. For this reason, the plot of $\langle E\rangle(T)$ in the range 180–240 K (Fig. 4) can reflect the variation of the contributions of the elementary processes as the temperature rises. In both cases the features of the thermally activated recombination process in the temperature range 180–240 K in LiB_3O_5 crystals cannot be associated with any manifestations of the suprathermal emission of electrons and photons.

CONCLUSIONS

Thus, the thermally stimulated exoelectron emission of lithium triborate crystals has been discovered for the first time and studied. The thermally stimulated emission of electrons and photons from LiB_3O_5 following various radiation treatments have been measured simultaneously and compared.

Two new trapping centers, whose thermal annealing (at 450 and 515 K) is accompanied by the emission of electrons, have been discovered. There are no manifestations of these centers on the TSL curve, possibly because the luminescence of the crystal is thermally quenched at those temperatures.

The experimental results provide a basis for discussing the presence of a field fluctuation process in LiB_3O_5 crystals in the temperature range 100–140 K; this process, which involves a lowering of the potential barrier, is presumably due to thermally stimulated relaxation of the radiation-induced charge.

A significant dependence of the quantitative parameters of the thermally stimulated relaxation processes on the type, energy, and dose of the ionizing radiation has been discovered.

We thank V. S. Kortov for his support, A. Yu. Kuznetsov and A. E. Monakhov for their assistance and participation in this work, and V. A. Maslov for supplying the LiB_3O_5 crystals.

[1] I. N. Ogorodnikov, V. I. Kirpa, and A. V. Kruzhalov, Zh. Tekh. Fiz. **61**(7), 67 (1991) [Sov. Phys. Tech. Phys. **36**, 759 (1991)].
[2] I. N. Ogorodnikov, V. I. Kirpa, and A. V. Kruzhalov, Zh. Tekh. Fiz. **63**(5), 70 (1993) [Tech. Phys. **38**, 404 (1993)].
[3] I. N. Ogorodnikov and A. V. Kruzhalov, Zh. Tekh. Fiz. **65**(6), 64 (1995) [Tech. Phys. **40**, 556 (1995)].
[4] I. N. Ogorodnikov, A. Yu. Kuznetsov, A. V. Kruzhalov, and V. A. Maslov, Radiat. Meas. **24**, 423 (1995).
[5] A. Yu. Kuznetsov, V. Yu. Ivanov, I. N. Ogorodnikov *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **359**, 339 (1995).
[6] V. A. Maslov, L. A. Olkhovaya, V. V. Osiko, and E. A. Shcherbakov, in *ICCG-10. 10th International Conference on Crystal Growth*, San Diego, California, 1992, p. 11.
[7] H. König and R. Hoppe, Z. Anorg. Allg. Chem. **439**, 71 (1978).
[8] S. F. Radaev, E. A. Genkina, V. A. Lomonov *et al.*, Kristallografiya **36**, 1419 (1991) [Sov. Phys. Crystallogr. **36**, 803 (1991)].

[9] V. I. Kirpa, A. S. Kuz'minykh, and V. V. Popov, in *Radiation-Stimulated Phenomena in Solids. An Interinstitute Collection* [in Russian], Sverdlovsk (1983), pp. 33–38.
[10] A. Yu. Kuznetsov, A. B. Sobolev, I. N. Ogorodnikov, and A. V. Kruzhalov, Fiz. Tverd. Tela (St. Petersburg) **36**, 3530 (1994) [Phys. Solid State **36**, 1876 (1994)].
[11] I. N. Ogorodnikov, V. Yu. Ivanov, A. Yu. Kuznetsov *et al.*, Pis'ma Zh. Tekh. Fiz. **19**(11), 1 (1993) [Tech. Phys. Lett. **19**, 325 (1993)].
[12] R. Chen and Y. Kirsh, *Analysis of Thermally Stimulated Processes*, Pergamon Press, Oxford (1981), 360 p.
[13] I. N. Ogorodnikov, A. Yu. Kuznetsov, and A. V. Porotnikov, Pis'ma Zh. Tekh. Fiz. **20**(13), 66 (1994) [Tech. Phys. Lett. **20**, 549 (1994)].

Translated by P. Shelnitz

Simultaneous self-focusing of two laser beams in a subthreshold coherent population trapping regime

I. V. Kazinets, B. G. Matisov and A. Yu. Snegirev

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
(Submitted November 19, 1996)
Zh. Tekh. Fiz. **67**, 126–129 (July 1997)

The propagation of two-component laser radiation in a medium consisting of atoms with a Λ level scheme is investigated. The simultaneous self-focusing of two beams is considered. The main features of this phenomenon are: 1) lowering of the self-focusing threshold by several orders of magnitude in comparison with the known case of saturation of a transition in a two-level atom; 2) a strong dependence of the character of the propagation of the radiation on the difference between the detunings of the two frequency components of the field from resonance, which is associated with fulfillment of the two-photon resonance condition. © *1997 American Institute of Physics.* [S1063-7842(97)02407-0]

INTRODUCTION

The self-focusing of light beams in nonlinear media was discovered quite long ago.[1–3] Being a very general nonlinear-wave phenomenon, it occurs not only in optics, but also in acoustics.[4–6] The results of an investigation in the model of a two-level atom with a nonlinear dielectric constant, for example, were generalized in Refs. 7 and 8. In recent years interest has been sparked in self-focusing, because the use of multilevel media to observe this phenomena has revealed new important features: lowering of the nonlinearity threshold[9–11] and the practically complete absence of absorption in a narrow frequency range where self-focusing is observed.[12] These effects are attributed to the possibility of quantum interference between different excitation channels in multilevel media. Consequently, a new stationary superposition state, whose decay rate is considerably smaller than the spontaneous relaxation rate of the excited levels can appear under certain conditions. Furthermore, this causes saturation to appear in the system at far smaller intensities of the laser beams than in the case of a two-level system owing to the phenomenon of coherent population trapping (for further details, see Ref. 13). According to theory, self-focusing should be eliminated on passage to the superthreshold regime, because bleaching of the medium occurs, i.e., the medium scarcely absorbs. Experimental support for this deduction was presented in Ref. 14. On the other hand, in the subthreshold regime self-focusing depends strongly on the difference between the detunings of the two spectral components of the field. This dependence was investigated in Ref. 15. However, the case in which only one beam is focused was considered in that work, while the intensity of the other beam was assumed to be constant along the cell. In the present work we examine the case of the simultaneous focusing of two beams and consider the elimination of the focusing of two beams when the difference between the detunings increases.

BASIC EQUATIONS

Let us assume that laser radiation containing two spectral components with the frequencies ω_1 and ω_2 (the detunings from resonance equal Ω_1 and Ω_2, respectively) and the complex amplitudes E_1 and E_2 propagates along the z axis through a medium consisting of atoms with a Λ level scheme.

We assume that wave 1 interacts with an atom only in the $|1\rangle - |3\rangle$ transition and that wave 2 interacts in the $|2\rangle - |3\rangle$ transition.

The abridged Maxwell equations with consideration of the finite transverse dimensions of the laser beam have the form[15]

$$2ik_m \frac{\partial}{\partial z} E_m + \Delta_\perp E_m = -4\pi N k_m^2 d_{m3} \rho_{3m}, \quad m = 1, 2, \quad (1)$$

where $k_m = \omega_m / c$ denotes the wave numbers, N is the concentration of atoms, d_{m3} is the dipole moment matrix element of the atom, and $\Delta_\perp$ is the two-dimensional transverse Laplacian

$$\Delta_\perp = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} \quad (2)$$

in Cartesian and cylindrical coordinates, respectively. The off-diagonal elements ρ_{3m} of the atomic density matrix are found from the stationary solution of the system of kinetic equations.[13] The expressions for them are presented in the Appendix to Ref. 15. We assume that the lower long-lived sublevels $|1\rangle$ and $|2\rangle$ either differ only with respect to the magnetic quantum number or belong to different components of the hyperfine structure, so that $k_1 \cong k_2 \equiv k$ and we can neglect the Doppler broadening of the two-quantum $|1\rangle - |2\rangle$ transition.

Equations (1) must be supplemented by the initial conditions

 1063-7842/97/070837-04$10.00

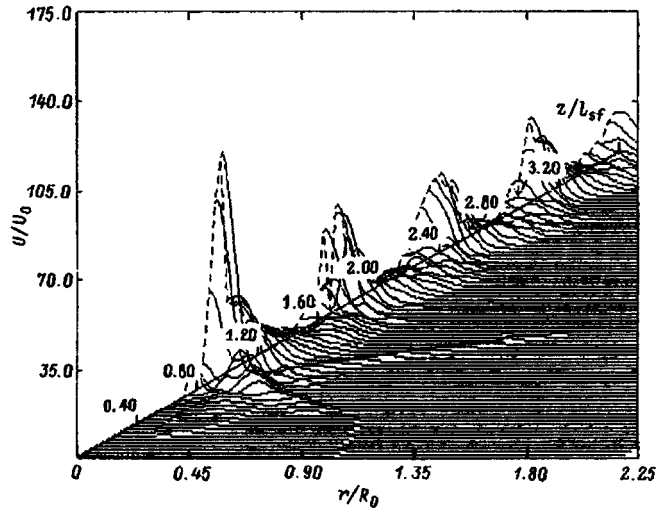

FIG. 1. Dependence of the intensity U of the first beam on the depth z and radius r for the following parameters of the radiation and the medium: $\gamma_1=\gamma_2=10^7$ s^{-1}, $\Gamma=0.5\times10^5$ s^{-1}, $\Gamma_{11}=10^4$ s^{-1}, $\Omega_1-\Omega_2=0$, $R_0=3\times10^{-3}$ cm, $k=10^5$ cm^{-1}, $N=10^{14}$ cm^{-3}, $\frac{1}{2}(\Omega_1+\Omega_2)=10^9$ s^{-1}, $g_1=g_2=0.7\times10^8$ s^{-1}, $V_T=0$, where γ_m is the spontaneous relaxation rate along the $|3\rangle-|m\rangle$ channel, Γ_{11} is the longitudinal relaxation rate between levels 1 and 2, and g_m is the Rabi frequency ($m=1,2$).

$$E_m|_{z=0}=E_{m0}\exp\left(-\frac{r^2}{2R_{m0}^2}\right),\quad m=1,2 \tag{3}$$

and the boundary conditions

$$E_m|_{r=0}\ \text{is bound};\quad E_m|_{r\to\infty}\to 0,\quad m=1,2$$

(for simplicity, we assume axial symmetry, i.e., that $\partial E_m/\partial\varphi=0$).

