Numerical analysis of scattering of the electrostatic field of a two-electrode cell at conducting surfaces

S. I. Safronov and R. P. Tarasov

Research Institute of Pulse Technology, 115304 Moscow, Russia

(Submitted November 4, 1997; resubmitted March 27, 1998)
Zh. Tekh. Fiz. **69**, 1–10 (June 1999)

A numerical analysis is made of the electrostatic field of a TEM cell and scattering of the initial field of the TEM cell by ideally conducting surfaces of revolution and by a parallelepiped.

INTRODUCTION

An integral part of the development, construction, and operation of a TEM cell[1] involves the construction of suitable numerical models: preliminary calculations of the field of a TEM cell having a particular profile can be used to assess the possibilities of this design and ultimately to synthesize a device which best satisfies these particular characteristics. The tolerances for the assembly of a real structure can also be determined. A numerical analysis of the scattering of the field of a TEM cell at conducting surfaces having a particular profile can also be used to estimate the perturbations of the initial field introduced by various measuring devices. Moreover, if the numerical analysis of this problem can be performed with a high degree of accuracy, it may be possible to calibrate the measuring devices to the extent of creating a standard. A numerical analysis of scattering problems is equally important for the direct operation of a TEM cell.

Under low-frequency excitation, the electromagnetic field of a TEM cell is quasistatic so that various measuring devices can be calibrated at given electrostatic potentials of the electrodes (casing and internal wiring) of the TEM cell. Consequently, in this case it is sufficient to make a numerical analysis of the electrostatic field of the TEM cell and scattering of the initial field by various conducting objects inserted in the TEM cell.

The geometry of a TEM cell is such that the numerical simulation of its electrostatic field and to a greater extent, the numerical analysis of scattering of the initial field at conducting surfaces are essentially three-dimensional problems whose solution presents various fundamental difficulties caused by the complex geometry and the diversity of the boundary surface formed by the electrodes of the TEM cell and the corresponding shields. In turn, discretizing a given boundary surface with acceptable accuracy necessitates constructing economical and stable algorithms for the numerical solution of mesh problems of very large dimensions. Hence, known calculation procedures are ineffective in this case and cannot be applied to make numerical analyses of the fields of a TEM cell with the required accuracy. As a consequence, in practice, three-dimensional models are usually rejected in favor of simpler, less accurate, two-dimensional qualitative models.

The methods described in Refs. 2–6 can be used for highly accurate numerical analyses of essentially three-dimensional problems. In the present paper we use these methods as the basis for a numerical simulation of the electrostatic field of a Crawford TEM cell[1,7] and also for a numerical analysis of the scattering of the fields of a TEM cell at a parallelepiped and at surfaces of revolution having a piecewise-smooth generatrix. The accuracy of the calculations is of the order of fractions of a percent.

1. NUMERICAL SIMULATION OF THE ELECTROSTATIC FIELD OF A TEM CELL

We shall consider the problem of determining the electrostatic field of a TEM cell whose fundamental geometry is shown in the conventional form in Figs. 1a and 1b and in general form in Fig. 1c. We shall assume that the internal electrode of the TEM cell forming the surface S_1 and the external electrode (forming the surface S_2) are at given potentials C_1 and C_2, respectively. Then, the potential of the electrostatic field of this system $v(x)$, where $x=\{x_i\}$ are Cartesian coordinates in the three-dimensional space V_3, is the solution of the Dirichlet problem for the Laplace equation

$$\Delta_3 v(x)=0, \quad \Delta_3 \equiv \frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2},$$

$$x \in V_3/S, \quad v(x)=f(x), \quad x \in S. \tag{1}$$

In this case we have $S=S_1 \cup S_2$, $f(x)=\chi_{S_i}(x)C_i$, $x \in S_i$, $i=1,2$, where χ_{S_i} is the characteristic function of the set of surface points S_i. We know that a solution of the boundary-value problem (1) can be constructed using the potential of a single layer. In this case, the surface function of the charge density $u(x)$, $x \in S$ satisfies the following system of first-order integral equations:

$$\sum_{j=1}^{2} A_{ij}u_j=C_i, \quad [A_{ij}\Psi](x)=\int \frac{\Psi(y)}{|x-y|}ds_y,$$

$$x \in S_i, \quad y \in S_j, \quad i,j=1,2, \tag{2}$$

where $|x-y|$ is the Euclidean distance between points x and y in the space V_3.

Methods of obtaining an approximate solution of an equation of the type (2) based on a numerical solution of

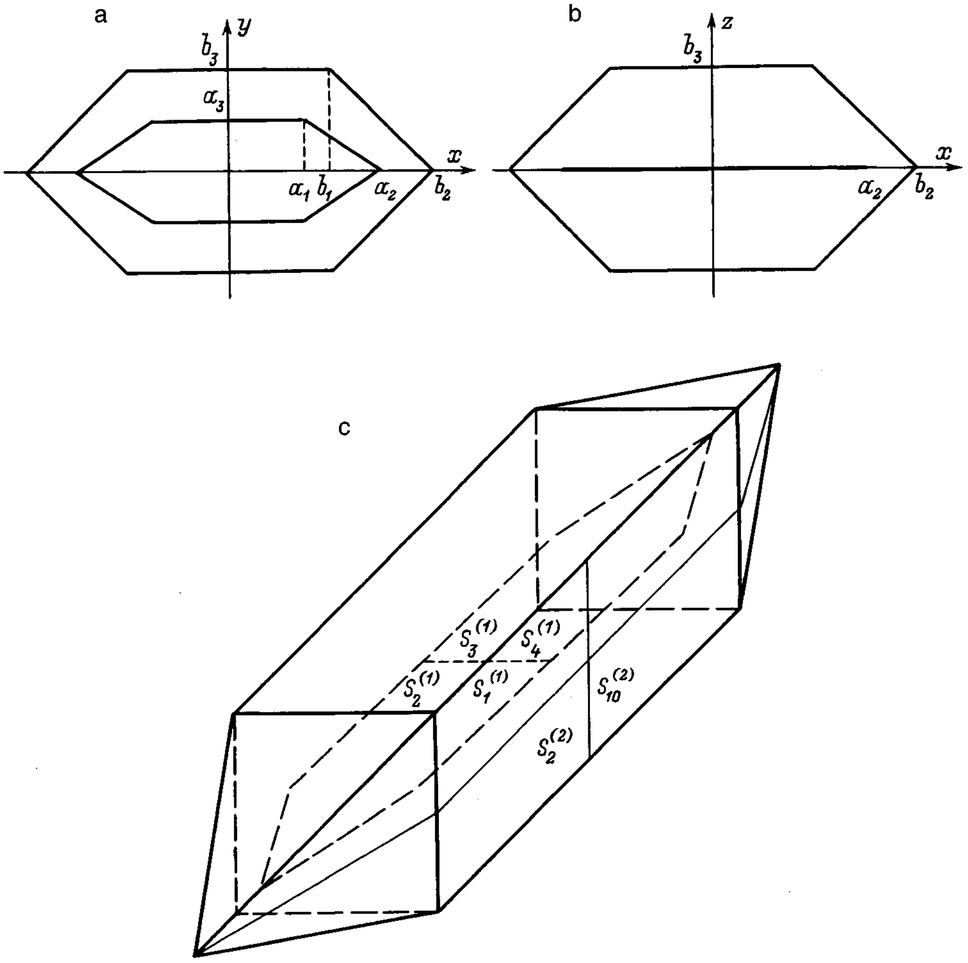


FIG. 1.

mesh problems of small dimensions M ($M \leqslant 500$) approximating an initial boundary-value equation are well-known.[8–10] For simply shaped surfaces S these methods demonstrated undoubted advantages over other approaches in a numerical implementation of the method of boundary-value integral equations. Moreover, since the number of arithmetical operations required to solve a mesh problem approximating an initial boundary-value equation is proportional to M^3, the computing costs increase substantially as the dimensions of the mesh problem increase. In addition, it is difficult to construct stable algorithms for the numerical inversion of mesh problems of large dimensions.

For boundary-value integral equations invariant with respect to symmetry transformations of some finite Abelian group $\{\tau_k\}$, $k=1,2,\ldots,N$, Demin and Tarasov[2] and Zakharov *et al.*[3] proposed methods of constructing a solution of the initial equation whose numerical implementation was significantly more stable and N^2 times more economical than conventional algorithms. These methods may be used for a numerical solution of Eq. (2) since the surface $S=S_1 \cup S_2$ can be partitioned into congruent components relative to the Klein fourth-order symmetry group. Allowance for this symmetry when constructing numerical algorithms can reduce the volume of calculations by a factor of 16 which, however, is insufficient to construct an approximate solution of the problem of scattering of the electrostatic field of a TEM cell at conducting surfaces with a high degree of accuracy.

The surface S_i (internal electrode) possesses an Abelian group of Klein symmetry transformations $\{\tau_k^{(1)}\}$ ($k=1,2,3,4$). The surface S_2 (external electrode) possesses a finite non-Abelian group of sixteenth-order symmetry transformations $\{\tau_k^{(2)}\}$ ($k=1,2,\ldots,16$). Thus, the symmetry of the surfaces S_1 and S_2 is best taken into account separately when constructing a solution of Eq. (2) using the highly efficient methods proposed in Refs. 2–5 for solving boundary-value integral equations with a commutative and noncommutative finite-order symmetry group.

Following Ref. 6, we shall seek a numerical solution of Eq. (2) using the iteration process

$$A_{11}u_1^{(n+1)}=(1-\omega)A_{11}u_1^{(n)}+\omega[C_1-A_{12}u_2^{(n)}],$$

$$A_{22}u_2^{(n+1)}=(1-\omega)A_{22}u_2^{(n)}+\omega[C_2-A_{21}u_1^{(n+1)}],$$

$$n=0,1,2,\ldots, \tag{3}$$

or, in a different form,

$$u_1^{(n+1)}=(1-\omega)u_1^{(n)}+\omega A_{11}^{-1}[C_1-A_{12}u_2^{(n)}],$$

$$u_2^{(n+1)}=(1-\omega)u_2^{(n)}+\omega A_{22}^{-1}[C_2-A_{21}u_1^{(n+1)}],$$

$$n=0,1,2,\ldots, \tag{3$'$}$$

where A_{11}^{-1} and A_{22}^{-1} are operators which are the inverse of A_{11} and A_{22}, respectively.

The invariance of the operators A_{11} and A_{22} relative to the end groups $\{\tau_k^{(1)}\}$ $(k=1,2,3,4)$ and $\{\tau_k^{(2)}\}$ $(k=1,2,\ldots,16)$ of the symmetry transformations of the surfaces S_1 and S_2 can be established directly. Consequently, the operators A_{11}^{-1} and A_{22}^{-1} can be constructed using the methods described in Refs. 2 and 3 and in Refs. 4 and 5, respectively.

Since algorithms for the inversion of operators invariant relative to transformations from the group of fourth-order Klein symmetries were analyzed in detail by Demin and Tarasov,[11] we shall merely give the final results here. Specifically, the surface S_1 can be represented in the form

$$S_1=\cup_{i=1}^{4} s_i^{(1)},\quad s_i^{(1)}\cap s_j^{(1)}=\emptyset \quad i\neq j,\quad i,j=1,2,3,4,$$

$$\tau_i^{(1)} s_1^{(1)}=s_i^{(1)},\quad i=1,2,3,4,\quad \mapsto \quad S_1=\cup_{i=1}^{4}\tau_i^{(1)} s_1^{(1)},$$

where $\tau_1^{(1)}=e$ is the identity transformation, $\tau_2^{(1)}$ and $\tau_4^{(1)}$ are the reflections relative to two orthogonal planes passing through the axes of rotation (in Fig. 1 the XZ and YZ planes), $\tau_3^{(1)}=\tau_2^{(1)}\tau_4^{(1)}$ is a rotation by the angle π (in Fig. 1 the Z axis is the axis of rotation), $\{s_i^{(1)}\}$ is a set of congruent components of the surface S_1 relative to the Abelian group $\{\tau_k^{(1)}\}$, $k=1,2,3,4$.