Let us first consider the case of two-photon resonance:

$$\Omega_1=\Omega_2\equiv\Omega. \tag{4}$$

Since under the assumptions adopted above the frequencies and dipole moments of both optical transitions in the Λ atom differ relatively weakly, we set $\gamma_1\cong\gamma_2\equiv\gamma$. When $\Omega\gg kV_T$, where V_T^2 is the ensemble-averaged square of the z projection of the velocity of the atoms, we obtain propagation equations in the form

$$2ik\frac{\partial E_m}{\partial z}+\Delta_\perp E_m=-2ik^2\frac{3\pi}{2}\nu\times\frac{E_m(1+i\eta)}{(1+\eta^2)(1+(|E_1|^2+|E_2|^2)/E_{\rm nl}^2)},\quad m=1,2, \tag{5}$$

where $\nu=Nk^{-3}$ and $\eta=\Omega/\gamma$ is the dimensionless detuning.

To within a numerical factor of order unity, the square $E_{\rm nl}^2$ of the amplitude of the light field which saturates the nonlinearity equals

$$E_{\rm nl}^2\cong\hbar\Gamma k^3(1+\eta^2). \tag{6}$$

The condition $|E_m|^2\leqslant E_{\rm nl}^2$ ($m=1,2$) means that the optical excitation rate is smaller than the low-frequency coherence decay rate ρ_{12}, i.e., the radiated intensity is below the threshold for the appearance of coherent population trapping: $U\leqslant U_c[1+(\Omega/\gamma)^2]$. Here $U_c=U_{\rm sat}\Gamma/(2\gamma)$, where $U_{\rm sat}$ is the saturation intensity of the optical transition.

A qualitative analysis of the solution of Eqs. (5) can be performed on the basis of the theory developed in Ref. 8. It should first be noted that self-focusing is possible only for a positive detuning $\eta>0$. If the squares of the field strengths of both spectral components at the entrance to the medium do not exceed $E_{\rm nl}^2$, the lower self-focusing threshold (which corresponds to the so-called critical wave amplitude $E_{\rm cr}$) can be estimated as

$$E_{\rm cr}^2=\frac{\hbar\Gamma k^3(1+\eta^2)^2}{\eta\nu(kR_{m0})^2}. \tag{7}$$

On the characteristic spatial scale (the self-focusing length)

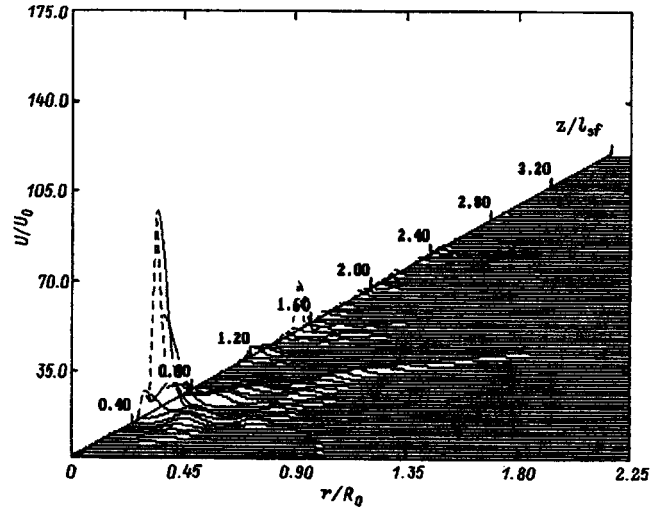

FIG. 2. Same as in Fig. 1, except that $\Gamma=10^5$ s^{-1}.

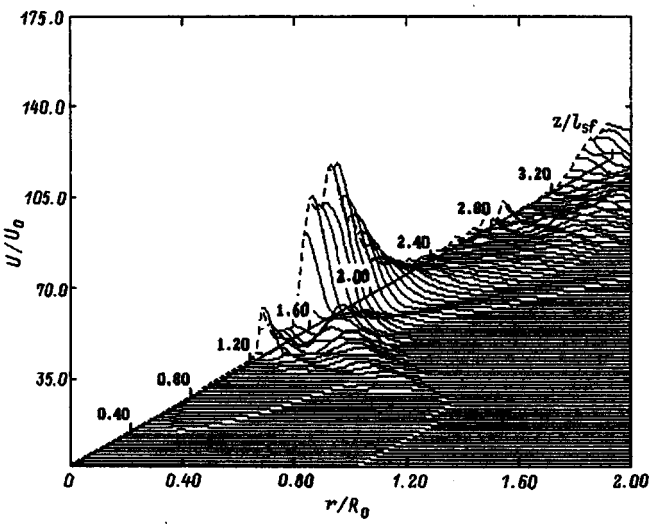

FIG. 3. Same as in Fig. 1, except that $\Gamma=10^5$ s^{-1} and $g_1=g_2=1.4\times10^8$ s^{-1}.

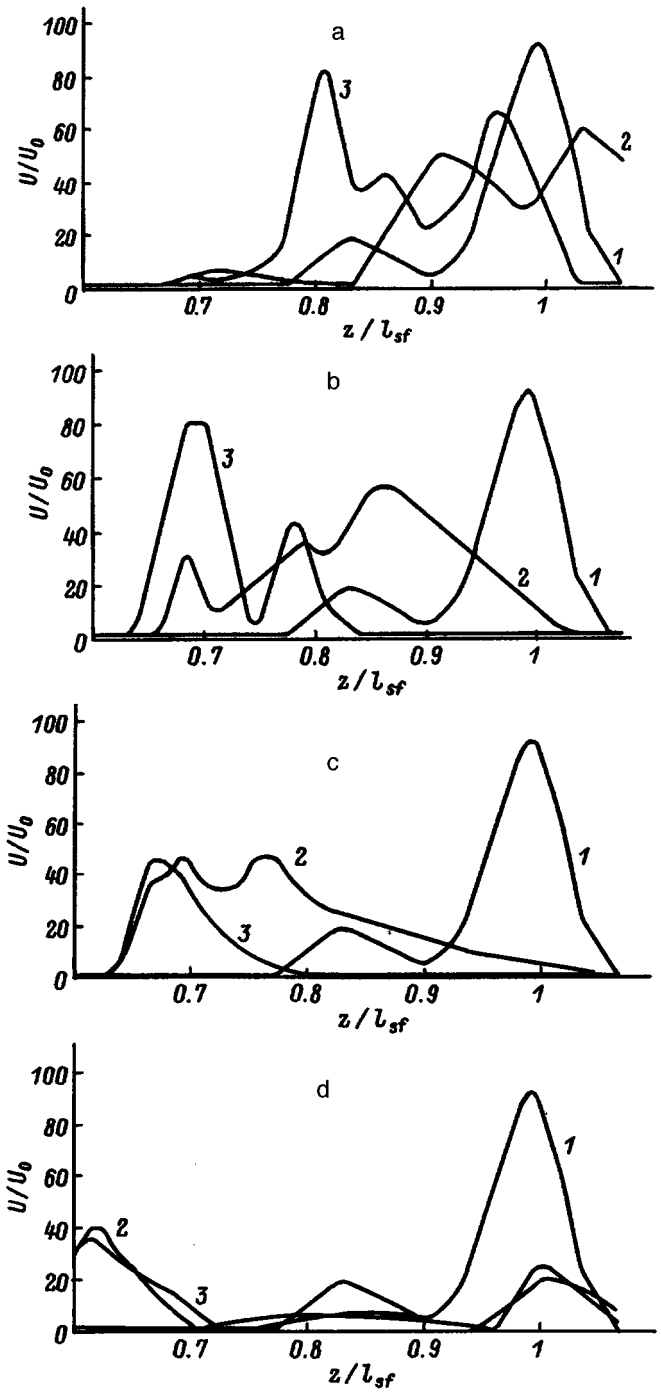

FIG. 4. Dependence of the intensity of the radiation at the beam axis on the depth z in the region of the first focus: *2* — first beam, *3* — second beam. The parameters are the same as in Fig. 1 with the exceptions: a — $\Omega_1-\Omega_2=1\times10^6$ s^{-1}, b — $\Omega_1-\Omega_2=2\times10^6$ s^{-1}, c — $\Omega_1-\Omega_2=6\times10^6$ s^{-1}, d — $\Omega_1-\Omega_2=1\times10^7$ s^{-1}. Curve *1* refers to the case $\Omega_1-\Omega_2=0$.

$$l_{\mathrm{sf},m}=R_{m0}\left(\frac{(1+\eta^2)E_{\mathrm{nl}}^2}{\eta\nu E_{m0}^2}\right)^{\frac{1}{2}}; \tag{8}$$

practically all the intensity of the mth spectral component is concentrated in a narrow region near the beam axis, i.e., a focus appears.

Let us consider the restrictions on the initial radius of the beams. First, the self-focusing length must be much smaller

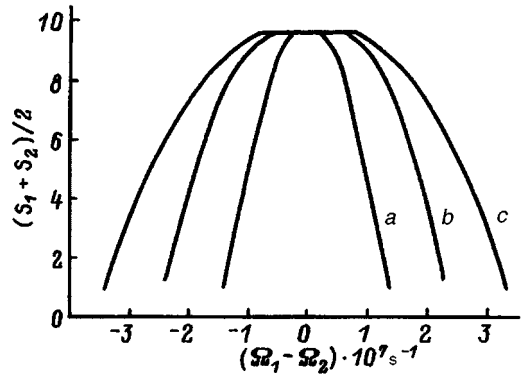

FIG. 5. Dependence of the mean area under the plot of the intensity at the beam axis in the region of the first focus on the difference between the detunings: a — $\Gamma=0.5\times10^5$, b — 1×10^5, c — $\Gamma=2\times10^5$ s^{-1}.

than the characteristic scale for absorption, i.e., $l_{\mathrm{sf},m}\ll(1+\eta^2)/(k\nu)$. A second restriction follows from the subthreshold character of the phenomenon with respect to coherent population trapping: $E_{m0}^2\ll E_{\mathrm{nl}}^2$. These two conditions can be written together in the form

$$\left(\frac{\eta}{\nu}\right)^{\frac{1}{2}}\ll kR_{m0}\ll\eta\left(\frac{\eta}{\nu}\right)^{\frac{1}{2}}. \tag{9}$$

If the radiation belongs to the visible range and the gas is at room temperature, Eq. (9) gives the numerical estimate

$$50\left(\frac{p^*}{p}\frac{\Omega}{\Omega^*}\right)^{\frac{1}{2}}\ll kR_{m0}\ll5\times10^3\left[\frac{p^*}{p}\left(\frac{\Omega}{\Omega^*}\right)^3\right]^{\frac{1}{2}}, \tag{10}$$

where $p^*=10^{-3}$ Torr, and $\Omega^*=10^9$ s^{-1}.

Taking the values $p\cong0.03$ Torr, $\Omega=10^{10}$ s^{-1}, $R_{m0}\cong0.1$ cm, and $\Gamma=10^5$ s^{-1}, we find that the critical field strength corresponds to an intensity of the order of 0.1mW/cm^2. If the intensity of the beams on the beam axis at the entrance to the medium is about 30mW/cm^2, the radiation is focused at a length of about 50 cm.

Thus, it is seen that, all other conditions being equal, the minimum intensity value at which self-focusing is possible in the subthreshold coherent population trapping regime is lower than in the case of ordinary saturation of the optical transition by a factor equal to $\gamma/\Gamma\cong10^2-10^4$. There is a simple physical explanation for this:[13] in a three-level Λ medium the atoms are pumped into the nonabsorbing state $|\Psi_{NC}\rangle$, from which they escape owing to the finite decay rate of the coherent superposition of the two low-lying atomic states $|1\rangle$ and $|2\rangle$.

RESULTS OF THE NUMERICAL CALCULATIONS

We present the results of some numerical calculations of the self-focusing in a three-level Λ medium that were performed for the parameters $R_{20}\approx R_{10}=R_0$ and $E_{20}^2\approx E_{10}^2=E_0^2$, i.e., when both beams are focused simultaneously. Figure 1 presents the dependence of the intensity of the first beam on the coordinates r and z. Natural conversion to dimensionless quantities is employed: $U(r,z)$ is measured in units of $U_0\equiv U(0,0)=|E_0|^2$, r is measured in units of

R_0, and z is measured in units of l_{sf}. As we see, in the first focus at $z \cong l_{sf}$ the intensity increases by two orders of magnitude. Subsequent foci are also observed. It is important to note here that the second beam is focused like the first; therefore, the total intensity of the two beams is twice as great as in Fig. 1. When the relaxation rate of the coherence between states $|1\rangle$ and $|2\rangle$ is doubled, the foci, beginning with the second, vanish (Fig. 2). If along with the doubling of Γ, the Rabi frequencies of both beams are increased by a factor of $\sqrt{2}$, the distribution of the foci remains unchanged (Fig. 1). When the Rabi frequencies are increased further by a factor of $\sqrt{2}$ (Fig. 3), the distribution of the foci is altered even with consideration of the change in the scale along the z axis [Eq. (8)].

The frequency dependence of the intensity in the first focus has a symmetric dependence relative to the sign of the difference between the detunings. Figure 4 presents plots of the dependence of the intensity of the fields on the beam axis for various differences between the detunings. The degree of elimination of the self-focusing can be estimated from the mean area under the plot of the intensity on the beam axis in the region of the first focus. It is seen from Fig. 5 that the dependence of the mean area on the difference between the detunings has a resonance form and is characterized by the half-width $\delta\omega_{sf}$ at half-maximum. A numerical calculation gives $\delta\omega_{sf} \approx 200\Gamma$.

Thus, the following conclusions can be drawn on the basis of the numerical calculations: 1) the simultaneous self-focusing of two beams is possible in the subthreshold coherent population trapping regime at a focusing threshold that is lower by a factor γ/Γ than in the case of saturation of the optical transition;[8] 2) in the case of simultaneous self-focusing: a) the focusing properties of the medium are more pronounced (in the regions of the foci the intensities are more than twice as high as when one beam is focused[15]); b) the required structure of the foci can be obtained by varying the intensities and the detunings of the beams; c) there is a sharp dependence of the spatial distribution of the light intensity on the difference between the detunings of the two laser fields from resonance [when the detuning of one of the beams deviates from the mean detuning by a few percent, the self-focusing vanishes (Fig. 5)].

This work was partially supported by Grant No. 5-5.5-139 from the State Committee of the Russian Federation for Higher Education.

[1] V. I. Talanov, Pis'ma Zh. Éksp. Teor. Fiz. **2**, 218 (1965) [JETP Lett. **2**, 138 (1965)].
[2] V. N. Lugovoĭ, Dokl. Akad. Nauk SSSR **179**, 58 (1967) [Sov. Phys. Dokl. **12**, 866 (1968)].
[3] P. I. Kelley, Phys. Rev. Lett. **15**, 1005 (1965).
[4] G. A. Askar'yan, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 144 (1966) [JETP Lett. **4**, 99 (1966)].
[5] G. A. Askar'yan and V. I. Pustovoĭt, Zh. Éksp. Teor. Fiz. **58**, 647 (1970) [Sov. Phys. JETP **31**, 346 (1970)].
[6] G. A. Askar'yan, Pis'ma Zh. Éksp. Teor. Fiz. **13**, 395 (1971) [JETP Lett. **13**, 283 (1971)].
[7] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon, Oxford (1984).
[8] V. N. Lugovoĭ and A. M. Prokhorov, Usp. Fiz. Nauk **111**(2), 203 (1973) [Sov. Phys. Usp. **16**, 658 (1974)].
[9] B. G. Matisov and I. E. Mazets, Pis'ma Zh. Tekh. Fiz. **20**(4), 16 (1994) [Tech. Phys. Lett. **20**, 140 (1994)].
[10] B. Matisov and L. Windholz, in *Abstracts of the 5th European Quantum Electronics Conference*, Amsterdam, 1994, p. 85.
[11] D. G. Akopyan and A. Zh. Muradyan, Opt. Orient. At. Mol. Sb. Fiz.-Tekh. Inst. (2) 162 (1990).
[12] U. Rathe, M. Fleischhauer, S. Y. Zhu *et al.*, Phys. Rev. A. **47**, 4994 (1993).
[13] B. D. Agap'ev, M. B. Gornyĭ, and B. G. Matisov, Usp. Fiz Nauk **163**(9), 1 (1993) [Phys. Usp. **36**, 763 (1993)].
[14] M. Jain, A. J. Merrian, A. Kasapi *et al.*, Phys. Rev. Lett. **75**, 4385 (1995).
[15] B. G. Matisov, I. E. Mazets, and A.Yu. Snegirev, Zh. Tekh. Fiz. **66**(7), 124 (1996) [Tech. Phys. **41**, 701 (1996)].

Translated by P. Shelnitz

Pulse characteristics of $Hg_{0.8}Cd_{0.2}Te$ n^+-p junctions

I. S. Virt

I. Franko Drogobych State Teachers Institute, 293720 Drogobych, Ukraine
(Submitted February 19, 1996)
Zh. Tekh. Fiz. **67**, 130–133 (July 1997)

The pulse characteristics of $Hg_{0.8}Cd_{0.2}Te$ n^+-p junctions are investigated. It is shown that the shape of the voltage pulse appearing in a junction on passage of a forward (reverse) current is determined by the recombination (generation) of nonequilibrium electrons in the hole region. An increase in the current pulse causes the appearance of an electric field, which draws electrons into the interior of the base region, and leads to variation of their lifetime because of the complex structure of the n^+-p junction. © *1997 American Institute of Physics.* [S1063-7842(97)02507-5]

1. INTRODUCTION

The pulse characteristics of $n-p$ junctions play a significant role in determining the speed of semiconductor photodiodes. Their speed depends on various factors: the doping of the base regions of the particular semiconductor, the dopant concentration, the capacitance of the junction, and the defect density in the near-junction region.[1] A special place belongs to the defects appearing during the industrial fabrication of $n-p$ junctions. This is important for the semiconductor material $Hg_{0.8}Cd_{0.2}Te$, in which intrinsic radiation defects of both the point and extended types readily form during the fabrication of $n-p$ junctions by ion implantation. The layered structure of the $n-p$ junctions is then fairly complicated.[2,3] The transient processes in n^+-p structures fabricated by ion implantation were investigated in Ref. 4.

In the present work we investigated the pulse characteristics of $Hg_{0.8}Cd_{0.2}Te$ n^+-p junctions as a function of the injection level. The n^+ region was created by implanting B^+ ions with an energy of 100 keV in a substrate with p-type conduction. The measurements were performed at $T=77$ K on the passage of current pulses in both the forward and reverse directions, and the results were displayed on an S8-13 oscillograph. The pulse duration was selected in the range between 10^{-5} and 10^{-7} s.

Oscillograms of the voltage on an n^+-p junction under forward and reverse biases are presented in Figs. 1a and 1b. The pulse which has passed through the sample exhibits some distortion on a background of the oscillogram. When a pulse is passed in the forward direction, minority charge carriers are injected through the junction, and they recombine in the base regions. Of course, when Auger recombination dominates, the voltage relaxation time in a given temperature range after completion of the current pulse is determined by the longer-lived electrons in the less doped hole region of the semiconductor. The falling edge of the pulse initially has a linear time dependence, which corresponds to a high injection level.[5] This is true for injection levels at which $qU \gtrsim kT$ (q is the charge of the electron, k is Boltzmann's constant, and U is the voltage drop on the $n-p$ junction). The concentration of nonequilibrium electrons in the p region is described by the expression

$$\Delta n_p = C(e^{\frac{qU}{kT}} - 1). \tag{1}$$

In this case for a high injection level ($qU/kT \gtrsim 1$; $e^{qU/kT} \gg 1$) we have

$$\Delta n_p \approx C e^{\frac{qU}{kT}}. \tag{2}$$

On the other hand, the relaxation of Δn_p follows the law

$$\Delta n_p = C' e^{-\frac{t}{\tau_n}}, \tag{3}$$

where C and C' are constants, and τ_n is the lifetime of the nonequilibrium electrons in the p region.

Comparing (2) and (3), we find that the time dependence of the postinjection voltage at a high injection level has the form

$$U(t) = -\frac{kT}{q}\frac{t}{\tau_n}. \tag{4}$$

At large observation times the linear decay becomes exponential. The changeover occurs when the voltage level reaches $U(t) \lesssim (kT)/q$. In this case in Eq. (1) we have

$$e^{\frac{qU}{kT}} \approx 1 + \frac{qU}{kT},$$

and Eq. (4) transforms into the expression

$$\ln U(t) = \ln\frac{kT}{q} - \frac{t}{\tau_n}. \tag{5}$$

The decay of the postinjection emf is plotted on a logarithmic scale in Fig. 2a. The lifetime of the minority charge carriers was determined on both the linear portion from Eq. (4) and the exponential portion from Eq. (5). When $qU/kT \gg 1$, an attractive electric field appears in the base region and must be taken into account in determining τ_n at ultrahigh injection levels:[5]

$$U(t) = \frac{kT}{q}\frac{t}{\tau_n}\frac{2b}{b+1}, \tag{6}$$

where $b = \mu_n/\mu_p$ is the ratio between the electron and hole mobilities.