Demin and Tarasov[11] then showed that as a result of unitary transformations the operator $A_{11}=\|a_{ij}\|$ $(i,j=1,2,3,4)$ may be reduced to the diagonal form

$$A_{11} \quad \mapsto \quad \hat{A}_{11}=\mathrm{diag}(b_1,b_2,b_3,b_4), \tag{4}$$

where

$$b_1=a_1+a_2+a_3+a_4,\quad b_2=a_1-a_2-a_3+a_4,$$

$$b_3=a_1+a_2-a_3-a_4,\quad b_4=a_1-a_2+a_3-a_4,$$

$$a_i=a_{1i}T_i^{(1)},\quad T_i^{(1)}u(x)=u(\tau_i^{(1)^{-1}}x),\quad i=1,2,3,4,$$

$$[a_{ij}\Psi](x)=\int\frac{\Psi(y)}{|x-y|}ds_y,$$

$$x\in s_i^{(1)},\quad y\in s_j^{(1)},\quad i,j=1,2,3,4.$$

Thus, the algorithm for inversion of the operator A_{11} reduces to constructing the operators b_i^{-1}, which are the inverse of b_i, $i=1,2,3,4$. This procedure can reduce the number of calculations by a factor of 16 by using direct numerical methods of inverting numerical matrices.

The surface S_2 is described by the non-Abelian group of symmetry transformations $\{\tau_k^{(2)}\}$, $k=1,2,\ldots,16$, i.e., it can be partitioned as

$$S_2=\cup_{i=1}^{16} s_i^{(2)},\quad s_i^{(2)}\cap s_j^{(2)}=\emptyset,\quad i\neq j,\quad i,j=1,2,\ldots,16,$$

$$\tau_i^{(2)} s_1^{(2)}=s_i^{(2)},\quad i=1,2,\ldots,16,\quad \mapsto \quad S_2=\cup_{i=1}^{16}\tau_i^{(2)} s_1^{(2)}$$

into congruent components $\{s_i^{(2)}\}$ relative to the group $\{\tau_k^{(2)}\}$, $k=1,2,\ldots,16$, where

$$\tau_1^{(2)}=e,\quad \tau_2^{(2)}=\sigma_1,\quad \tau_3^{(2)}=C,\quad \tau_4^{(2)}=C\sigma_1,\quad \tau_5^{(2)}=C^2,$$

$$\tau_6^{(2)}=C^2\sigma_1,\quad \tau_7^{(2)}=C^3,\quad \tau_8^{(2)}=C^3\sigma_1,\quad \tau_9^{(2)}=\sigma_2,$$

$$\tau_{10}^{(2)}=\sigma_2\sigma_1,\quad \tau_{11}^{(2)}=\sigma_2 C,\quad \tau_{12}^{(2)}=\sigma_2 C\sigma_1,\quad \tau_{13}^{(2)}=\sigma_2 C^2,$$

$$\tau_{14}^{(2)}=\sigma_2 C^2\sigma_1,\quad \tau_{15}^{(2)}=\sigma_2 C^3,\quad \tau_{16}^{(2)}=\sigma_2 C^3\sigma_1,$$

and the transformations of the Euclidean space e, C, σ_1, and σ_2 are defined as follows: e is an identity transformation, C is a rotation through the angle $2\pi/4$ (the X axis is the axis of rotation in Fig. 1), σ_1 is a specular reflection operation relative to the plane passing through the axis of rotation, and σ_2 is a specular reflection relative to the plane orthogonal to the axis of rotation (the YZ plane in Fig. 1). Then, if the coordinate functions corresponding to the irreducible representations of the $\{\tau_k^{(2)}\}$ group are known, following Refs. 4 and 5, we can convert from the second equation (3) to the equations for its canonical representation relative to the $\{\tau_k^{(2)}\}$ group.

TABLE I.

$u_i(\tau_k)$	e	σ_1	C	$C\sigma_1$	C^2	$C^2\sigma_1$	C^3	$C^3\sigma_1$
u_1	1	1	1	1	1	1	1	1
u_2	1	-1	1	-1	1	-1	1	-1
u_3	1	1	-1	-1	1	1	-1	-1
u_4	1	-1	-1	1	1	-1	-1	1

Since the $\{\tau_k^{(2)}\}$ $(k=1,2,\ldots,16)$ groups can be represented as the direct product of two of its subgroups—a second-order Abelian group $\{e,\sigma_2\}$ and an eighth-order non-Abelian group (square group) $\{e,\sigma_1,C,C\sigma_1,C^2,C^2\sigma_1,C^3,C^3\sigma_1\}$, calculation of the coordinate functions $\{\tau_k^{(2)}\}$ reduces to calculating the coordinate functions of these two subgroups. The coordinate functions of the irreducible regular representations of an Abelian group are its characters. The characters of any Abelian group are known and for a second-order group these have the values 1 and -1 (for further details see Sec. 3). For its part, the square group $\{\tau_k\}$ $(k=1,2,\ldots,8)$ has three generatrices: $\tau_1=e$, $\tau_2=\sigma_1$, $\tau_3=C$, which allows us to calculate directly the coordinate functions $u_i(\tau_k)$, $i=1,2,\ldots,5$, corresponding to five irreducible regular representations of this group. The values of $u_i(\tau_k)$, $(i=1,2,\ldots,4,\ k=1,2,\ldots,8)$ are given in Table I.

The coordinate functions $u_5(\tau_k)$ $(k=1,2,\ldots,8)$ may be represented in the form

$$u_5(e)=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad u_5(\sigma_1)=\begin{pmatrix}1&0\\0&-1\end{pmatrix},$$

$$u_5(C)=\begin{pmatrix}0&-1\\1&0\end{pmatrix},\quad u_5(C\sigma_1)=\begin{pmatrix}0&-1\\-1&0\end{pmatrix},$$

$$u_5(C^2)=\begin{pmatrix}-1&0\\0&-1\end{pmatrix},\quad u_5(C^2\sigma_1)=\begin{pmatrix}-1&0\\0&1\end{pmatrix},$$

$$u_5(C^3)=\begin{pmatrix}0&1\\-1&0\end{pmatrix},\quad u_5(C^3\sigma_1)=\begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

The calculated values $u_i(\tau_k)$ can be used to write the equations of the canonical representation relative to the $\{\tau_k^{(2)}\}$ group for the second equation in (3)

$$(\tilde{A}_i+\tilde{B}_i)\tilde{u}_i(x)=\tilde{f}_i(x),\quad i=1,2,\ldots,10,\quad x\in s_1^{(2)}, \tag{5}$$

where $\tilde{A}_i=\tilde{A}_{i+5}$, $\tilde{B}_i=-\tilde{B}_{i+5}$, $i=1,2,\ldots,5$.

The following relations hold for the operators $\tilde{A}_i$ and $\tilde{B}_i$ $(i=1,2,\dots,5)$

$$\tilde{A}_1=\hat{A}_1+\hat{A}_2+\hat{A}_3+\hat{A}_4+\hat{A}_5+\hat{A}_6+\hat{A}_7+\hat{A}_8,$$

$$\tilde{A}_2=\hat{A}_1-\hat{A}_2+\hat{A}_3-\hat{A}_4+\hat{A}_5-\hat{A}_6+\hat{A}_7-\hat{A}_8,$$

$$\tilde{A}_3=\hat{A}_1+\hat{A}_2-\hat{A}_3-\hat{A}_4+\hat{A}_5+\hat{A}_6-\hat{A}_7-\hat{A}_8,$$

$$\tilde{A}_4=\hat{A}_1-\hat{A}_2-\hat{A}_3+\hat{A}_4+\hat{A}_5-\hat{A}_6-\hat{A}_7+\hat{A}_8,$$

$$\tilde{A}_5=\begin{bmatrix}\hat{A}_1+\hat{A}_2-\hat{A}_5-\hat{A}_6 & -\hat{A}_3-\hat{A}_4+\hat{A}_7+\hat{A}_8\\ \hat{A}_3-\hat{A}_4-\hat{A}_7+\hat{A}_8 & \hat{A}_1-\hat{A}_2-\hat{A}_5+\hat{A}_6\end{bmatrix}$$

$$\tilde{B}_1=\hat{A}_9+\hat{A}_{10}+\hat{A}_{11}+\hat{A}_{12}+\hat{A}_{13}+\hat{A}_{14}+\hat{A}_{15}+\hat{A}_{16},$$

$$\tilde{B}_2=\hat{A}_9-\hat{A}_{10}+\hat{A}_{11}-\hat{A}_{12}+\hat{A}_{13}-\hat{A}_{14}+\hat{A}_{15}-\hat{A}_{16},$$

$$\tilde{B}_4=\hat{A}_9+\hat{A}_{10}-\hat{A}_{11}-\hat{A}_{12}+\hat{A}_{13}+\hat{A}_{14}-\hat{A}_{15}-\hat{A}_{16},$$

$$\tilde{B}_4=\hat{A}_9-\hat{A}_{10}-\hat{A}_{11}+\hat{A}_{12}+\hat{A}_{13}-\hat{A}_{14}-\hat{A}_{15}+\hat{A}_{16},$$

$$\tilde{B}_5=\begin{bmatrix}\hat{A}_9+\hat{A}_{10}-\hat{A}_{13}-\hat{A}_{14} & -\hat{A}_{11}-\hat{A}_{12}+\hat{A}_{15}+\hat{A}_{16}\\ \hat{A}_{11}-\hat{A}_{12}-\hat{A}_{15}+\hat{A}_{16} & \hat{A}_9-\hat{A}_{10}-\hat{A}_{13}+\hat{A}_{14}\end{bmatrix}, \tag{6}$$

where

$$\hat{A}_i=T_i^{(2)}A_{i1},\quad T_i^{(2)}u(x)=u(\tau_i^{(2)^{-1}}x),\quad i=1,2,\dots,16,$$

$$[A_{ij}\Psi](x)=\int\frac{\Psi(y)}{|x-y|}ds_y,$$

$$x\in s_i^{(2)},\quad y\in s_j^{(2)},\quad i,j=1,2,\dots,16. \tag{7}$$

Converting from the second equation to Eqs. (5)–(7) reduces the volume of calculations for the construction of A_{22}^{-1} by a factor of $\cong 170$. In addition, this procedure for constructing the inverse to the mesh operator approximating A_{22} has a significantly higher calculation stability compared with existing procedures.

By implementing the iteration process ($3'$) with the operators A_{11}^{-1} and A_{22}^{-1} constructed using relations (4) and (5)–(7), it was possible to obtain stable, highly accurate solutions of mesh problems which approximate the equations (2) with a discretization order $M\cong 10^4$. Therefore, values of the distribution density function of the electrostatic charges $u(x)$ at the electrodes of the TEM cell S_1 and S_2 can be calculated with an error not exceeding fractions of percent and consequently at any point x in the Euclidean space V_3 the potential $v(x)$ and the intensity vector $E(x)=-\mathrm{grad}\,v(x)$ of the electrostatic field of the TEM cell can be determined by

$$v(x)=\int\frac{u(y)}{|x-y|}ds_y,\quad E(x)=-\int \mathrm{grad}_x\frac{u(y)}{|x-y|}ds_y,$$

$$x\in V_3\backslash S,\quad y\in S=S_1\cup S_2.$$

The algorithms for calculating the electrostatic field of a TEM cell described in the present section can be used to determine the position and required characteristics of the working volumes of the TEM cell efficiently and highly accurately, and the procedure for determining these working volumes can be adequately visualized.

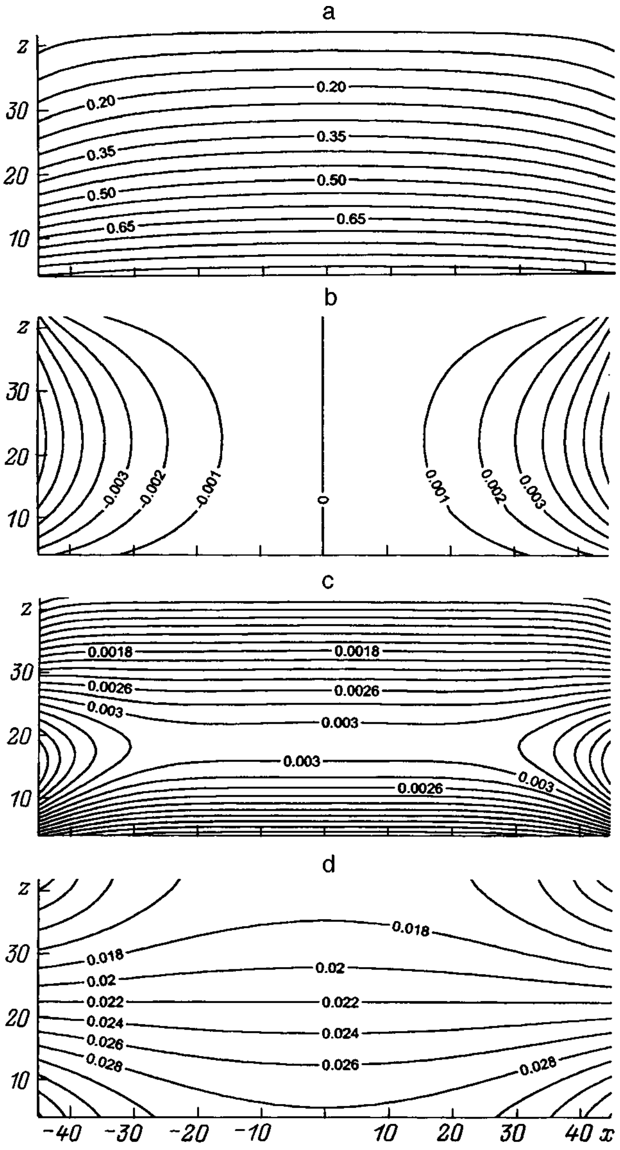


FIG. 2.

Figures 2 and 3 show isolines for the potential $v(x,y,z)$ (where x, y, and z are the coordinates of the point in the XYZ Cartesian system shown in Fig. 1) and the components $E_x(x,y,z)$, $E_y(x,y,z)$, and $E_z(x,y,z)$ of the intensity vector $E(x,y,z)$ of the electrostatic field of the TEM cell shown in Fig. 1. A unit potential was applied to the internal electrode S_1 while the potential at the external electrode (the surface S_2) was assumed to be zero. Figure 2a gives the equipotential distribution in the rectangle $y=0.2b_3$, $x\in[-a_1,a_1]$, $z\in[0,b_3]$, while Fig. 3a gives that in the rectangle $x=0.2b_3$, $y\in[-b_3,b_3]$, $z\in[0,b_3]$. Figures 2b and Fig. 3b show the isolines for $E_x(x,y,z)$ in the same rectangles; Figs. 2c and 3c give those for $E_y(x,y,z)$, while Figs 2d and 3d give those for $E_z(x,y,z)$.

In these calculations the number of nodes on the boundary surface $S=S_1\cup S_2$ was 8000, i.e., the initial boundary-

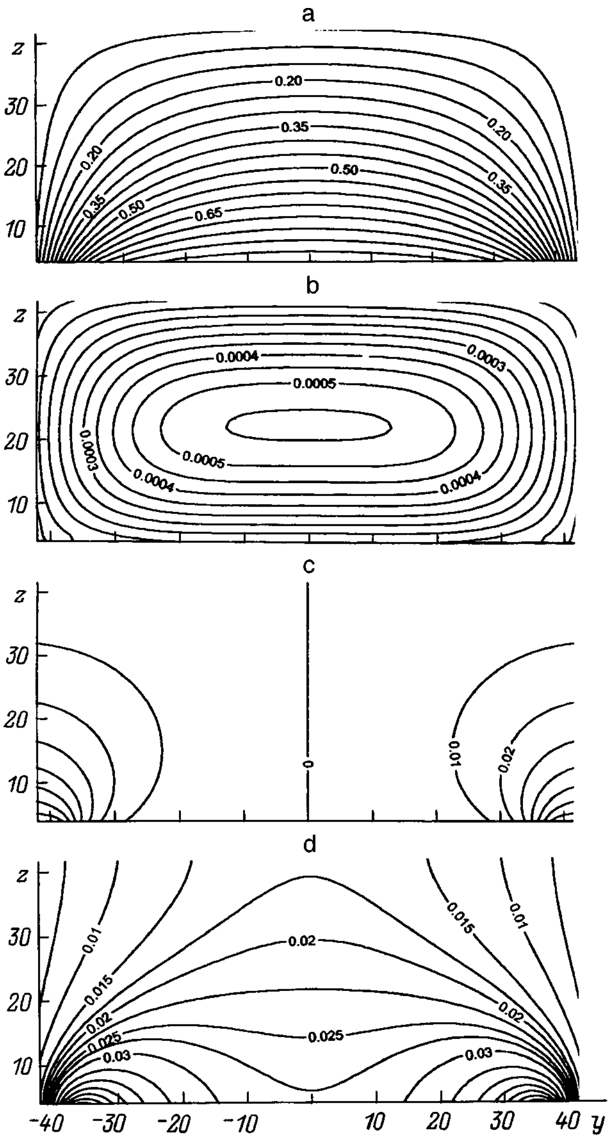


FIG. 3.

value integral operator (2) was approximated by a densely filled (8000×8000) matrix.

The components of the electrostatic field intensity vector $E_x(x,y,z)$ and $E_y(x,y,z)$ are usually neglected when developing TEM cells since it is assumed that their values are negligible compared with $E_z(x,y,z)$. However, no reliable quantitative estimates have been obtained to confirm this.

A comparison between the isolines shown in Fig. 2 with the $v(x,y,z)$, $E_x(x,y,z)$, and $E_z(x,y,z)$ isolines obtained in the rectangle $y=0$, $x\in[-a_1,a_1]$, $z\in[0,b_3]$ (in this case we have $E_y(x,y,z)\equiv 0$) revealed qualitative differences. This also applies to the $v(x,y,z)$, $E_y(x,y,z)$, and $E_z(x,y,z)$ isolines obtained in the rectangle $x=0$, $y\in[-b_3,b_3]$, $z\in[0,b_3]$ and the isolines shown in Fig. 3. The values of the electrostatic field potential $v(x,y,z)$ and the corresponding components of the intensity vector $E_x(x,y,z)$ at the points $(x,0,z)$ and $(x,0.2b_3,z)$ for $x\in[-a_1,a_1]$ and $z\in[0,b_3]$ and also at the points $(0,y,z)$ and $(0.2b_3,y,z)$ for $y\in[-b_3,b_3]$ and $z\in[0,b_3]$ differed by a few percent. This indicates that under certain conditions, a simulation of the fields of a TEM cell based on electrostatic problems on a plane can give a qualitatively (but not quantitatively) true picture of the potential distribution and the component of the electrostatic field intensity vector in the XY and XZ symmetry planes, although this is very approximate for simulating the scattering of the field of a TEM cell at conducting surfaces.

2. SCATTERING OF THE ELECTROSTATIC FIELD OF A TEM CELL AT A CONDUCTING SURFACE OF REVOLUTION

The problem of determining the perturbation of the field of a TEM cell caused by various conducting objects inserted in the cell is of major practical interest. In this section we analyze the scattering of the field in a TEM cell by infinitely conducting shields and objects comprising various surfaces of revolution (such as cylinders, cones, spheres, ellipsoids of revolution, and so on). In this case, we need to find the solution of the boundary-value problem (1) whose set of boundary points forms a surface S having the following form: $S=S_1\cup S_2\cup S_3$, where S_1 and S_2 are the electrodes of the TEM cell considered previously and S_3 is some surface of revolution, given that the potential $v(x)$ of the electrostatic field of this system has the following values at the boundary S:

$$v(x)=f(x),\quad x\in S,\quad f(x)=\chi_{S_i}(x)C_i,$$

$$x\in S_i,\quad i=1,2,3,$$

where C_1 and C_2 are given constants and C_3 is a generally unknown constant.

By analogy with Sec. 1, we convert from the boundary-value problem to an analysis of the following system of boundary-value integral equations:

$$\sum_{j=1}^{3} A_{ij}u_j=C_i,\quad [A_{ij},\Psi](x)=\int\frac{\Psi(y)}{|x-y|}ds_y,$$

$$x\in S_i,\quad y\in S_j,\quad i,j=1,2,3. \tag{8}$$

We shall construct a numerical solution of the equations (8) on the basis of the iteration process

$$A_{11}u_1^{(n+1)}=(1-\omega)A_{11}u_1^{(n)}+\omega[C_1-A_{12}u_2^{(n)}-A_{13}u_3^{(n)}],$$

$$A_{22}u_2^{(n+1)}=(1-\omega)A_{22}u_2^{(n)}+\omega[C_2-A_{21}u_1^{(n+1)}-A_{23}u_3^{(n)}],$$

$$A_{33}u_3^{(n+1)}=(1-\omega)A_{33}u_3^{(n)}+\omega[C_3-A_{31}u_1^{(n+1)}-A_{32}u_2^{(n+1)}],$$

$$n=0,1,2,\ldots. \tag{9}$$

We shall consider the construction of an operator A_{33}^{-1} the inverse of A_{33}, since the operators A_{11}^{-1} and A_{22}^{-1} are constructed using relations (4) and (5)–(7), respectively.

The surface of revolution S_3 is described by the C_∞ group of symmetry transformations of Euclidean space and consequently it has a finite Abelian subgroup C_n of the C_∞ group for any n $(n=1,2,3,\ldots)$; C_n is a cyclic group of symmetry transformations having the following form:

$\{e,\tau,\ldots,\tau^{(n-1)}\}$, where τ is a rotation through the angle $2\pi/n$ with the axis of rotation matched with the symmetry axis of the surface of revolution. Since C_n is a finite Abelian group of symmetry transformations of the surface S_3 and the operator A_{33} is invariant with respect to transformations of the motion from the symmetry group of the boundary surface, the operator A_{33}^{-1} is best constructed using the methods described by Demin and Tarasov[2] and Zakharov *et al.*[3]

By virtue of the symmetry properties, the surface S_3 can be partitioned as follows

$$S_3=\cup_{i=1}^{n}s_i^{(3)},\quad s_i^{(3)}\cap s_j^{(3)}=\emptyset,\quad i\neq j,\quad i,j=1,2,\ldots,n,$$

$$\tau^{i-1}s_i^{(3)}=s_1^{(3)},\quad i=1,2,\ldots,n,\ \mapsto\ S_3=\cup_{i=1}^{n}\tau^{n-i+1}s_1^{(3)}$$

into congruent components $\{s_i^{(3)}\}$ relative to the C_n group. Demin and Tarasov[2] and Zakharov *et al.*[3] also showed that as a result of known unitary transformations, the operator $A_{33}=\|a_{ij}\|$ $(i,j=1,2,\ldots,n)$ can be reduced to the diagonal form[12]

$$A_{33}\quad\mapsto\quad \hat{A}_{33}=\mathrm{diag}(b_1,b_2,\ldots,b_n),\tag{10}$$

where

$$b_i=\sum_{j=1}^{n}a_j\varepsilon^{n-(j-1)(i-1)},\quad i=1,2,\ldots,n,$$

$$\varepsilon=\exp(2\pi I/n),\quad I^2=1,$$

$$a_j=a_{1j}T^{n-j+1},\quad T^{n-j+1}u(x)=u(\tau^{j-1}x),\quad j=1,2,\ldots,n,$$

$$[a_{ij}\Psi](x)=\int\frac{\Psi(y)}{|x-y|}ds_y,$$

$$x\in s_i^{(3)},\quad y\in s_j^{(3)},\quad i,j=1,2,\ldots,n.$$

Constructing A_{33}^{-1} using the relations (10) can reduce the number of arithmetic operations involved in inverting the mesh operators approximating S_{33} by a factor of $\cong n^2$. For instance, for $n=100$ the number of operations will be reduced $\cong 10^4$ times when direct methods are used to invert the numerical matrices.

Computational experiments showed that if the value of the constant C_3 is known (the potential of the surface of revolution), the iteration algorithm (9) where the operators $A_{11}^{(-1)}$, $A_{22}^{(-1)}$, and $A_{33}^{(-1)}$ are calculated using relations (4), (5)–(7), and (10), respectively, can be used to determine the electrostatic field created by a system of electrodes $S=S_1\cup S_2\cup S_3$ having given potentials C_1, C_2, and C_3 with an error not exceeding fractions of percent.

We shall consider the scattering of the electrostatic field of a TEM cell by an infinitely conducting scatterer. We know that this problem can be solved approximately by means of a numerical solution of the equations (8) supplemented by the following condition (charge conservation condition):

$$\int u(y)ds_y=0,\quad y\in S_3.\tag{11}$$

Condition (11) can naturally be incorporated in the iteration process (9) to determine the constant C_3 without any real increase in the number of arithmetic operations, i.e.,

$$A_{11}u_1^{(n+1)}=(1-\omega)A_{11}u_1^{(n)}+\omega[C_1-A_{12}u_2^{(n)}-A_{13}u_3^{(n)}],$$

$$A_{22}u_2^{(n+1)}=(1-\omega)A_{22}u_2^{(n)}+\omega[C_2-A_{21}u_1^{(n+1)}-A_{23}u_3^{(n)}],$$

$$A_{33}\tilde{u}_3^{(n+1)}=(1-\omega)A_{33}\tilde{u}_3^{(n)}-\omega[A_{31}u_1^{(n+1)}+A_{32}u_2^{(n+1)}],$$

$$u_3^{(n+1)}=C_3^{(n+1)}\hat{u}_3+\tilde{u}_3^{(n+1)},\quad C_3^{(n+1)}=-\frac{\int\tilde{u}_3^{(n+1)}ds_y}{\int\hat{u}_3ds_y},$$

$$n=0,1,2,\ldots,\tag{12}$$

where $\hat{u}_3$ is the solution of $A_{33}\hat{u}_3=1$.