Since $b \gg 1$ in $Hg_{0.8}Cd_{0.2}Te$, Eq. (6) can be simplified:

 1063-7842/97/070841-04$10.00

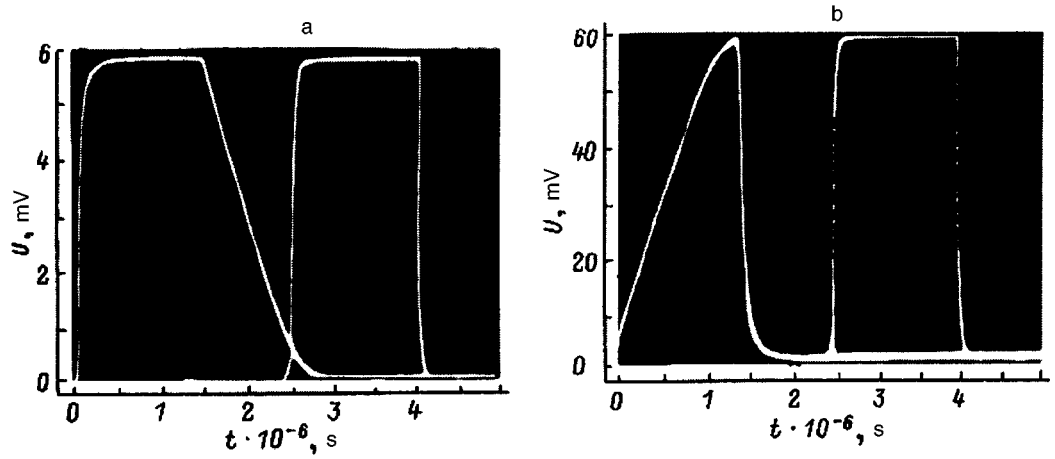

FIG. 1. Oscillograms of the voltage on a n^+-p junction on the forward (a) and reverse (b) passage of a pulse (the form of the rectangular input pulse is shown for comparison).

$$U(t)=\frac{kT}{q}\frac{2t}{\tau_n}. \tag{7}$$

The linear slope of the falling edge varies only slightly with time, attesting to the insignificance of the diffusion processes, i.e., the variation of the coordinate distribution of the minority charge carriers is determined mainly by their recombination.

The reverse pulse (when the n^+-p junction is subjected to a reverse bias) consists of two portions (Fig. 1a): an exponential portion at short times $t<\tau_n$ and a linear portion at $t>\tau_n$. The rising edge of the voltage pulse is governed by the generation of electrons in the p region, and the falling edge is determined by Maxwell relaxation of the electrons in the n^+ region during a time $\tau_m \ll \tau_n$. The generation of electron at small biases ($qU/kT \ll 1$) is described by the expression

$$U(t)=\frac{kT}{q}(1-e^{-\frac{t}{\tau_n}}). \tag{8}$$

When a voltage $U(t) \gtrsim (kT)/q$ is achieved, for the case of generation in Eq. (3) we have

$$\Delta n_p = C'(1-e^{-\frac{t}{\tau_n}}). \tag{9}$$

Thus, since

$$-\Delta n_p = C(e^{-\frac{qU}{kT}}-1) \tag{10}$$

at these bias levels,

$$U(t)=\frac{kT}{q}\frac{t}{\tau_n}. \tag{11}$$

Logarithmic plots of the rise of the voltage on an n^+-p junction for the reverse passage of a pulse are presented in Fig. 2b in $\ln(U_{\max}-U)-t$ coordinates, where $U_{\max}$ is the maximum value of the voltage. The values of τ_n determined on the linear and exponential portions were approximately equal to the values determined from the injection pulses. At large reverse biases ($|U|>0.3$ V) the dependence of the rising edge of the pulse is exponential over almost the entire segment. At these biases $e^{-qU/kT} \ll 1$ and $\Delta n_p \sim C \sim J_{\text{rev}}$. Since the tunneling current is dominant at such biases, i.e., $J_{\text{rev}} \sim U^3$ (Ref. 6), the voltage across the n^+-p junction varies according to the exponential law

$$U(t) \sim (1-e^{-\frac{t}{\tau_n}})^{1/3}. \tag{12}$$

Plots of the dependence of the lifetime of the minority charge carriers determined from the linear and exponential portions of the postinjection falling edge on the magnitude of the pulse are presented in Fig. 3. At small injection levels the

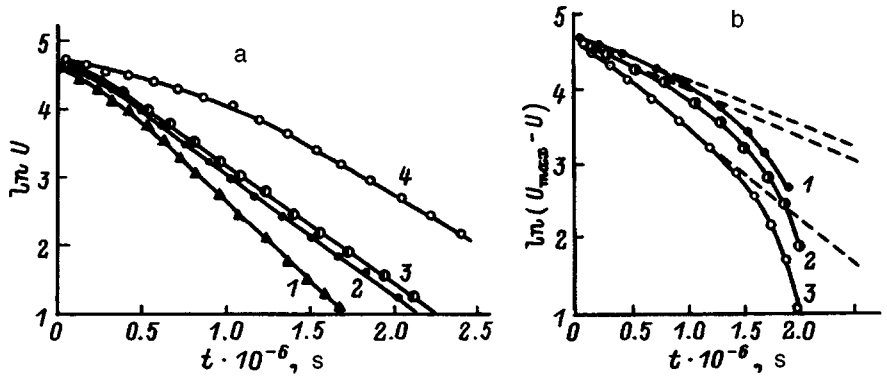

FIG. 2. Time dependence of the voltage on a n^+-p junction after completion of a forward current pulse (a) for $U_{\text{for}}=3$ (*1*), 6 (*2*), 12 (*3*), and 30 mV (*4*) and after the supply of a reverse current (b) for $U_{\text{rev}}=12$ (*1*), 60 (*2*), and 300 mV (*3*).

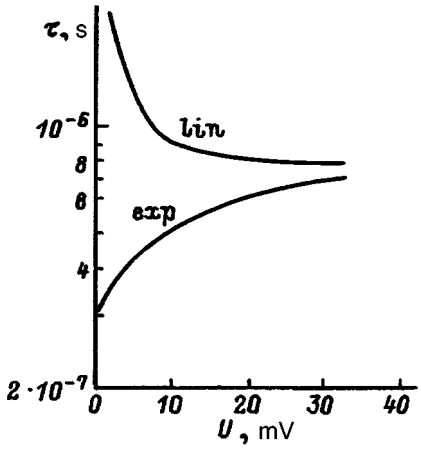

FIG. 3. Dependence of the lifetime of the electrons on the amplitude of the injection pulse calculated for the linear (lin) and exponential (exp) portions of the postinjection falling edge.

values of τ_n differ strongly. This difference is probably associated with the layered structure of the n^+-p junction. At small ($qU \ll kT$) and high ($qU \gg kT$) injection levels the variation of the electric field in the junction can be disregarded, and it can be assumed that recombination occurs at a distance approximately equal to the diffusion length of the nonequilibrium electrons L_n in the p region. The charges recombine in the high-resistivity n_i region,[2,3] whose dimensions can reach 10 μm, with a time $\tau_n \approx 2\times 10^{-6}$ s. At small injection levels recombination takes place on the interface between the n^+ and n_i regions with a high defect density (of a radiation nature) and is characterized by a comparatively low value of the electron lifetime τ_n.

At ultrahigh injection levels the electric field appearing because of the high electron concentration gradient in the p region must be taken into account. The presence of this electric field means that electrons are drawn into the interior of the hole base to a distance $l = L_n + q\mu_n|\varepsilon|$, where ε is the field strength in the n^+-p junction. This distance amounts to hundreds of microns for the known characteristics of electrons and a given composition x. Recombination takes place mainly in the interior of the p region with a low defect density, as is observed experimentally when $U > 0.15$ V.

It should be noted that the determination of the lifetime of the minority charge carriers is influenced by the homogeneity of the n^+-p junction. For example, the presence of shunting channels, which alter the current–voltage characteristics of the junction, is also manifested in the pulse characteristics. When such channels are present, the reverse branches of the characteristics can be described by a linear dependence with a characteristic resistance R_{sh} (Fig. 4a, curve *3*). Upon injection both the passage and recombination of electrons take place mainly on inhomogeneities. Accordingly, in the case of Auger recombination, which predominates in the impurity region of $Hg_{0.8}Cd_{0.2}Te$ crystals, the low-resistivity current-passing regions reduce the lifetime of the minority charge carriers (Fig. 4b, curve *3*).

Thus, in $Hg_{0.8}Cd_{0.2}Te$ n^+-p junctions the passage of a current pulse is associated with the recombination (during forward passage) and generation (during reverse passage) of

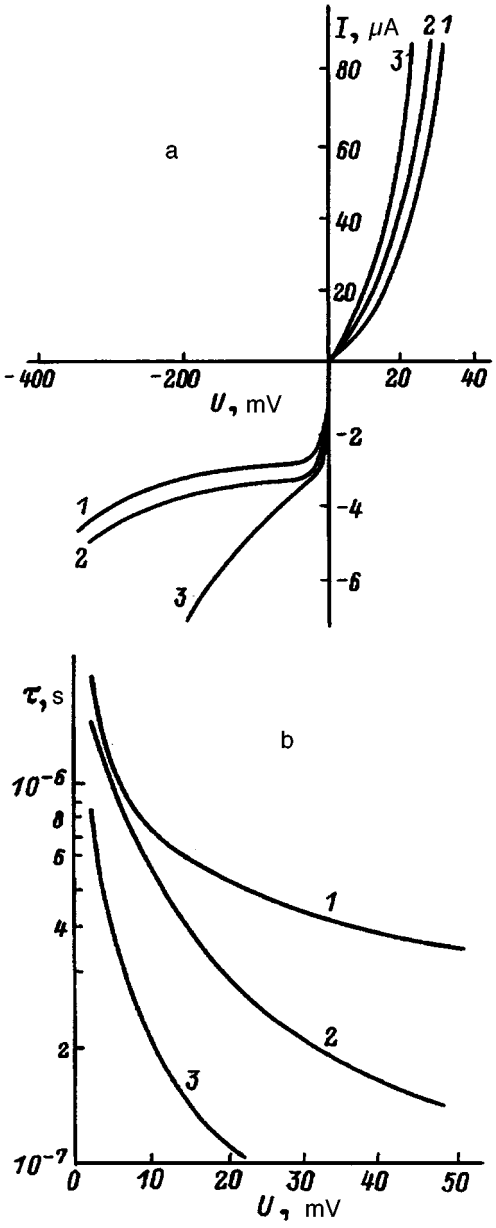

FIG. 4. Current–voltage characteristics of various (*1–3*) n^+-p junctions (a) and dependence of their relaxation time on the injection level (b).

electrons in the hole base region. The dependence of the relaxation time of a pulse on its amplitude shows that the lifetime of the minority charge carriers is influenced by the complex structure of the n^+-p junction, as well as by the appearance of an electric field associated with their concentration gradient.

We thank S. V. Belotelov for supplying the experimental samples.

[1] É. A. Shevtsov and M. E. Belkin, *Photodetection Devices for Fiber-Optic Communication Systems* [in Russian], Radio i Svyaz', Moscow (1992).
[2] S. V. Belotelov, V. I. Ivanov-Omskiĭ, A. I. Izhnin, and V. A. Smirnov, Fiz. Tekh. Poluprovodn. **25**, 1058 (1991) [Sov. Phys. Semicond. **25**, 637 (1991)].
[3] D. L. Spears, in *Lasers and Electrons. Conference Proceedings*, Orlando, Florida, New York, (1989), pp. 113–116.