It is easy to see that implementing the iteration process (12) requires almost no increase in the number of operations compared with process (9), i.e., the iteration process (12) has the same advantages for solving scattering problems as the iteration process (9) used to determine the electrostatic field created by the surface $S=S_1\cup S_2\cup S_3$. We stress that for iteration processes (9) and (12) the relative spatial configuration of the surfaces S_1, S_2, and S_3 is arbitrary.

We shall analyze the scattering of the electrostatic field of a TEM cell (Figs. 2 and 3) at a cylinder of height $H=0.065b_3$ and radius $R=0.18b_3$. Approximate solutions of this problem for various positions of the cylinder inside the TEM cell were obtained using the iteration process (12). For instance, Fig. 4a shows field potential isolines $v(x,y,z)$ of this system in the rectangle $y=0$, $x\in x[-a_1,a_1]$, $z\in[0,b_3]$. Isolines for the component of the electrostatic field intensity vector $E_z(x,y,z)$ in this rectangle are shown in Fig. 4b while Fig. 4c gives $E_z(x,y,z)$ isolines in the rectangle

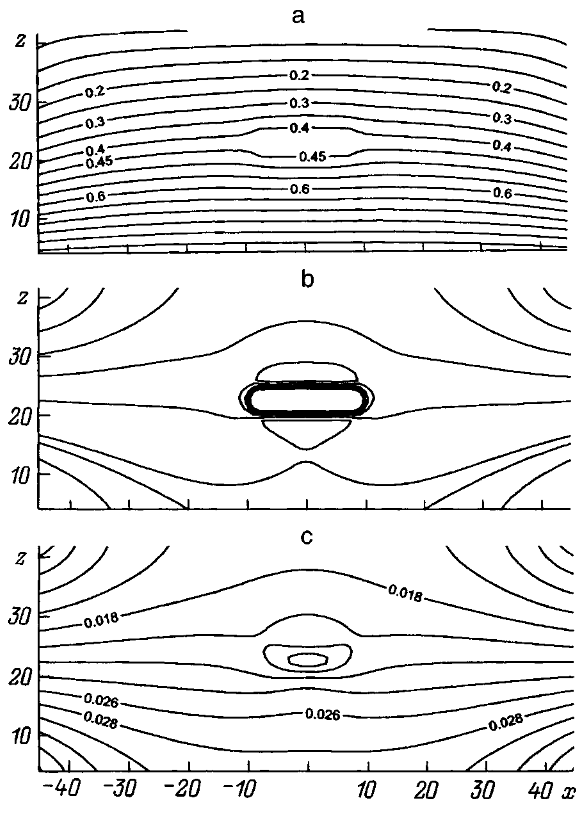


FIG. 4.

$y=0.2b_3$, $x\in[-a_1,a_1]$, $z\in[0,b_3]$. In this case, the axis of rotation of the cylinder coincided with the Z axis (Figs. 1a and 1b) and the distance between the internal electrode of the TEM cell and the lower end of the cylinder was $h=0.45b_3$.

Let us assume that this cylinder is positioned as follows: the axis of rotation of the cylinder is parallel to the Y axis and intersects the Z axis at the height $h=0.35b_3$ and one of the ends of the cylinder lies in the XZ plane. For this cylinder configuration Fig. 5a gives the isolines characterizing the potential distribution $v(x,y,z)$ in the rectangle $x=0$, $y\in[-b_3,b_3]$, $z\in[0,b_3]$. Figure 5b shows the $E_z(x,y,z)$ isolines in the same cross section and Fig. 5c gives the $E_z(x,y,z)$ isolines in the rectangle $x=0.2b_3$, $y\in[-b_3,b_3]$, $z\in[0,b_3]$.

By comparing Figs. 4 and 5 with Figs. 2 and 3 we can determine some characteristic features of the perturbations of the electrostatic field of a TEM cell caused by the presence of an ideally conducting cylinder in some part of this cell. Since the electrostatic field is potential and the ideally conducting cylinder is an equipotential surface, we can talk of the field flowing round the surface of the cylinder whose nature can be determined directly from Figs. 4 and 5 and the degree of perturbation of the electrostatic field of the TEM cell caused by the cylinder and its rate of attenuation in space can easily be assessed.

In these calculations the number of nodes on the boundary surface $S=S_1\cup S_2\cup S_3$ was 9000, i.e., the initial boundary-value integral operator (8) was approximated by a densely filled (9000×9000) matrix.

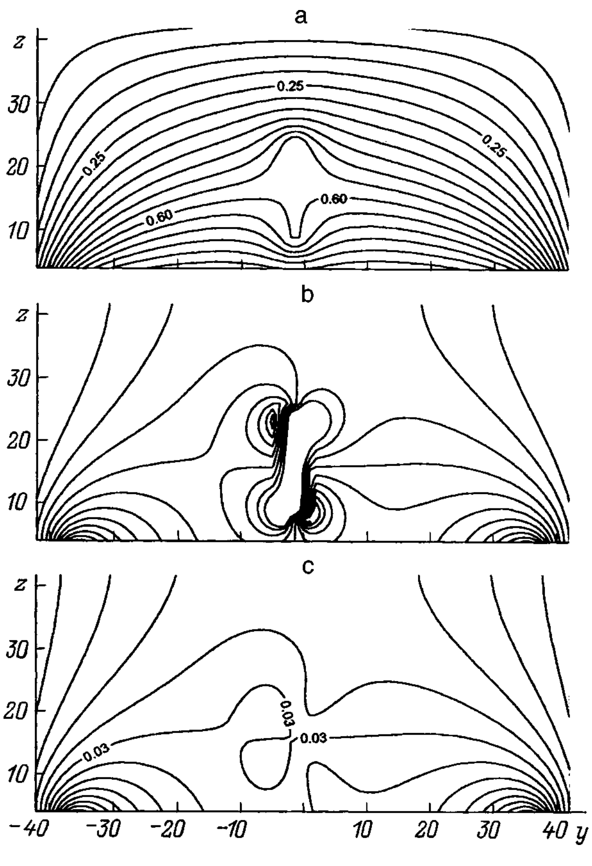


FIG. 5.

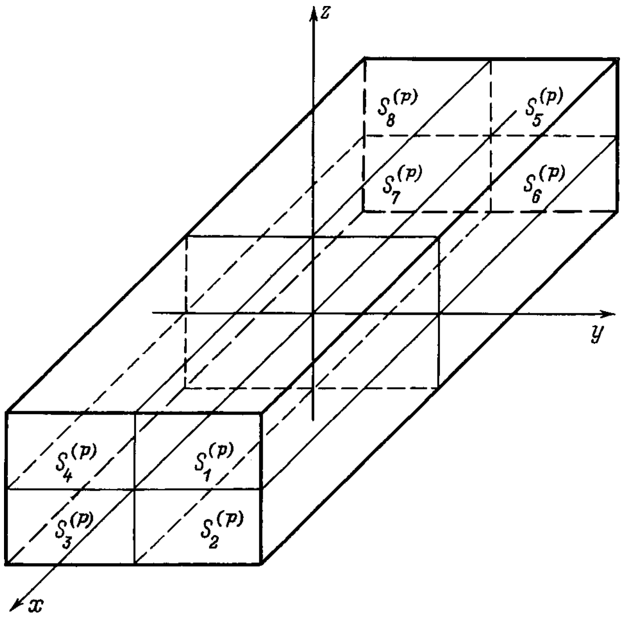


FIG. 6.

To conclude this section, we note that we have confined ourselves to developing programs to simulate the scattering of the field of a TEM cell at surfaces of revolution having a piecewise-smooth generatrix represented as a finite combination of line segments and arcs of circles (closed and open cylindrical and conical surfaces, spheres, and so on). However, this program can also be used as the basis for analyzing surfaces of revolution with any types of generatrices since these algorithms implement a specific type of symmetry and are not related to a specific surface profile. This implies that the development of programs to solve the problem of scattering of the field of a TEM cell for surfaces of revolution having a specific type of generatrix can be reduced to a program describing merely a class of these surfaces.

3. SCATTERING OF THE ELECTROSTATIC FIELD OF A TEM CELL AT A PARALLELEPIPED

If a parallelepiped (Fig. 6) is taken as the component S_3 of the boundary surface $S=S_1\cup S_2\cup S_3$ considered in the previous section, approximate values of the electrostatic field potential for known constants C_1, C_2, and C_3 may be obtained using the iteration process (9). Then, the problem of scattering of the electrostatic field of a TEM cell at an infinitely conducting parallelepiped can be solved numerically using the iteration process (12). In addition, calculations based on the iteration algorithms (9) or (12) can be predicted to have a high accuracy provided that effective algorithms are available for inverting the operator A_{33}

TABLE II.

$\chi_i(\tau_j)$	$\tau_1=e$	$\tau_2=\sigma_z$	$\tau_3=\sigma_y\sigma_z$	$\tau_4=\sigma_y$	$\tau_5=\sigma_x$	$\tau_6=\sigma_x\sigma_y$	$\tau_7=\sigma_x\sigma_y\sigma_z$	$\tau_8=\sigma_x\sigma_z$
χ_1	1	1	1	1	1	1	1	1
χ_2	1	1	-1	-1	1	1	-1	-1
χ_3	1	-1	-1	1	1	-1	-1	1
χ_4	1	-1	1	-1	1	-1	1	-1
χ_5	1	1	1	1	-1	-1	-1	-1
χ_6	1	1	-1	-1	-1	-1	1	1
χ_7	1	-1	-1	1	-1	1	1	-1
χ_8	1	-1	1	-1	-1	1	-1	1

$$[A_{33}\Psi](x)=\int\frac{\Psi(y)}{|x-y|}ds_y,\quad x,y\in S_3,$$

where S_3 is a parallelepiped.

The surface S_3 is described by the Abelian symmetry group $\{\tau_k^{(3)}\}$ $(k=1,2,\dots,8)$ (Fig. 6) and the invariance of the operator A_{33} to transformations of the motion of the Euclidean space from the eighth-order group $\{\tau_k^{(3)}\}$ can be established directly. Zakharov *et al.*[13] analyzed three existing abstract eighth-order Abelian groups and obtained canonical representations of the operators invariant with respect to transformations from these groups. The results[13] form the basis of the algorithm for constructing the operator A_{33}^{-1} which is the inverse of A_{33}.

The eighth-order Abelian group of symmetry transformations of the surface S_3 $\{\tau_k^{(3)}\}$ has the following form: $\{e,\sigma_z,\sigma_y\sigma_z,\sigma_y,\sigma_x,\sigma_x\sigma_z,\sigma_x\sigma_y\sigma_z,\sigma_x\sigma_z\}$, where σ_x, σ_y, and σ_z are reflections relative to the three pairwise orthogonal planes $\{YZ\}$, $\{XZ\}$, and $\{XY\}$, respectively (Fig. 6). Since $\sigma_x^2=\sigma_y^2=\sigma_z^2=e$, e is an identity transformation, the group $\{\tau_k^{(3)}\}$ is the direct product of its three Abelian subgroups $\{e,\sigma_x\}$, $\{e,\sigma_y\}$, and $\{e,\sigma_z\}$. Since the subgroups $\{e,\sigma_x\}$, $\{e,\sigma_y\}$, and $\{e,\sigma_z\}$ are cyclic groups, the tables of their characters have the same form: $\chi_1(e)=1$, $\chi_2(e)=1$, $\chi_1(\tau)=1$, $\chi_2(\tau)=-1$, and $\tau=\{\sigma_x,\sigma_y,\sigma_z\}$, which can be used to directly calculate the characters $\chi_i(\tau_k^{(3)})$ of the $\{\tau_k^{(3)}\}$, $k=1,2,\dots,8$ group (Table II).