[4] M. G. Andrukhiv, S. V. Belotelov, and I. S. Virt, Fiz. Tekh. Poluprovodn. **27**, 1863 (1993) [Semiconductors **27**, 1026 (1993)].
[5] Yu. R. Nosov, *Switching in Semiconductor Diodes*, Plenum Press, New York (1969).
[6] A. Rogalski and Y. Piotrowski, Prog. Quantum Electron. **12** (2–3), 87 (1988).

Translated by P. Shelnitz

Use of moiré effects in the optical treatment of images of a periodic amplitude grating distorted by a reconstructed wave front

A. I. But' and A. M. Lyalikov

Ya. Kupala Grodno State University, 230023 Grodno, Belarus
(Submitted February 21, 1996; resubmitted August 12, 1996)
Zh. Tekh. Fiz. **67**, 134–136 (July 1997)

The use of moiré effects in superimposing images of an amplitude grating distorted by a reconstructed wave front for increasing the sensitivity of measurements of the deflection angles of the reconstructed rays is proposed. © *1997 American Institute of Physics.*
[S1063-7842(97)02607-X]

The schlieren methods for investigating a reconstructed wave front are widely used in the holographic interferometry of nonstationary processes in the presence of phase inhomogeneities which deflect light rays by large angles.[1,2] Various modifications of the defocused diaphragm method, which is a quantitative schlieren method, permit determination of the deflection angles of light rays caused by inhomogeneities in the phase object.[3]

The present work is devoted to further development of a quantitative schlieren method for investigating the distortions of a reconstructed wave front by obtaining the distorted image of a withdrawn defocused amplitude grating.[4] The possibility of increasing the amount of information provided by the method and the accuracy of the quantitative measurements by increasing the sensitivity of the method is considered on the basis of moiré effects.

Figure 1 presents a diagram of the reconstruction of a wave front by a singly exposed hologram of a focused image of a phase object and its subsequent investigation. A visualizing amplitude grating, which is illuminated along the normal by a collimated beam of monochromatic light with the wavelength λ, is located at position *1*. After passing through grating *1*, the amplitude-modulated plane wave front impinges on object hologram *2*, which is located at a distance L from the grating *1*. The amplitude transmission of the singly exposed hologram of the focused image of the phase object is described by the expression[1]

$$\tau(x,y)\sim 1+\cos\left[\frac{2\pi x}{P}+\phi(x,y)\right], \tag{1}$$

where x and y are the coordinates in the plane of the hologram, the y axis is parallel to the holographic fringes with a period P, and $\phi(x,y)$ describes the phase distortions caused by the object under investigation.

Since the phase distortions of the waves reconstructed in the $+1$ and -1 diffraction orders have opposite signs, the deflection angles of the light rays from rectilinear propagation caused by inhomogeneities in the phase object will also have opposite signs. For example, the x components of the angles for the reconstructed waves in the $+1$ and -1 diffraction orders will be proportional, respectively, to

$$\frac{\partial\phi(x,y)}{\partial x} \text{ and } -\frac{\partial\phi(x,y)}{\partial x}.$$

We assume that the increase in the telescopic system formed by objective lenses *3* and *5* is equal to unity and that the periods of the hologram fringes (P) and the amplitude grating (T) satisfy the condition $P \ll T$. To determine the x components of the deflection angles of the rays reconstructed by the hologram from the rectilinear direction, i.e., ε_x, the bars of the grating must be oriented parallel to the y axis. If the complex conjugate ± 1 diffraction orders of the waves on the hologram are singled out in the Fourier plane of objective lens *3* by spatial filter *4* (a diaphragm with openings), it can be shown with consideration of relation (1) that in plane *6* which is optically conjugate to the hologram the distribution of the complex wave amplitudes will have the respective forms

$$A_{+1}(x,y)=\cos\left[\frac{2\pi x}{T}+\frac{2\pi L\varepsilon_x}{T}\right]\times\exp\left\{1\left[\frac{2\pi x\cos\alpha}{\lambda}+\phi(x,y)\right]\right\}, \tag{2}$$

$$A_{-1}(x,y)=\cos\left[\frac{2\pi x}{T}-\frac{2\pi L\varepsilon_x}{T}\right]\times\exp\left\{-1\left[\frac{2\pi x\cos\alpha}{\lambda}+\phi(x,y)\right]\right\}, \tag{3}$$

where $\cos\alpha$ is the direction cosine of a wave diffracted in the first order, which is defined as $\cos\alpha=\lambda/P$.

The first cofactors in expressions (2) and (3) specify the real amplitudes $A_{+1}(x,y)$ and $A_{-1}(x,y)$ of the waves. Spatial modulation of the real amplitudes is caused by the presence of the low-frequency amplitude grating *1*.

When the waves described by Eqs. (2) and (3) are recorded separately in plane *6*, the distributions of the illuminance in the patterns for the $+1$ and -1 diffraction orders, respectively, have the forms

$$I_{+1}=A_{+1}A_{+1}^{*}=\cos^2\left[\frac{2\pi x}{T}+\frac{2\pi L\varepsilon_x}{T}\right], \tag{4}$$

$$I_{-1}=A_{-1}A_{-1}^{*}=\cos^2\left[\frac{2\pi x}{T}-\frac{2\pi L\varepsilon_x}{T}\right]. \tag{5}$$

Figures 2a and 2b show the patterns of the distorted images of an amplitude grating observed through the hologram of a sphere on a ballistic trajectory. The pattern in Fig.

 1063-7842/97/070845-03$10.00

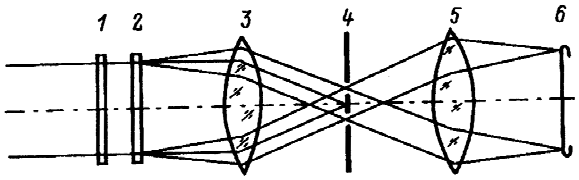

FIG. 1. Optical diagram of the investigation of a wave front reconstructed by a hologram of a phase object.

2a was obtained using only the $+1$ diffraction order, and the illuminance in it is described by Eq. (4). The other pattern, in Fig. 2b, was obtained using the -1 diffraction order and is described by Eq. (5).

It is seen from Eqs. (4) and (5) and the patterns presented (Figs. 2a and 2b) that the deviations of the bars due to the phase inhomogeneities in the image of the visualizing amplitude grating have different signs for the same points in the object. A similar property, but with the formation of interference patterns, was previously utilized in the holographic interferometry of phase objects to increase the sensitivity of the measurements when the hologram is recorded by three beams and moiré patterns are obtained.[5] Superposition of the schlieren patterns formed by the waves reconstructed in the complex-conjugate orders was used to compensate for the inhomogeneities in the hologram substrate, as well as to increase the sensitivity of the measurements in Ref. 6.

The differences in the signs of the deviations of the images of the lattice bars can be used to increase the sensitivity of the measurements of the angles ε_x in treating patterns of forms (4) and (5). Let us consider the formation of moiré patterns when distorted grating images (4) and (5) are superimposed on one another. The moiré patterns can be observed by three methods.

1. We assumed that waves of forms (2) and (3) are superimposed on one another in plane *6* (Fig. 1). Their coherent summation gives an illuminance distribution of the form

$$I=(A_{+1}+A_{-1})(A^*_{+1}+A^*_{-1}). \tag{6}$$

2. If patterns (4) and (5) are successively recorded on a photographic material, according to the double-exposure method, the amplitude transmission of the double-exposure photograph has the form

$$\tau=I^{-\gamma/2}=(I_{+1}+I_{-1})^{-\gamma/2}, \tag{7}$$

where γ is the contrast coefficient of the photographic material.

3. If photographs of patterns (4) and (5) are obtained separately, the resultant amplitude transmission of such combined photographs can be represented in the form

$$\tau=\tau_1\tau_2=I_{+1}^{-\gamma/2}+I_{-1}^{-\gamma/2}. \tag{8}$$

It can be shown for all three cases[7] that the low-frequency modulation of the illuminance in (6), as well as of the amplitude transmissions in (7) and (8), will be described by a single term of the form

$$\cos\frac{\pi x}{T}\cos\left(\frac{2\pi L\varepsilon_x}{T}\right). \tag{9}$$

The high-frequency component described by the first cofactor will be modulated slowly by the varying function (the second cofactor). It follows from (9) that the visibility of the fringes is lowest at the points where

$$\frac{2\pi L\varepsilon_x}{T}=N\pi+\pi/2, \quad N=0,1,2,\ldots. \tag{10}$$

Therefore, a region of low visibility (a moiré fringe) is the geometric locus of the points at which ε_x is constant. Condition (10) leads to a working formula for determining the x component of the deflection angles of the rays

$$\varepsilon_x=\frac{T(N+1/2)}{2L}. \tag{11}$$

When patterns of forms (4) and (5) and moiré patterns of form (9) are interpreted, it is seen that a single measurement of ε_x will correspond to different changes in the order of a fringe. The change in the order of the fringe will be two times larger in the latter case than in the former. Thus, the sensitivity of the measurements will be twice as high in the latter case then in the case described in Ref. 4. The increase in the sensitivity of the measurements is associated with the increase in the number of moiré fringes when two grating images with diametrically opposite displacements of the bars are superimposed.

Figures 3a and 3b present moiré patterns obtained by the methods described. The first pattern (Fig. 3a) corresponds to the illuminance in the recording plane *6* (Fig. 1) when waves of forms (2) and (3) are coherently summed, i.e., it was obtained using the first method. When the second method is employed, the form of the moiré pattern formed when the double-exposure photograph is illuminated corresponds exactly to the pattern presented in Fig. 3a. Use of the third method, unlike the first two, also makes it possible to adjust the tuning of the field of the unperturbed zone of the phase object. When the two separate photographs are juxtaposed exactly and illuminated by either coherent or incoherent light, the distribution of the illuminance in the moiré pattern will also correspond to the pattern presented in Fig. 3a. However, when one photograph is shifted relative to the other

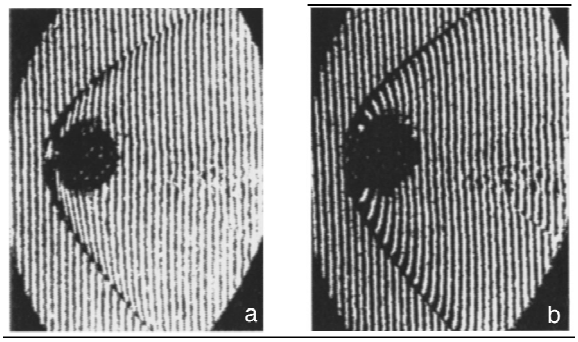

FIG. 2. Patterns visualizing the distribution of the x component of the deflection angles of light rays reconstructed by the hologram of a sphere on a ballistic trajectory obtained with separate recording in the $+1$ (a) and -1 (b) diffraction orders.