The surface S_3 can then be partitioned as

$$S_3=\cup_{i=1}^{8}s_i^{(p)},\quad s_i^{(p)}\cap s_j^{(p)}=\emptyset,\quad i\neq j,$$

$$\tau_i^{(3)}s_1^{(p)}=s_i^{(p)},$$

$$i=1,2,\dots,8,\quad\mapsto\quad S_3=\cup_{i=1}^{8}\tau_i^{(3)}s_1^{(p)}$$

into congruent components $\{s_k^{(p)}\}$ relative to the $\{\tau_k^{(3)}\}$; $k=1,2,\dots,8$ group. Taking into account these relations for each n $(n=0,1,2,\dots)$, we rewrite the third equation of the iteration processes (9) and (12) in the form

$$\sum_{j=1}^{8}A_{ij}U_j(x)=F_i(x),\quad i=1,2,\dots,8,\quad x\in s_1^{(p)},\tag{13}$$

where the following notation is introduced for the third equation (9)

$$U_j(x)=u_3^{(n+1)}(\tau_j^{(3)},x),\quad u_3^{(n+1)}(\tau_j^{(3)},x)=u_3^{(n+1)}(\tau_j^{(3)}x),$$

$$x\in s_1^{(p)},\quad j=1,2,\dots,8,$$

$$F_i(x)=f(\tau_i^{(3)},x),\quad f(\tau_i^{(3)},x)=f(\tau_i^{(3)}x),$$

$$x\in s_1^{(p)},\quad i=1,2,\dots,8,$$

$$f(x)=(1-\omega)A_{33}u_3^{(n)}(x)+\omega[C_3-A_{31}u_1^{(n+1)}(x)$$
$$-A_{32}u_2^{(n+1)}(x)],\quad x\in S_3,$$

$$[A_{ij}\Psi](x)=\int\frac{\Psi(\tau_j^{(3)}y)}{|\tau_i^{(3)}x-\tau_j^{(3)}y|}ds_y,$$

$$x,y\in s_1^{(p)},\quad i,j=1,2,\dots,8.$$

Similar notation is suitably introduced for the third equation in the iteration process (9).

The system (13) can be converted to eight independent equation by means of unitary transformations

$$A(\hat{\sigma}_i)U(\hat{\sigma}_i)=F(\hat{\sigma}_i),\quad i=1,2,\dots,8,\tag{14}$$

where

$$A(\hat{\sigma}_i)=\sum_{j=1}^{8}A_{j1}\chi_i(\tau_j^{(3)}),\quad U(\hat{\sigma}_i)=\sum_{j=1}^{8}U_j\chi_i(\tau_j^{(3)}),$$

$$F(\hat{\sigma}_i)=\sum_{j=1}^{8}F_j\chi_i(\tau_j^{(3)}),\quad i=1,2,\dots,8.$$

If the solution of the equations (14) $U(\hat{\sigma}_i)$ $(i=1,2,\dots,8)$ is known, the solution of the system (13) can be reconstructed as follows:

$$U_i=\sum_{j=1}^{8}U(\hat{\sigma}_j)\chi_j(\tau_i^{(3)}),\quad i=1,2,\dots,8.$$

Hence, the solution of the third equation in the iteration processes (9) or (12) is reduced to solving the equations (14), i.e., inversion of the operator A_{33} is replaced by constructing $A^{-1}(\hat{\sigma}_i)$ operators the inverse of $A(\hat{\sigma}_i)$, $i=1,2,\dots,8$. This procedure can reduce the volume of calculations required to construct the mesh operators approximating A_{33}^{-1} by a factor of $\cong 64$. The problem of scattering of an electrostatic field by an arbitrarily oriented parallelepiped can therefore be solved with a high degree of accuracy by means of the iteration process (12) where the operators A_{11}, A_{22}, and A_{33} are inverted in accordance with Eqs. (4), (5)–(7), and (14).

We shall now consider the scattering of the electrostatic field of a TEM cell (Figs. 2 and 3) at a parallelepiped having the following typical dimensions: $a_x=0.18b_3$, $b_y=0.18b_3$, and $c_z=0.0325b_3$, i.e., this parallelepiped is a square plate whose base length is equal to the diameter and whose thickness is equal to the thickness of the cylinder described in the previous section. This plate was successively inserted in the same volumes as the cylinder. The electrostatic fields of these two systems with the parallelepiped was calculated for 11 500 nodes suitably selected on the boundary surface $S=S_1\cup S_2\cup S_3$. Graphs of the isolines in the same cross sections as in Figs. 4 and 5 were then plotted for the electrostatic field potential $v(x,y,z)$ and the corresponding components of the intensity vector $E(x,y,z)$. A comparison between these graphs of the isolines of the electrostatic field of a TEM cell scattered by a parallelepiped and those of the field isolines scattered by a cylinder (Figs. 4 and 5) revealed a definite qualitative similarity.

Zakharov *et al.*[13] considered quadrupole systems which have the same symmetry group as a parallelepiped, i.e., the class of surfaces having this type of symmetry which is of interest for practical applications is fairly extensive. Problems involving the scattering of the electrostatic field of a TEM cell at surfaces in this class can then be solved using these algorithms. It is only necessary to obtain a program describing this class of surfaces which does not present any fundamental difficulties and requires relatively little computing time.

To conclude we note that these algorithms may be taken as the basis for developing schemes for obtaining numerical solutions of problems involving the scattering of electrostatic fields of known types of TEM cells at various systems of closed and open shields.

[1] M. L. Crawford, IEEE Trans. Electromagn. Compat. **EMC-16**(4), 189 (1974).

[2] S. K. Demin and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **29**, 1308 (1989).

[3] E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Dokl. Akad. Nauk SSSR **314**, 589 (1990) [Sov. Phys. Dokl. **35**, 799 (1990)].

[4] R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **31**, 1515 (1992).

[5] R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **33**, 1815 (1993).

[6] E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Z. Phys. Chem. (Munich) **33**, 1030 (1993).

[7] C. H. Wan, IEEE Trans. Electromagn. Compat. **35**, 109 (1993).

[8] M. A. Jaswon and G. T. Symm, Proc. R. Soc. London, Ser. A **275**, 23 (1963).

[9] M. A. Jaswon and A. R. Ponter, Proc. R. Soc. London, Ser. A **273**, 237 (1963).

[10] R. F. Harrington, *Field Computation by Moment Methods* (Macmillan, New York, 1968).

[11] S. K. Demin and R. P. Tarasov, Mat. Model. **5**(7), 113 (1993).

[12] E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **31**, 40 (1992).

[13] E. V. Zakharov, S. I. Safronov, and R. P. Tarasov, Zh. Vychisl. Mat. Mat. Fiz. **30**, 1631 (1990).

Translated by R. M. Durham

Efficiency of reverse cycles

Yu. B. Barkovskiĭ, E. N. Moos, and N. I. Tabunov

P. A. Kostychev Ryazan State Agricultural Academy, 390044 Ryazan, Russia

(Submitted January 5, 1997; resubmitted April 22, 1998)
Zh. Tekh. Fiz. **69**, 11–14 (June 1999)

It is demonstrated that the generally accepted treatment of the second law of thermodynamics is incorrect when the statement of the maximum limiting efficiency of the forward Carnot cycle is applied to the reverse cycles. An analysis of reversible cycles compared with the Carnot cycle shows that the reverse Carnot cycle has the lowest efficiency of all reverse cycles. A new characteristic is proposed — the thermal efficiency of reverse cycles, and a generalized theorem is put forward for the additivity of the thermal efficiencies of the forward and reverse cycles. A formulation of the second law of thermodynamics is suggested from the point of view of the efficiency of reverse cycles.
[S1063-7842(99)00206-8]

INTRODUCTION

Erroneous statements and conclusions are frequently encountered in treatments of the principles of thermodynamics.[1] In particular, the thermal efficiency of reverse equilibrium processes and the position of the Carnot cycle in this group of cyclic processes have yet to be determined. One formulation of the second law of thermodynamics indicates that some limit exists for thermal efficiency whose boundary is determined by the most efficient forward Carnot cycle.

In the forward and reverse cycles the conversion of some types of energy into others is considered under conditions where the conservation law is satisfied. Nonequilibrium is allowed for by a correction in analyses of processes for real cycles. The efficiency and direction of energy conversion can be formulated in terms of efficiency, and in terms of entropy. Thus, analyses of forward and reverse cycles may be useful in chemical and technical thermodynamics.

Reversible cycles (subsequently we shall not specifically refer to the condition of reversibility) are compared with the Carnot cycle as the standard[2] which is considered to be the most favorable in the forward and reverse forms. It is postulated[3–6] that the reverse Carnot cycle has superior economic characteristics, i.e., cooling and heating coefficients. This is demonstrated either on the basis of fallacious premises or the conditions of comparison are taken to be arbitrary.[7–9] The statement (see Refs. 3 and 5) that the Carnot cycle has the highest cooling coefficient compared with other reverse cycles at given heat-source temperatures is incorrect. In particular, the following inequality[9] is assumed to hold for reverse cycles

$$Q_2/W_c < T_2/(T_1 - T_2),$$

where $Q_2/W_c = \varepsilon$ is the cooling coefficient, W_c is the work of the cycle, and T_1 and T_2 are the temperatures of the hot and cold heat sources.

However, this expression contradicts the ''principle of the heat engine '' (the second Carnot theorem), i.e., the inequality

$$W_c/Q_1 < (T_1 - T_2)/T_1.$$

At the same time, the relationship between the cooling coefficient and the thermal efficiency is such[10] that the more efficient the forward cycle, the less efficient will be the reverse cycle.

It is difficult to envisage the position of the Carnot cycle among the other reverse cycles because of the absence of any characteristic similar to the thermal efficiency of forward cycles. Thus, an explanation of the impossibility of spontaneous concentration of energy using a heat pump in violation of the second law of thermodynamics is of no significance.[11]

It is impossible to formulate the second law of thermodynamics in terms of the existing efficiency characteristics of reverse cycles. As will be shown, the statement that the Carnot cycle is the most efficient of the forward cycles cannot be transferred ''mechanically'' to reverse cycles.

COMPARISON OF REVERSE CYCLES

For the analysis we give the well-known theorems.

First Carnot theorem: the thermal efficiency $\widehat{\eta}_C$ of the forward Carnot cycle depends only on the temperatures of the heat sources

$$\widehat{\eta}_C = 1 - T_2/T_1.$$

Second Carnot theorem: the forward Carnot cycle has the highest thermal efficiency of any cycle for given temperatures T_1 and T_2

$$\widehat{\eta}_C > \eta^{\frown}. \qquad (1)$$

The efficiency of reverse cycles is characterized by the heating coefficient ψ

$$\psi = Q_1/W_C = Q_1/(Q_1 - Q_2).$$

Since the thermal efficiency of an arbitrary forward cycle is

$$\widehat{\eta} = (Q_1 - Q_2)/Q_1,$$

it follows that

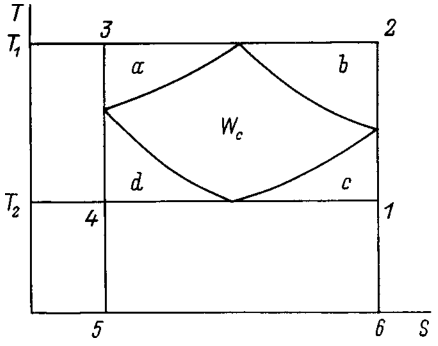


FIG. 1.

$$\psi = 1/\hat{\eta}.$$

The second Carnot theorem indicates that $1/\psi_c > 1/\psi$ or

$$\psi_c < \psi. \tag{2}$$

Thus, any reverse cycle with given temperature limits has a higher heating coefficient that the corresponding Carnot cycle. The same result can be achieved by comparing the Carnot cycle and an arbitrary reverse cycle on the B,T diagram. The cycles being compared should be situated between the limiting temperatures T_1 and T_2 in order to eliminate any ambiguity.

The other efficiency characteristic of reverse cycles is the cooling coefficient

$$\varepsilon = Q_2/W_c = Q_2/(Q_1 - Q_2).$$

Having inscribed an arbitrary cycle into the Carnot cycle (Fig. 1), we can see that $Q_{2,c}$ is equal to the area of the rectangle 1456 whereas the same heat of the arbitrary cycle Q_2 is the sum of the areas $1456 + c + a$, so that $Q_{2,c} < Q_2$. The work of the Carnot cycle $W_{c,C}$, which is equal to the area of the rectangle 1234, exceeds the work of the arbitrary cycle W_c by the amount $a + b + c + d$, whence $W_{c,C} > W_c$. Consequently we find $Q_{2,C}/W_{c,C} < Q_2/W_c$ or

$$\varepsilon_C < \varepsilon,$$

which is consistent with the conclusion (2). This means that the heating and cooling coefficients of an arbitrary reverse cycle are greater than the corresponding characteristics of the reverse Carnot cycle.