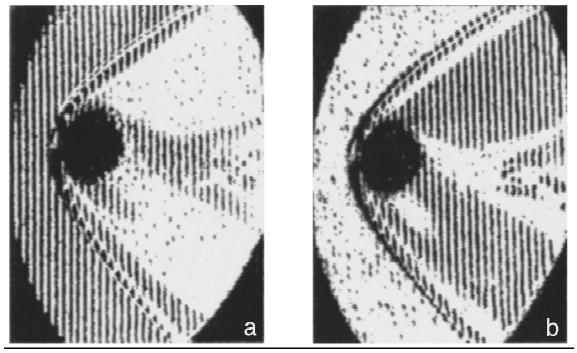

FIG. 3. Moiré patterns visualizing the distribution of the x component of the deflection angles of light rays reconstructed by the hologram of a sphere on a ballistic course, which were obtained by different methods: a — coherent summation of the waves propagating in the $+1$ and -1 diffraction orders; b — juxtaposition and incoherent illumination of two separate photographs, one of which is shifted along the x axis by an amount equal to $T/2$.

along the x axis by an amount equal to $T/2$, the zones of minimum visibility in (9) will be replaced by zones of maximum visibility. This can be utilized to eliminate the ambiguity in the determination of the number of each moiré fringe. Figure 3b presents the moiré pattern obtained by the third method with illumination of the juxtaposed photographs by white light and tuning of the unperturbed field to a light moiré fringe.

To determine the y component of the deflection angle of the rays reconstructed by a hologram, i.e., ε_y, the ''rulings'' of grating *1* in Fig. 1 must be oriented parallel to the x axis, and the working formula (11) must be used to interpret the moiré pattern.

Thus, superposition of the distorted images of an amplitude grating observed in complex-conjugate orders reconstructed by a hologram makes it possible to obtain a moiré pattern with an increased measurement sensitivity.

In conclusion, we would like to express our thanks to I. S. Zeĭlikovich for providing the hologram of a sphere on a ballistic trajectory.

[1] A. K. Beketova, A. F. Belozerov, A. N. Berezkin *et al.*, in *Holographic Interferometry of Phase Objects* [in Russian], Nauka, Leningrad (1979).
[2] I. S. Zeĭlikovich and N. M. Spornik, *Holographic Diagnostics of Transparent Media* [in Russian], Univ. Izd., Minsk (1988).
[3] L. A. Vasil'ev, *Schlieren Methods*, Israel Program for Scientific Translations, Jerusalem (1971).
[4] A. I. But' and A. M. Lyalikov, Tech. Phys. **42**, 287 (1997).
[5] L. Sevigny, Appl. Phys. Lett. **10** (3), 78 (1967).
[6] I. S. Zeĭlikovich and A. M. Lyalikov, Opt. Spektrosk. **68**, 197 (1990) [Opt. Spectrosc. (USSR) **68**, 112 (1990)].
[7] C. M. Vest, *Holographic Interferometry*, Wiley, New York (1979) [Russ. transl., Mir, Moscow (1982)].

Translated by P. Shelnitz

Influence of the state of the bulk of the sample on the formation and thermal stability of the surface silicide on W(100)

N. R. Gall', E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted April 4, 1996)
Zh. Tekh. Fiz. **67**, 137–140 (July 1997)

The formation and destruction of the surface silicide on W(100) after cleaning of the sample surface and bulk in various regimes is studied by high-resolution Auger electron spectroscopy. It is shown that the cleanness of the bulk has practically no influence on the laws governing the formation of the surface silicide when Si atoms are adsorbed on a heated W surface and that almost up to completion of its formation all the silicon atoms impinging on the surface, from the very first, remain on it and are incorporated into the surface silicide. The destruction of the surface silicide depends in a definite manner on the state of the bulk, and at $T=1400$ K it is apparently limited in the early stages by the passage of Si atoms from the surface to the subsurface layer and in subsequent stages by the diffusion of silicon within the substrate. The bulk silicon density that limits the destruction of the surface silicide is estimated. © *1997 American Institute of Physics.* [S1063-7842(97)02707-4]

INTRODUCTION

A surface chemical compound of silicon with a refractory metal, i.e., a surface silicide, was first discovered in the study of the adsorption of silicon on the W(100) surface in Ref. 1 and was found to have the stoichiometry WSi. The surface silicide discovered had some remarkable properties: in the temperature range $T=1100-1350$ K its formation preceded all other transport processes and solid-phase reactions taking place when Si atoms are adsorbed on the metal surface, and all the silicon atoms impinging on the surface are incorporated into the surface silicide until its formation is completed, while all the Si atoms arriving after its formation dissolve completely in the W bulk. Surface silicides having similar properties were soon discovered on Re$(10\bar{1}0)$,[2] Ir(111),[3] Mo(100),[4] and Ta(100).[5] High-temperature thermostable surface silicides with strict stoichiometry were observed in Ref. 6 on an Ni surface (Ni_2Si) and in Ref. 7 on a Pt surface ($PtSi_2$). The surface carbides of Re,[8] W,[9] and Mo[10] and the surface sulfides of W[11] and Mo[12] also have properties that are similar in many respects to those of surface silicides.

The studies of the properties of surface silicides in Refs. 1–7 focused mainly on the processes taking place on the surface, while the role of the substrate bulk and the possible influence of its state on the formation and stability of the surface silicide were not taken into account. At the same time, as was shown in Ref. 10, the formation and thermal stability of a closely related compound, viz., the surface carbide on Mo(100), are decisively dependent on the density of carbon dissolved in the Mo bulk, the densities needed to maintain the surface carbide on the surface lying in the ''trace'' range: $n_c=0.01$ at.%.

The role of the state of the bulk, particularly the density of silicon dissolved in the bulk or, at least, in its subsurface region is apparently important for understanding the nature of a surface silicide. The study of this influence is the purpose of the present work.

INSTRUMENTS AND EXPERIMENTAL METHODS

The experiments were performed in a high-resolution prism Auger spectrometer.[8] The samples used were strips of tungsten foil measuring $1\times0.02\times40$ mm, which were heated by passing a variable current through them. The strips were cleaned by prolonged annealing alternately in oxygen ($p=10^{-6}$ Torr) at $T=1500$ K and in an ultrahigh vacuum at $T=2600$ K. After the cleaning, only Auger peaks of tungsten were observed on the surface. Along with the cleaning, the strips were textured, and the (100) face emerged on the surface, the work function of the surface being equal to 4.65 eV (a value which is characteristic of that face[13]). The surface was homogeneous with respect to the work function. The temperature of the strips was measured by an optical micropyrometer and was determined in the nonpyrometric region by extrapolating the dependence of the temperature on the heating current. The Auger spectra could be obtained directly at strip temperatures as high as $T=2100$ K. The peak-to-peak amplitudes of the silicon Auger peak with $E=92$ eV and the tungsten peak with $E=179$ eV were used in the measurements. Since the surface densities of silicon definitely did not exceed the monolayer level during the formation of the surface silicide, the intensity of the Auger signal was assumed to be directly proportional to the surface density of the adsorbate.

Silicon was sputtered onto the surface from a strand measuring $1\times1\times40$ mm. The silicon flux was calibrated absolutely according to the method described in Ref. 10. The accuracy of the calibration was $\pm10\%$.

We do not know of any direct methods that permit measurement of the bulk densities of silicon dissolved in the subsurface region of a metal in the range $n_{Si}=0.1-0.01$ at.%, especially under conditions for which the total number of Si atoms on the surface is comparable to or exceeds the number of these atoms dissolved in the bulk. Therefore, the cleanness of the bulk was evaluated on the basis of the sample cleaning regime employed between series of experi-

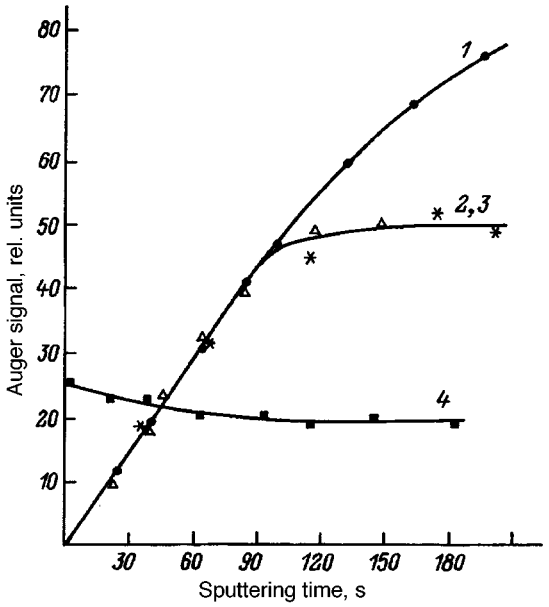

FIG. 1. Dependence of the Auger signals of silicon (*1*–3) and W (*4*) on the sputtering time for the sputtering of silicon onto W(100) at $T=300$ (*1*) and 1300 K (*2*–*4*). The sample was cleaned by annealing at $T=2500$ K for 100 s (asterisks) and at 2000 K for 10 s (triangles). The flux density of the silicon atoms was $\nu_{Si}=1\times10^{13}$ atoms/cm^{2}s.

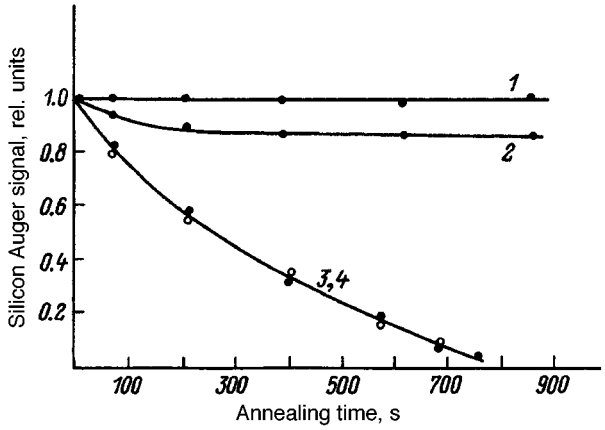

FIG. 2. Dependence of the relative Auger signal of silicon on the annealing time for silicon from the surface silicide on W(100) and various states of the bulk of the sample. *1* — initial state — surface silicide on a sample containing a silicon dose equal to 100 times the dose corresponding the surface silicide at $T=1400$ K; *2* — initial state — surface silicide on a sample that was cleaned by annealing at $T=2000$ K for 10 s; *3* (filled circles) — initial state — $N_{Si}=2.3\times10^{14}$ atoms/cm^2 (0.23 of the surface silicide dose) on a sample that was cleaned at $T=2500$ K for 100 s; *4* (unfilled circles) — initial state — $N_{Si}=2\times10^{14}$ atoms/cm^2 (0.2 of the surface silicide dose) on a sample that was cleaned at $T=2000$ K for 90 s.

ments. For example, annealing of the strips at $T=2100$ K for 30 s (this cleaning regime is similar to the one employed in Ref. 1) led to complete removal of the Si atoms from the surface, but it apparently did not remove them from the bulk of the metal. Annealing at $T=2500$ K for 100 sec was employed to completely clean both the surface and the W bulk. As follows from the data in Ref. 14, in which direct mass-spectrometric detection of the Si atoms desorbed from a W surface was performed, the desorption flux of Si atoms from W does, in fact, end during the annealings just described. Let us evaluate the diffusion lengths for Si atoms in a W lattice under such conditions. The diffusion length is given by the following expression

$$L_{dif}=(D_{Si}\cdot t)1/2=(D_0)^{1/2}\exp[-E_{dif}/2kT]\sqrt{t}. \quad (1)$$

Setting $D_0=10^{-3}$ cm^2/s and $E_{dif}=2.5$ eV for the diffusion of Si atoms in refractory metals,[15] as well as $T=2500$ K and the time of a single experiment to ~100 s, we obtain $L_{dif}=100$ μm, which is an order of magnitude greater than the strip thickness. Thus, the absence of a desorption flux apparently does attest to the cleaning of both the surface and the bulk of the sample.