The characteristics ε and ψ are convenient for estimating the useful properties of the reverse cycle but do not reflect the appearance of the second law, as may be achieved using the thermal efficiency of the forward cycle. The absence of such a characteristic, i.e., the thermal efficiency of the reverse cycle, makes it difficult to obtain unambiguous results in thermodynamic analyses.

THERMAL EFFICIENCY OF REVERSE CYCLES

By definition, in all cases the efficiency should be a proper fraction[12] given by the ratio of the useful effect to the energy consumption.

Heat pumps operating on reverse cycles can transfer heat from cold to hot media. The efficiency for heat engines should be defined as the ratio of the difference between the energy Q_1 supplied to the working medium and the compensation energy E to the energy Q_1, i.e.,

$$\eta = (Q_1 - E)/Q_1. \tag{4}$$

For forward cycles E is the heat Q_2 removed from the working medium and then

$$\hat{\eta} = (Q_1 - Q_2)/Q_1,$$

whereas for reverse cycles it is the work (mechanical energy) W_c needed to transfer heat from the cold to the hot medium so that

$$\hat{\eta} = (Q_1 - W_c)/Q_1 = Q_2/Q_1.$$

The second Carnot theorem gives $1 - T_2/T_1 = \eta_c > \hat{\eta} = 1 - Q_2/Q_1$, and $T_2/T_1 < Q_2/Q_1$, and we obtain

$$\hat{\eta}_C < \hat{\eta}. \tag{5}$$

Inequality (5) is consistent with inequalities (2) and (3), i.e., the reverse Carnot cycle is most inefficient.

MULTIPLICITY OF EFFICIENCY CHARACTERISTICS OF FORWARD AND REVERSE CYCLES AND THEIR INTERRELATIONSHIP

It can be shown that the introduction of a third efficiency characteristic for the reverse cycle is excessive. However, the thermal coefficient ε_t is known for the forward cycle which indicates the possibility of utilizing the heat dumped in the forward cycle

$$\varepsilon_t = \hat{\eta} + Q_2/Q_1.$$

For the forward cycle we can also suggest the ecological characteristic

$$\varepsilon_e = W_c/Q_2,$$

which at the same time characterizes the working efficiency of the engine (if $\varepsilon_e > 1$, the engine operates in the most perfect cycle).

As a result, the forward and reverse cycles can be characterized by a system of interrelated characteristics. We know[13] that

$$\hat{\eta}(\varepsilon + 1) = 1. \tag{6}$$

However, this relation is not unique since

$$\hat{\eta} \cdot \psi = \varepsilon \cdot \varepsilon_e = \varepsilon_t = 1, \tag{7}$$

and also

$$\hat{\eta}(\varepsilon_e + 1) = 1, \tag{8}$$

which is equivalent to Eq. (6). The efficiency characteristics of the reverse cycles are related analytically

$$\hat{\eta} = \varepsilon/\psi. \tag{9}$$

It follows from Eqs. (6)–(8) that for the Carnot cycle, ε_e and $\hat{\eta}$ have the highest values while ψ, ε, and $\hat{\eta}$ have the lowest values compared with other cycles. Hence, the forward Carnot cycle is most efficient in terms of all characteristics while the reverse cycle is least efficient.

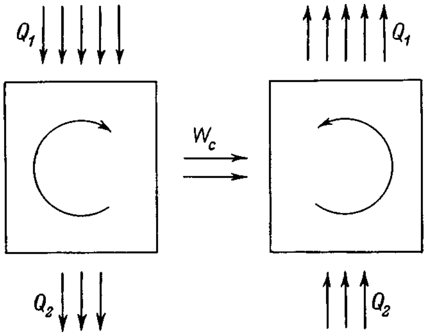


FIG. 2.

THEOREM OF ADDITIVITY OF THE EFFICIENCY OF FORWARD AND REVERSE CYCLES

The property of additivity of the efficiencies is the most general from the point of view of the first law of thermodynamics. It is easily shown by introducing $\widehat{\eta}$ that in any cyclic process the sum of the thermal efficiencies is one

$$\widehat{\eta}+\widehat{\eta}=1. \tag{10}$$

This statement can be taken as the theorem of additivity of the efficiencies of the forward and reverse cycles whose analytic expression together with Eqs. (6)–(9) combines the characteristics of both groups of cyclic processes. As the difference between the temperatures of the hot and cold sources in any cycle increases, $\widehat{\eta}$ increases and $\widehat{\eta}$ decreases. Formula (10) is a rational expression for the condition of reversibility of the heat and work conversion in a cyclic process conducted in the forward and then in the reverse direction or in a system of two identical coupled cycles, one being forward and the other reverse (Fig. 2), and is equivalent to writing this condition in the form (6)–(8).

The theorem (10) shows the erroneous nature of the conclusion[14] that the efficiencies of coupled cycles are equal.

The second law of thermodynamics can be formulated from the point of view of the efficiency of the reverse cycle: it is impossible to completely transfer energy from a cold to a hot medium without an additional expenditure of energy (compensation energy) in the implementation of this process, i.e., it is impossible to achieve a value of unity for $\widehat{\eta}$.

Equilibrium conditions give the highest efficiency for the forward and reverse cycles.

We obtain a certain ''inverse'' Carnot principle: the greater the temperature difference between the two sources, the lower the efficiency of engines and devices operating in the reverse cycle.

For nonequilibrium processes we clearly have

$$\widehat{\eta}_i+\widehat{\eta}<1$$

or

$$\widehat{\eta}+\widehat{\eta}_i<1,$$

where the subscript *i* refers to irreversible processes.

CONCLUSIONS

In reverse cycles the ratio Q_2/Q_1 indicates the efficiency of transferring heat from a cold source to a hot one by analogy with the expression $1-Q_2/Q_1$, which describes the efficiency of transferring heat into work in accordance with the second law of thermodynamics. From this similarity we suggest that the ratio Q_2/Q_1 should be called the thermal efficiency of the reverse cycle.

The introduction of the thermal efficiency $\widehat{\eta}$ as a universal characteristic means that the reverse cycle can be assessed from common viewpoints regardless of the conditions of application of the heat pump (cooling, heating, and combined heating-cooling devices) and the condition for reversibility of the energy conversions in a cyclic process can be expressed rationally. The conclusion is confirmed that among the reverse cycles the Carnot cycle ceases to be a standard cycle. The proposed efficiency $\widehat{\eta}$ means that this can be made equal to $\widehat{\eta}$ by formulating the second law of thermodynamics for equilibrium and nonequilibrium processes.

[1] I. P. Bazarov, *Delusions and Errors in Thermodynamics* [in Russian], Moscow State University Press, Moscow (1993), 56 pp.
[2] G. N. Alekseev, *General Heat Engineering* [in Russian], Vysshaya Shkola, Moscow (1980), 552 pp.
[3] M. P. Vukalovich and I. P. Novikov, *Technical Thermodynamics* [in Russian], Gosénergoizdat, Moscow (1955), 336 pp.
[4] V. S. Zhukovskiĭ, *Thermodynamics* [in Russian], Énergoatomizdat, Moscow (1983), 304 pp.
[5] V. V. Nashchokin, *Technical Thermodynamics and Heat Transfer*, 3rd ed. [in Russian], Vysshaya Shkola, Moscow (1980), 469 pp.
[6] E. Schmidt, *Thermodynamics. Principles and Applications to Engineering*, transl. from the 3rd German ed. (Clarendon Press, Oxford, 1949) [Russ. transl., later ed., Énergiya, Moscow, 1965, 392 pp.].
[7] V. A. Kirillin, V. V. Sychev, and A. E. Sheĭndlin, *Technical Thermodynamics*, 4th ed. [in Russian], Énergoatomizdat, Moscow (1983), 416 pp.
[8] *Technical Thermodynamics: Textbook for Universities*, edited by V. I. Krutov, 2nd ed. [in Russian], Vysshaya Shkola, Moscow (1981), 493 pp.
[9] J. B. Fenn, *Engines, Energy, and Entropy: a Thermodynamics Prmer* [Freeman, San Francisco (1982); Mir, Moscow (1986), 336 pp.].
[10] I. A. Kvasnikov, *Thermodynamics and Statistical Physics. Theory of Equilibrium Systems* [in Russian], Moscow State University Press, Moscow (1991), 800 pp.
[11] L. Artsimovich, P. Kapitsa, and I. Tamm, Pravda No. 326, (1959).
[12] *Physics Encyclopedic Dictionary* [in Russian], Sov. Éntsiklopediya, Moscow (1983).
[13] N. I. Belokon', *Basic Principles of Thermodynamics* [in Russian], Nedra, Moscow (1968), 110 pp.
[14] J. Orear, *Physics* [Macmillan, New York (1979); Vol. 1, Mir, Moscow (1981), 336 pp.].

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Characteristic features of the formation of the vibrational distribution function of H_2 molecules in a hydrogen stream

F. G. Baksht and V. G. Ivanov

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

(Submitted April 15, 1998)
Zh. Tekh. Fiz. **69**, 15–21 (June 1999)

A theoretical analysis is made of the flow of vibrationally excited hydrogen in a channel. It is shown that coverage of the channel walls with adsorbed hydrogen atoms can substantially increase the concentration of vibrationally excited molecules in the stream. The possibility of applying these results to bulk sources of negative H^- hydrogen ions is discussed. It is shown that the rate of H^- ion generation in the source may be enhanced appreciably under conditions where this generation is achieved by dissociative attachment of thermal electrons to H_2 molecules injected into the discharge chamber, whose vibrational distribution function has been pre-enriched in excited molecules by suitably organizing the hydrogen flow in the channel.
[S1063-7842(99)00306-2]

1. The present paper is concerned with the generation of vibrationally excited H_2 molecules in a stream of molecular hydrogen. Interest in the generation of these molecules has arisen as a result of various plasma-chemical applications,[1,2] the use of vibrationally excited H_2 molecules in negative hydrogen ion sources,[3] and other factors. Various types of gas discharge are being actively used to generate vibrationally excited molecules (see, for instance, Refs. 1–4). In these discharges the formation of vibrationally excited molecules is generally determined by the vibrational pumping of lower vibrational levels by $e-v$ exchange and subsequent diffusion of vibrational quanta into regions of high vibrational numbers v as a result of $v-v$ exchange. An important factor is that in many cases, fairly efficient pumping of the lower vibrational levels of molecules is accompanied by appreciable heating of the gas. This reduces the role of $v-v$ exchange in filling the upper levels and increases their $v-t$ relaxation which ultimates reduces the populations of the higher vibrational states.

In Refs. 5 and 6 the present authors suggested using a flow of vibrationally excited hydrogen in a channel to enhance the populations of highly excited vibrational states. In this case, the vibrational distribution function f_v of the H_2 molecules forms in two stages. First, molecular hydrogen flows across the discharge where the hydrogen molecules acquire a fairly high average vibrational energy $\langle E_v \rangle$ as a result of $e-v$ exchange. The vibrationally excited hydrogen then flows through a channel whose walls are kept at a low temperature T_s such as room temperature. An additional, and for some vibrational levels, very substantial increase in the populations N_v is achieved by $v-v$ exchange in the cold gas in the channel, i.e., by comparatively inefficient $v-t$ exchange. Note that similar effects caused by pumping high vibrational levels of H_2 molecules were observed experimentally and investigated theoretically in the afterglow of pure hydrogen[7] and cesium–hydrogen discharges.[8,9]

2. In order to reduce the losses of vibrationally excited molecules to the channel walls, it is natural to use wall materials having the highest possible potential barrier for adsorption of molecular hydrogen. Such materials specifically include transition metals such as copper, gold, and silver. As in Ref. 6, we shall subsequently analyze a flow of vibrationally pumped hydrogen in a copper channel since the probabilities of adsorption and desorption of molecular hydrogen, including vibrationally excited hydrogen, have been comparatively thoroughly studied for copper surfaces.[10–19] The most comprehensive theoretical analysis of the interaction between vibrationally excited H_2 molecules and a copper surface was made by Cacciatore and Billing[10] for the Cu(111) face. These authors[10] showed that at low kinetic energies ($E_{\mathrm{kin}} < 0.1$ eV) the surface deactivation of H_2 molecules is mainly caused by tunneling of molecules through the potential barrier in the surface layer followed by dissociation, i.e., attachment of H atoms to the surface. Cacciatore and Billing[10] also determined the probability of attachment $w_v(E_{\mathrm{kin}})$ of a molecule vibrationally excited to the v level, having the energy E_{kin} along the normal to the surface.

In the present study we consider a low-voltage cesium–hydrogen discharge as the discharge in which the initial vibrational pumping of H_2 takes place.[4,20] This type of discharge was selected because its parameters can be determined theoretically to quite acceptable accuracy (see Refs. 20–22, for instance). This method of enhancing the concentration N_v of vibrationally excited molecules in a stream is naturally also promising for pure hydrogen discharges.

In Ref. 6 we reported a theoretical analysis of the flow of vibrationally excited hydrogen in a channel using the mechanism for the surface deactivation of molecules indicated above. However, in Ref. 6 we neglected effects caused by the finite coverage Θ of the metal surface with adsorbed hydrogen atoms. The aim of the present study is to take these effects into account. We shall show that when these effects are taken into account, the populations N_v of a whole range of vibrationally excited states of H_2 molecules increase substantially in the upper part of the spectrum. This then leads to a considerable increase in the effective rate constant for dissociative attachment of thermal electrons to vibrationally excited H_2 molecules compared with the case $\Theta=0$ considered in Ref. 6.

3. The basic equations describing the flow of vibrationally excited hydrogen in a channel are written as follows (for further details see Ref. 6). The distribution of the molecular hydrogen pressure $p(x)$ over the length of a planar channel is given by

$$p(x)=[p_0^2-(p_0^2-p_s^2)x/h]^{1/2}, \tag{1}$$

where $p_0=p(0)$ is the hydrogen pressure at the entrance to the channel $(x=0)$, i.e., approximately the pressure in the discharge, and p_s is the pressure at the channel exit $(x=h)$.

Then p_0 and p_s are related by

$$h/L=(R_s/24)(c_p/c_v)^{-1}[(p_0/p_s)^2-1], \tag{2}$$

where $R_s=\rho_s L V_s/\eta$ is the Reynolds number calculated using the channel cross section L, the molecular hydrogen viscosity η, the density ρ_s, and the velocity of sound V_s in the exit cross section of the channel.

The vibrational level populations $N_v(x)$ are determined from the system of equations

$$\frac{d}{dx}(N_v V)=I_v^{(vv)}\{N_v\}+I_{vM}^{(vt)}\{N_v\}+I_{vA}^{(vt)}\{N_v\}+I_v^{(w)}\{N_v\}$$

$$(v=0,1,2,\ldots,14), \tag{3}$$

where $V(x)$ is the gasdynamic velocity averaged over the channel cross section.

The terms on the right-hand side of Eq. (3) systematically allow for $v-v$ exchange, $v-t$ exchange with hydrogen molecules and atoms,[23] and vibrational relaxation of molecules at the channel walls

$$I_v^{(w)}\{N_v\}=-N_v/\tau_v+\Psi_v/L. \tag{4}$$

The first term on the right-hand side of Eq. (4) describes the losses of vibrationally excited molecules caused by their drift from the stream toward the channel walls and τ_v is the corresponding effective lifetime of a vibrationally excited molecule in the channel. The second term on the right-hand side describes the increase in the concentration of vibrationally excited H_2 molecules in the stream as a result of desorption of hydrogen from the channel walls, where Ψ_v is twice the density of the molecular desorption flux from the walls.

The concentration $N_H(x)$ of hydrogen atoms in the channel was determined from

$$\frac{d}{dx}(N_H V)=-N_H/\tau_H, \tag{5}$$

where τ_H is the lifetime of hydrogen atoms in the channel caused by their losses to the walls.

In Eqs. (3)–(5) $N_v(x)$ and $N_H(x)$ are the concentrations of vibrationally excited H_2 molecules and H atoms averaged over the channel cross section: the true distributions of N_v and N_H over the channel cross section are generally inhomogeneous because of the removal of particles to the channel walls and the resulting depleted concentration in the wall zones. Effects arising from the depleted concentrations are taken into account by introducing the effective lifetimes of vibrationally excited molecules τ_v and atoms τ_H, which are given by[6,24]

$$\tau_v=\frac{L^2}{\pi^2 D_{sd}}+\frac{L}{v_{H_2}}\frac{2-\gamma_v}{\gamma_v},$$

$$\tau_H=\frac{L^2}{\pi^2 D_{12}}+\frac{L}{v_H}\frac{2-\gamma_H}{\gamma_H}, \tag{6), (7}$$

where D_{sd} and D_{12} are the coefficient of self-diffusion of H_2 molecules and the coefficient of diffusion of H atoms in H_2 molecules,[25,26] and $v=\sqrt{8kT/\pi\times M}$.

The coefficients γ_v and γ_H determine the fractions of excited H_2 molecules lost to the surface of the channel walls from the total number of molecules or atoms incident on the surface, respectively; γ_v and γ_H depend on the surface coverage Θ with adsorbed hydrogen atoms. If $\gamma\sim 1$ in Eqs. (6) or (7), the first term on the right-hand side is substantially greater than the second. In this case, the lifetime of an excited molecule or atom is limited by particle diffusion from the channel to the walls and a characteristic diffusion concentration profile is established between the walls. In the opposite case, when $\gamma\ll 1$ (for further details see Ref. 6), only the last terms are important in Eqs. (6) or (7). This corresponds to comparatively fast diffusion which equalizes the particle concentration over the channel cross section. In this last case, the particle flux from the gas to the surface of each channel wall is $\frac{1}{4}N\overline{v}\gamma$ and is small compared with the random current $\frac{1}{4}N\overline{v}$.

4. We shall now determine the values of γ_v, γ_H, Ψ_v, and the coverage Θ. The value of γ_v, which is the probability of a vibrationally excited molecule being adsorbed at the surface, depends on the attachment probability $w_v(E_{\text{kin}})$ and the surface coverage Θ with adsorbed hydrogen atoms[27–29]

$$\gamma_v(\Theta)=\langle w_v(E_{\text{kin}})\rangle_T(1-\Theta)^2. \tag{8}$$

Here $\langle w_v(E_{\text{kin}})\rangle_T$ denotes the attachment probability averaged over a semi-Maxwellian molecular distribution function in a random stream at gas temperature T. Since the $w_v(E_{\text{kin}})$ values given in Ref. 10 is insufficient to calculate $\langle w_v\rangle$, we made the approximation that everywhere $\langle w_v(E_{\text{kin}})\rangle\cong w_v(kT)$. The values of $w_v(kT)$ in the required range of temperatures T were determined by extrapolating the data from Ref. 10.

Under the conditions being considered here, the probability γ_H, which determines the fraction of hydrogen atoms lost to the surface is written as

$$\gamma_H(\Theta)=w_H(1-\Theta)+\sigma_{ER}^{(\text{eff})}\sigma_H\Theta. \tag{9}$$

The first term on the right-hand side of Eq. (9) describes the attachment of atomic hydrogen to the surface and is similar to expression (8); w_H is the probability of attachment of atomic hydrogen to the surface. The second term on the right-hand side describes the probability of an H_2 molecule being formed at the surface as a result of surface recombination of an incident H atom with an adsorbed hydrogen atom by the Eley–Rideal mechanism.[18] Here $\sigma_{\rm H}\cong 1.5\times 10^{15}\,{\rm cm}^{-2}$ is the density of sorption centers at the surface of the Cu (111) face[14] and $\sigma_{\rm ER}^{(\rm eff)}$ is the effective cross section of Eley–Rideal surface recombination.

The probability $w_{\rm H}$ of attachment of hydrogen atoms to the copper surface was assumed to be one because of the low gas temperature and the very small thermal energy spread in the atomic hydrogen stream penetrating into the metal. Consequently, the probability of atoms returning from the metal to the gas is low. Note that similar values of $w_{\rm H}$ were calculated by Bischler *et al.*[16] for room-temperature gas and the Cu(110) surface allowing only for a phonon mechanism of energy loss for a hydrogen atom in the metal.

The effective cross section for surface recombination of a hydrogen atom by the Eley–Rideal mechanism followed by desorption of a vibrationally excited H_2 molecule was assumed to be $\sigma_{\rm ER}^{(\rm eff)}=5\,{\rm \AA}^2$. The value used for the cross section is the result of an analysis of the experimental data[14] made by Persson and Jackson.[18] We note that the cross section $\sigma_{\rm ER}^{(\rm eff)}$ thus determined from the experimental data is substantially higher than the calculated cross section for Eley–Rideal recombination obtained in Ref. 18. A quite obvious reason for this (as noted by Persson and Jackson in Ref. 19) is the extremely large cross section for capture of an H atom in the surface layer of the metal. Consequently, Eley–Rideal recombination involving trapped but not yet thermalized hydrogen atoms at the surface is much more efficient than direct Eley–Rideal recombination of an atom incident on the surface for which the cross section was calculated in Ref. 18.

For these values of $w_{\rm H}$ and $\sigma_{\rm ER}^{(\rm eff)}$, we find $\gamma_{\rm H}\sim 1$, so that the first term on the right-hand side of Eq. (7) for any coverage Θ is substantially greater than the second. Thus, for the calculations we assumed $\tau_{\rm H}=\tau_d^{(\rm H)}$, where $\tau_d^{(\rm H)}=L^2\pi^2 D_{12}$ is the time for diffusion of atomic hydrogen from the stream to the channel walls. As a result, $\tau_{\rm H}$ is almost independent of Θ and the accurate values of the constants on the right-hand side of Eq. (9). On the contrary, the lifetime τ_v of an excited H_2 molecule in the channel depends strongly on Θ.

The coverage Θ of the surface of the channel walls with adsorbed hydrogen atoms is determined from

$$\left[2\sum_v N_v/\tau_v(\Theta)+N_{\rm H}/\tau_d^{(H)}\right]L=2\sum_v \Psi_v(\Theta). \tag{10}$$

The left-hand side of Eq. (10) is the number of hydrogen atoms adsorbed per unit time on both walls of the channel per unit surface area. The right-hand side of Eq. (10) gives the number of hydrogen atoms entering the stream as a result of desorption of molecules from the channel walls; Ψ_v is twice the flux density of molecules excited to level v desorbed from each channel wall.

With reference to the notation used in expressions (4) and (10), we note that in the adopted formalism, the desorption of vibrationally excited molecules from the walls should be taken into account, not by including the flux density Ψ_v from the channel walls as a separate term on the right-hand side of expressions (4) and (10), but by suitably redefining γ_v and τ_v. We also note that the effective lifetime τ_v can generally only be introduced if, for given v, the emission of vibrationally excited molecules from the gas stream to the walls exceeds the reverse desorption flux. Allowance for all these factors should make the calculation scheme significantly more complex. In the situation being analyzed here, however, these complexities can be avoided because in the calculations presented below, Eley–Rideal desorption will play a dominant role.[18] The density of the desorption flux of vibrationally excited H_2 molecules from the walls is described by

$$\Psi_v^{(\rm ER)}=L\,\frac{N_{\rm H}}{\tau_d^{(H)}}\,\frac{\sigma_{\rm ER}^{(\rm eff)}\sigma_{\rm H}\Theta}{\gamma_{\rm H}(\Theta)}\,f_v^{(\rm ER)}. \tag{11}$$

Here the cofactor $\sigma_{\rm ER}^{(\rm eff)}\sigma_{\rm H}\Theta/\gamma_{\rm H}$ determines the fraction of hydrogen atoms incident on the surface of the channel walls and participating in surface recombination by the Eley–Rideal mechanism and $f_v^{(\rm ER)}$ is the vibrational distribution function of the molecules desorbed by the Eley–Rideal mechanism, normalized to unity $(\Sigma_v f_v^{(\rm ER)}=1)$. The distribution function $f_v^{(\rm ER)}$ used in the calculations (see Fig. 2b) is taken from Ref. 30: $f_v^{(\rm ER)}$ is also similar to one of the calculated variants[18] (see Fig. 10c in Ref. 18). An important factor is that the vibrational distribution function $f_v^{(\rm ER)}$ is almost only nonzero for $v=1$–3 whereas the attachment probabilities w_v and therefore the values of $1/\tau_v$ in expressions (4) and (10) are only nonzero for comparatively high vibrational levels $v\geqslant 5$ (Ref. 10). As a result, the effective lifetimes τ_v given by expressions (6) and (8) will only be used in the calculations when the corresponding desorption fluxes are $\Psi_v=0$. This means that expression (6) can be used for τ_v. Allowance for desorption by the Langmuir–Hinshelwood mechanism in cases where this mechanism is important requires a special analysis (see Sec. 6).