EXPERIMENTAL RESULTS

a) Formation of the surface silicide. Figure 1 presents the variation of the silicon and tungsten Auger peaks during the sputtering of Si atoms onto a W(100) surface at $T=1300$ K, i.e., in a regime where the surface silicide forms, according to Ref. 1. In fact, the plot of $I_{Si}=I_{Si}(t)$ is initially linear (practically up to $t=100$ s) and reaches saturation at a level corresponding to a surface silicide with a sputtered dose of Si atoms $N_{Si}=1\times10^{15}$ cm^{-2}, which corresponds to the stoichiometry of the surface silicide WSi. The high-temperature sputtering curve closely coincides with the sputtering curve obtained at room temperature up to $t=90$ s, providing evidence that almost up to completion of the formation of the surface silicide all (to within an experimental error of 1–2%) the silicon atoms impinging on the surface, from the very first, remain on it and are incorporated into the surface silicide.

It is noteworthy that no differences stemming from the preliminary sample cleaning regimes could be detected in the laws governing the formation of the surface silicide: this process had an identical course on a surface of W in whose bulk a silicon dose $N_{Si}=10^{17}$ cm^{-2} (i.e., about 100 times the dose corresponding to the surface silicide) was dissolved at $T=2000$ K and on a surface of a metal with a bulk practically free of dissolved silicon.

b) Thermal stability of the surface silicide. Figure 2 presents the results of the annealing of the surface silicide formed on the surface of tungsten with different densities of dissolved silicon in the bulk. As we see, in the first case (curve *1*), in which the surface silicide was formed by dissolving a thick silicon film with $N_{Si}=10^{17}$ atoms/cm^2 and there was a considerable quantity of dissolved silicon in the bulk, the surface silicide is stable up to $T=1400$ K over the course of at least 1000 s, and no tendency for weakening of the silicon Auger signal is observed.

At the same time, under conditions in which the surface silicide forms on a sample that has been freed of silicon to a considerable extent (curves *2*–*4*), the surface silicide loses its thermal stability already at $T=1400$ K: under conditions in which the bulk is not completely cleaned, the silicon Auger signal decreases and stabilizes at a new value, but it decreases to zero, if the bulk is more thoroughly cleaned. As was shown in Ref. 14, the thermal desorption of silicon from W takes place only at $T>1500$ K; therefore, in our case the

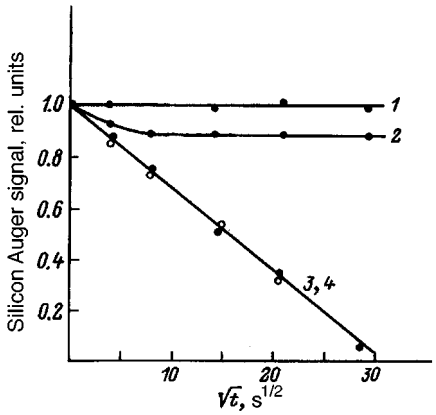

FIG. 3. Dependence of the relative Auger signal of silicon from the surface silicide on the square root of the annealing time. The numbering of the curves is the same as in Fig. 2.

departure of silicon from the surface can be attributed only to its dissolution in the bulk of the metal.

The departure of silicon from the surface can be limited both by diffusion in the bulk and by the dissolution process proper, i.e., by the passage of atoms from the surface into the bulk of the sample. The curves themselves, which are presented in Fig. 2, do not permit a choice between these possibilities. To elucidate this question, Fig. 3 presents the same curves reconstructed in different coordinates. As before, the intensity of the silicon Auger signal (which is proportional to the surface silicon density) is plotted along the vertical axis, but the square root of the annealing time is plotted along the horizontal axis. As we see, the form of curve *2* has changed only slightly, while curves *3* and *4*, which characterize the destruction of the surface silicide on a sample with a ''silicon-free'' bulk, have transformed into straight lines, which are characteristic of the case in which bulk diffusion determines the kinetics of the process.[16]

DISCUSSION OF RESULTS

Let us discuss the results obtained. We, first of all, evaluate the bulk densities of silicon in the subsurface region of W when the surface silicide dissolves. Dissolved silicon atoms from the surface silicide with a surface density N_{Si} are located at a depth equal to the diffusion length L_{dif}, and the corresponding bulk density is

$$n_{\mathrm{Si}}=N_{\mathrm{Si}}/L_{\mathrm{dif}}. \tag{2}$$

We can utilize of Eq. (1) to calculate it. When $T=1400$ K and the dissolution time equals 400 s, $L_{\mathrm{dif}}=2\times10^{-5}$ cm$=2000$ Å. If the silicon atoms that formerly comprised the surface silicide and had a surface density $N_{\mathrm{Si}}=1\times10^{15}$ cm^{-2} (which corresponds to the surface silicide WSi) are now accommodated in this volume, we obtain $n_{\mathrm{Si}}=5\times10^{19}$ cm$^{-3}=0.1$ at.% for the corresponding bulk density, since the bulk density of tungsten is $n_{\mathrm{W}}=6.3\times10^{22}$ cm^{-3} (Ref. 17). There are no data in the literature on the limiting solubility of silicon in tungsten in the temperature range cited, but, extrapolating the values obtained for higher temperatures equal to 1900–2300 K into this range, we can expect that the limiting solubility must be $n_{\mathrm{Si}}=\sim1$ at.%. Thus, bulk densities of dissolved silicon that are an order of magnitude smaller than the expected value of the limiting solubility already limit the removal of Si atoms in the surface silicide from the surface.

It would be interesting to understand whether the available experimental data can somehow be used to evaluate E_{s1}, i.e., the activation energy for the passage of a Si adatom from the surface into the first subsurface layer of the bulk. It can be hypothesized that just this process determines the rate of dissolution of the very first Si atoms in the surface silicide before there is any dissolved silicon in the bulk. Assuming that this hypothesis is correct, let us estimate E_{s1} from the slope of the initial portions of the $I_{\mathrm{Si}}=I_{\mathrm{Si}}(t)$ curves (curves *3* and *4* in Fig. 2) under the additional assumption that we can use the Arrhenius relation for the lifetime of a particle on the surface with respect to the dissolution process τ_{s1}

$$\tau_{\mathrm{s1}}=\tau_0\ \exp[E_{\mathrm{s1}}/kT]. \tag{3}$$

Setting $\tau_0=10^{-13}$ s, we obtain $E_{\mathrm{s1}}=4.3$ eV, where the accuracy of the evaluation with consideration of the indefiniteness of the value of τ_0 amounts to $\Delta E=\pm0.2$ eV.

The data presented indicate that the results in Ref. 1 pertaining to the physical picture of the destruction of a surface silicide are in need of some revision. In particular, the desorption of silicon from a W surface is apparently preceded by its partial predissolution, whose possibility was pointed out in Ref. 14. Apparently, the arguments developed above are also relevant for other adsorption systems consisting of Si and a surface of a refractory metal, as is confirmed, in particular, by the data in Ref. 5, where a dependence of the thermal stability of the surface silicide of tantalum on the density of silicon dissolved in the bulk of the sample was observed.

BRIEF RESULTS AND CONCLUSIONS

The influence of silicon dissolved in the bulk of a metal on the processes involved in the formation of the corresponding surface silicide and on its thermal stability has been studied. It has been shown that the formation of the surface silicide of tungsten as a result of the high-temperature adsorption of silicon on its surface scarcely depends on the state of the bulk, and thus the very first Si atoms impinging on the surface are incorporated into the surface silicide. Conversely, the destruction of the surface silicide is decisively dependent on the state of the bulk. For example, at $T=1400$ K the dissolution rate is initially determined by the passage of Si atoms from the surface into the subsurface layer and is subsequently limited by the bulk diffusion of silicon in tungsten.

This work was supported by the Russian State Program ''Surface Atomic Structures.''

[1] V. N. Ageev, E. Yu. Afanas'eva, N. R. Gall' *et al.*, Poverkhnost', No. 5, 7 (1987).
[2] N. R. Gall', E. V. Rut'kov, and A. Ya. Tontegode, Zh. Tekh. Fiz. **60** (4), 125 (1990) [Sov. Phys. Tech. Phys. **35**, 475 (1990)].
[3] N. R. Gall', E. V. Rut'kov, and A. Ya. Tontegode, Poverkhnost'. No. 10, 47 (1989).

[4] N. R. Gall, E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Phys. Low-Dimens. Struct. **9**, (1994).
[5] N. R. Gall, E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Phys. Low-Dimens. Struct. (1996) (in press).
[6] V. M. Bermudes, Surf. Sci. Lett. **230**, L155 (1990).
[7] H. P. Bonzel, A. M. Franken, and C. Paris, Surf. Sci. **104**, 625 (1981).
[8] N. R. Gall, S. N. Mikhailov, E. V. Rut'kov, and A. Ya. Tontegode, Surf. Sci. **191**, 185 (1987).
[9] V. N. Ageev, E. Yu. Afanas'eva, N. R. Gall' *et al.*, Pis'ma Zh. Tekh. Fiz. **12**, 565 (1986) [Sov. Tech. Phys. Lett. **12**, 231 (1986)].
[10] E. V. Rut'kov, A. Ya. Tontegode, M. M. Usufov, and N. R. Gall, Appl. Surf. Sci. **78**, 179 (1994).
[11] E. V. Rut'kov, A. Ya. Tontegode, M. M. Usufov, and N. R. Gall, Appl. Surf. Sci. (1996) (in press).
[12] N. R. Gall', E. V. Rut'kov, A. Ya. Tontegode, and M. M. Usufov, Zh. Tekh. Fiz. **66**(5), 143 (1996) [Tech. Phys. **41**, 483 (1996)].
[13] V. S. Fomenko, *Emission Properties of Materials. A Handbook* [in Russian], Naukova Dumka, Kiev (1980).
[14] V. N. Ageev and E. Yu. Afanas'ev, Poverkhnost', No. 7, 30 (1987).
[15] G. V. Samsonov and I. M. Vinitskii, *Handbook of Refractory Compounds*, Plenum Press, New York (1980) [Russ. original, Metallurgiya, Moscow (1976), pp. 232-240].
[16] R. Rosenberg, M. J. Sullivan, and J. K. Howard, ''Effect of thin film interactions on silicon device technology,'' in *Thin Films — Interdiffusion and Reactions*, J. M. Poate, K. N. Tu, and J. W. Mayer (Eds.), Wiley-Interscience, New York (1978), pp. 13–55 [Russ. transl., Mir, Moscow (1982), pp. 17–58].
[17] C. Kittel, *Introduction to Solid State Physics*, 5th ed., Wiley, New York (1976) [Russ. transl., Mir, Moscow (1978), p. 55].

Translated by P. Shelnitz

Effect of surface diffusion of adsorbed molecules on energy transfer through a porous layer

V. V. Levdanskiĭ

Academic Scientific Complex, A. V. Lykov Institute of Heat and Mass Transfer, Academy of Sciences of Belarus, Belarus

(Submitted July 8, 1996)
Zh. Tekh. Fiz. **67**, No. 7, 141–142 (July 1997)

[S1063-7842(97)02807-9]

It is well known that gas flow through a porous body is accompanied by energy transfer. In the case of bodies with small pores the interaction of the gas particles with a surface greatly influences the transfer process.[1] For example, it is well known that surface diffusion of molecules can influence mass transfer in microporous bodies (it can also initiate mass transfer in an initially equilibrium system[2]). The effect of segregation and surface diffusion of molecules in pores on the thermal conductivity of porous bodies was investigated in Ref. 3. This paper examines the question of the appearance of a temperature differential in a porous layer as a result of adsorption, desorption, and surface diffusion processes. This effect arises as follows. Molecules adsorbed on one side of a porous layer can diffuse along the surface, enter pores, and cross to a different surface of the layer, whence they deosrb. In this case a source of heat appears on one side of the porous layer. The appearance of this heat source is due to the fact that the energy released as molecules are adsorbed from the gas phase is greater than the energy going into desorption of the molecules (since some adsorbed molecules enter the porous layer). The opposite situation occurs on the other side of the layer, where molecules migrate from pores onto the surface of the porous layer and then desorb. In this manner, surface diffusion of the molecules leads to energy transfer and to the appearance of a temperature differential in the porous layer.

In this paper we shall examine the case of a free-molecular gas flow through a capillary-porous body of thickness L. We assume that the pores are all cylindrical and possess the same radius r (nuclear filters are characterized by such properties). The flux of molecules through the transverse section of a quite long capillary has the form

$$j_r = j_{\mathrm{Kn}} + j_s, \tag{1}$$

where j_{Kn} is the particle flux in the gas phase (Knudsen diffusion) and j_s is the surface flux of particles.

For simplicity, let us assume at first that the average surface diffusion length l_s of the adsorbed particles ($l_s \sim D_s\tau$, where D_s is the surface diffusion coefficient and τ is the adsorption time) is much longer than the capillary radius. In this case the first term in Eq. (1) can be neglected compared with the second term and the flux j_r can be expressed as

$$j_r = j_s = -2\pi r D_s \frac{dn_a}{dX}, \tag{2}$$

where n_a is the density of the adsorbed particles.

Assuming that the sticking coefficient for molecules at the surface equals 1 and the adsorption of the molecules is described by Henry's law, we obtain for $n_a(0)$ and $n_a(L)$

$$n_a(0) = \frac{P(0)\tau(0)}{(2\pi m k T_g)^{1/2}}, \quad n_a(L) = \frac{P(L)\tau(L)}{(2\pi m k T_g)^{1/2}}, \tag{3}$$

where $P(0)$ and $P(L)$ are the gas pressure at $X=0$ and $X=L$ (we assume below that $P(L)/P(0)>1$), T_g is the temperature of the surrounding gas, m is the mass of a molecule, and k is Boltzmann's constant.

The quantities τ and D_s can be written in the form[1]

$$\tau = \tau_0 \exp\{Q_a/kT\}, \quad D_s = D_{s0}\exp\{-Q_d/kT\}, \tag{4}$$

where Q_a is the energy of adsorption, Q_d is the activation energy of surface diffusion, and τ_0 and D_{s0} are pre-exponential factors, which depend on the vibrational frequencies of the adsorbed particles.

For the problem described above, in the stationary case the temperature distribution over the thickness of the porous layer is found from the equation

$$\frac{d}{dX}\left(\lambda \frac{dT}{dX}\right) = 0 \tag{5}$$

with the boundary conditions

$$\lambda \frac{dT}{dX}\bigg|_{X=0} = \alpha[T(0) - T_g] + \gamma,$$

$$\lambda \frac{dT}{dX}\bigg|_{X=L} = \alpha[T_g - T(L)] + \gamma. \tag{6}$$

Here λ is the thermal conductivity, which is assumed to be constant; α is the heat-transfer coefficient; $\gamma = j_s N_s Q_a$, where N_s is the density of the capillary outlet points at the surface of the porous layer. From Eqs. (5) and (6) we find the temperature distribution in the layer as

$$T = A + Cx, \tag{7}$$

where

$$A = T_g + \frac{C}{\mathrm{Bi}} - \frac{\gamma}{\alpha}, \quad C = \frac{2\gamma L}{\lambda(2+\mathrm{Bi})},$$

$\mathrm{Bi} = \alpha L/\lambda$ is Biot's number and $x = X/L$.

It is evident from Eq. (7) that for $j_s \neq 0$ $T(0) < T_g$, $T(L) > T_g$, and

 1063-7842/97/070852-02$10.00

$$T(L)-T_g=T_g-T(0)=\frac{\gamma}{\alpha}\frac{\mathrm{Bi}}{2+\mathrm{Bi}}.$$

An expression for j_s follows from Eqs. (2)–(7) (for simplicity, we assume that $(Q_aC/kA^2)\ll 1$, $Q_dC/kA^2\ll 1$):

$$j_s=\frac{2\pi r}{L}\frac{P(0)\tau_0 D_{s0}}{(2\pi mkT_g)^{1/2}}\exp\{(Q_a-Q_d)/kA\}\times\left[\frac{P(L)}{P(0)}(1-Q_aC/kA^2)-1\right]. \quad (8)$$

Expressions for $T(0)$, $T(L)$, and $\Delta T=T(0)-T(L)$ can be found from Eqs. (7) and (8). It should be noted that the temperature differential will increase with the mass flux, the temperature being higher on the side with the higher gas pressure. This will in turn increase the resistance of the porous body to mass transfer.

The problem of finding the quantity j_s can be solved more accurately on the basis of an integrodifferential equation for n_a (Refs. 2 and 4). This equation makes it possible to take into account the effect of external fields on the transfer process and to describe correctly the simultaneous mass transfer in the gas and adsorption phases without assuming local equilibrium in each section of the capillary. On the basis of the assumptions made above, we obtain for n_a the equation

$$\frac{1}{L^2}\frac{d}{dx}\left(D_s\frac{dn_a}{dx}-LbFn_a\right)=\frac{n_a}{\tau}-\beta\left\{\int_0^1\left(\frac{n_a}{\tau}+I\right)\times K_1(|x-x'|)dx'+N_0K(x)+N_1K(1-x)\right\},$$

$$I=(1-\beta)\left\{\int_0^1\left(\frac{n_a}{\tau}+I\right)K_1(|x-x'|)dx'+N_0K(x)+N_1K_1(1-x)\right\}. \quad (9)$$

Here x is the dimensionless coordinate ($x=X/L$); β is the sticking coefficient for gas particles at the surface; N_0 and N_1 are the densities of the particle fluxes entering a capillary at $x=0$ and $x=1$; the functions K and K_1 characterize the probability that a particle is transferred from one element of the capillary surface to another;[4] F is the force acting on the adsorbed particles along the surface; and, b is the mobility of the adsorbed molecules ($b=D_s/kT$).

Assuming that an equilibrium density of the particles adsorbed on the end surfaces of the capillaries is established at a characteristic distance $(D_s\tau)^{1/2}$, the following boundary conditions can be imposed at $x=0$ and $x=L$:

$$D_s\left.\frac{dn_a}{dx}\right|_{x=0}-LbFn_a(0)=L\left(\frac{D_s}{\tau}\right)^{1/2}[n_a(0)-n_{a0}],$$

$$D_s\left.\frac{dn_a}{dx}\right|_{x=1}-LbFn_a(1)=L\left(\frac{D_s}{\tau}\right)^{1/2}[n_{a1}-n_a(1)],$$

where $n_{a0}=\beta\tau N_0$ and $n_{a1}=\beta\tau N_1$.

The surface flux of adsorbed particles at a capillary outlet can be found with the aid of Eq. (2). In Ref. 2 an equation of the type (9) was solved with an exponential approximation of the functions K and K_1. We note that the adsorption times τ and the surface diffusion coefficients D_s on the end surfaces of the capillaries emerging at opposite sides of the porous layer can differ as a result of the effect of external fields. Specifically, their values can change under the action of resonance radiation.[5] As a result of the fact that the indicated quantities have different values on opposite surfaces of the porous layer, a surface flux of molecules is initiated[6] and, as follows from the preceding discussion, a temperature differential arises in the system. In the case when the adsorbed molecules possess an electric charge or a dipole moment, mass transfer can be initiated by imposing an electric field on the system (then the force F in Eq. (9) is different from zero). Therefore external fields initiating mass transfer in a porous body will lead to energy transfer and, accordingly, to the appearance of a temperature differential in the porous body.

[1] J. H. de Boer, *The Dynamical Character of Adsorption*, Clarendon Press, Oxford, 1953 [Russian translation, Inostr. Lit., Moscow, 1962].
[2] V. V. Levdanskiĭ, Inzh. Fiz. Zh. **31**, No. 1, 97 (1976).
[3] T. Gambaryan, E. Litovsky, M. Shapiro, and A. Shavit, in *Proceedings of the 10th International Heat Transfer Conference*, Brighton, UK, 1994, Vol. 5, pp. 267–272.
[4] W. L. Winterbottom, J. Chem. Phys. **47**, 3546 (1967).
[5] N. V. Karlov and A. M. Prokhorov, Usp. Fiz. Nauk **123**, 57 (1977) [Sov. Phys. Usp. **20**, 721 (1977)].
[6] V. V. Levdankiĭ, Zh. Tekh. Fiz. **52**, 825 (1982) [Sov. Phys. Tech. Phys. **27**, 532 (1982)].

Translated by M. E. Alferieff