5. We shall now discuss the calculated distribution of the coverage $\Theta(x)$ over channel length. Figure 1 shows distributions of the molecular and atomic hydrogen concentrations $N_{\rm H_2}(x)$ and $N_{\rm H}(x)$ over the channel length determined from Eqs. (1) and (5) together with the distribution $\Theta(x)$, calculated by solving Eq. (10) for $\Psi_v=\Psi_v^{(\rm ER)}$. The initial values of the concentrations $N_{\rm H_2}(0)$ and $N_{\rm H}(0)$ at the channel entrance (allowing for a jump in temperature T at the channel entrance[6]) correspond to the conditions achieved in an initial low-voltage cesium–hydrogen discharge. The discharge parameters were calculated using a method described in Ref. 23. Figures 2a and 2b show calculated distributions of the vibrational distribution function of hydrogen molecules in the stream $f_v(x)=N_v(x)/\Sigma_v N_v(x)$ over the channel length, calculated by solving the system of equations (3), and the vibrational distribution function $f_v^{(\rm ER)}$ used in the calculations where desorption is described by the Eley–Rideal mechanism ($f_v(x)$ is given in relative units $f_v(x)/f_v(0)$).

We shall analyze the $\Theta(x)$ curve plotted in Fig. 1 for which we shall first assess the role of the various terms on the left-hand side of Eq. (10). Simple estimates show that in the initial section of the channel, where the vibrational distribution function $f_v(x)$ of the molecules in the stream, under those conditions making the major contribution to adsorp-

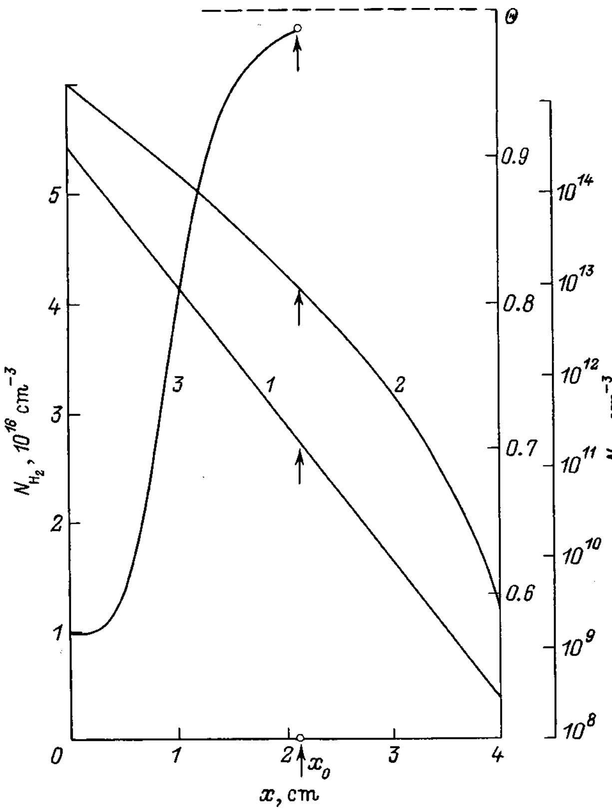


FIG. 1. Distributions of atomic and molecular hydrogen concentrations and coverage over channel length: *1* — N_H, *2* — N_{H_2}, *3* — Θ. Initial discharge parameters: $L = 0.3$ cm, $N_{H_2}^{(0)} = 3\times10^{16}$ cm^{-3}, $N_H^{(0)} = 1.8\times10^{14}$ cm^{-3}, $N_{Cs}^{(0)} = 10^{14}$ cm^{-3}, $n_e = 3.26\times10^{13}$ cm^{-3}, $N_H^{(0)} = 3.8\times10^{12}$ cm^{-3}, $T_e = 0.65$ eV, $\langle E_v\rangle = 0.335$ eV, $T_s = 300$ K. $j_s = 4.5$ A/cm^2, $j \cong 5$ A/cm^2, $U \cong 4.9$ V, $\varphi_1 = 8.65$ V, and $\varphi_2 = 0.75$ V. The gas temperature in the discharge was $T_0 = 0.06$ eV and in the channel $T = 0.03$ eV.

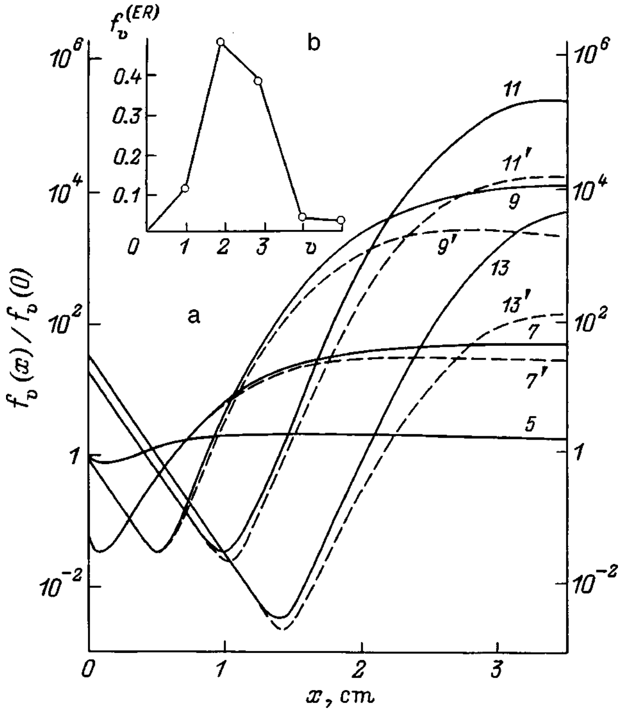


FIG. 2. a — Variation of the vibrational distribution function of H_2 molecules over the channel length: the number of the vibrational levels is indicated on the curves. The solid curves gives the results of the present study and the dashed curves give those of Ref. 6 obtained for $\Theta = 0$; the discharge and flow parameters in the channel are the same as in Fig. 1; b — vibrational distribution function of H_2 molecules desorbed from the walls by the Eley–Rideal mechanism.

tion, is still close to the initial vibrational distribution function $f_v(0)$ in the discharge, only the second term on the left-hand side of Eq. (10) describing the adsorption of atomic hydrogen is significant. This yields the coverage Θ, which does not depend on N_H and is given by $\Theta = 1/(1 + \sigma_H \sigma_{ER}^{(eff)})$. Then, as the populations N_v of the levels contributing to the adsorption of molecular hydrogen increase and the atomic hydrogen concentration N_H decreases, the first term on the left-hand side of Eq. (10) describing the adsorption of vibrationally excited H_2 molecules also becomes important. This has the result that the coverage Θ increases, approaching unity. A formal solution of Eq. (10) for $x \to \infty$ and $N_H \to 0$ gives $\Theta = 1$. The curve $\Theta(x)$ in Fig. 1 has a cutoff at the point x_0 where desorption by the Eley–Rideal mechanism is comparable with that by the Langmuir–Hinshelwood mechanism.

6. Desorption by the Langmuir–Hinshelwood mechanism only becomes appreciable at fairly low concentrations N_H and $\Theta \cong 1$. In this case, it follows from Eq. (8) that $\gamma_v \ll 1$, so that only the last term in Eq. (6) is important. As has been noted, this is consistent with the fact that the true concentration of vibrationally excited molecules becomes equalized over the channel cross section as a result of rapid diffusion and the flux of excited molecules from the gas to the channel walls becomes $(1/4) N_v \overline{v}_{H_2} \cdot \gamma_v$. In this approximation there is no need to use the effective lifetimes τ_v and the resultant molecular flux J_v at the gas–wall interface is calculated as

$$J_v = \frac{1}{4} N_v \overline{v}_{H_2}\, \gamma_v(\Theta) - \frac{1}{2} \Psi_v^{(LH)}, \tag{12}$$

$$\frac{1}{2} \Psi_v^{(LH)} = \frac{1}{2}\, \delta_{LH} \Theta^2 f_v^{(LH)} \tag{13}$$

i.e., the density of the Langmuir–Hinshelwood desorption flux of molecules excited to level v from one of the channel walls.

The value of $(1/2)\delta_{LH}\Theta^2$ is determined as in Ref. 31 and is the density of the total flux of H_2 molecules desorbed from the walls. Here $\delta_{LH}(T_s) = \nu \sigma_H^2 \exp(-E/kT_s)$ is the desorption rate constant, ν and E are the effective vibration frequency of an adsorbed atom and the desorption energy, respectively, and T_s is the wall temperature. Moreover, $f_v^{(LH)}$ is the vibrational distribution function of molecules desorbed by the Langmuir–Hinshelwood mechanism normalized to unity. The criterion for the validity of expression (12) is the inequality $\gamma_v \ll 2\pi^2 D_{sd}/L\overline{v}_{H_2}$ which for $x \geq x_0$ is satisfied by a large margin. The vibrational distribution function $f_v^{(LH)}$ is determined from the principle of detailed equilibrium

$$f_v^{(\mathrm{LH})}=[\langle w_v(E_{\mathrm{kin}})\rangle_{T_s}\exp(-E_v/kT_s)]/\left[\sum_v \langle w_v(E_{\mathrm{kin}})\rangle_{T_s}\exp(-E_v/kT_s)\right]. \qquad (14)$$

The surface temperature was assumed to be $T_s=300\,\mathrm{K}$. The rate constant $\delta_{\mathrm{LH}}(T_s)$ was calculated using the measured desorption spectrum for H_2/Cu (111) (see Fig. 7 in Ref. 31). Finally, instead of expression (10) for $x\geqslant x_0$ we obtain the following equation to determine the coverage Θ, which expresses the atom balance at each channel wall:

$$2\sum_v \frac{1}{4}N_v\bar{v}_{\mathrm{H}_2}\gamma_v(\Theta)+\frac{1}{2}\frac{N_{\mathrm{H}}L}{\tau_d^{(\mathrm{H})}}=\sum_v [\Psi_v^{(\mathrm{ER})}(\Theta)+\Psi_v^{(\mathrm{LH})}(\Theta)], \qquad (15)$$

where $\Psi_v^{(\mathrm{ER})}$ is described by expression (11) as before and $\Psi_v^{(\mathrm{LH})}$ was defined in expression (13).

The value of Θ obtained by solving Eq. (15) for $x\geqslant x_0$ is very close to unity. As a result for $x\geqslant x_0$ the probability of vibrational deactivation of H_2 molecules is very low so that the term $I_v^{(w)}\{N_v\}$ in Eq. (3) is insignificant here. Vibrational deactivation of molecules in this region is caused by $v-t$ exchange. Now, unlike Eq. (6) $v-t$ exchange limits the optimum channel length h for vibrational pumping of high levels. This means that a substantially higher level of vibrational excitation of the molecules can be achieved at higher v than in Ref. 6 both as a result of the lower probabilities $\gamma_v(\Theta)$ of vibrational deactivation of the molecules at the walls and as a result of the increased channel length h.

Figure 2a shows how the results of the calculations are influenced by the reduced probability $\gamma_v(\Theta)$, with two series of curves being plotted: the dashed curves give the results of Ref. 6 obtained for $\Theta=0$ and the solid curves give the results of the present study. It can be seen that for the same channel length $h=3.5\,\mathrm{cm}$, the vibrational distribution function $f_v(h)$ at the channel exit is significantly higher at high vibrational numbers v compared with $f_v(h)$ calculated in Ref. 6. This is also illustrated in Fig. 3, which gives the vibrational distribution functions $f_v(h)$ (curves *2, 2′* and *3*) at the channel exit for various calculations. In Fig. 3, curve *2* calculated in Ref. 6 corresponds to approximately the optimum channel length for the selected discharge and flow parameters in the channel. The distribution function $f_v(h)$ shows the highest level of vibrationally highly excited molecules. A comparison of curve *2′* with curves *2* and *3* shows that allowance for a finite coverage Θ leads to an increase in the vibrational population distribution at high v and also increases the optimum channel length.

7. To conclude, we shall discuss how modification of the vibrational distribution function $f_v(x)$ in the channel influences the rate of H^- generation as a result of dissociative attachment of plasma electrons to vibrationally excited H_2 molecules.[32] As in Ref. 6, we shall assume that a hydrogen flux pumped vibrationally in a channel is injected into a discharge chamber which contains thermal electrons having a Maxwellian distribution with the temperature T_e'. We shall determine the effective rate constant for dissociative attachment

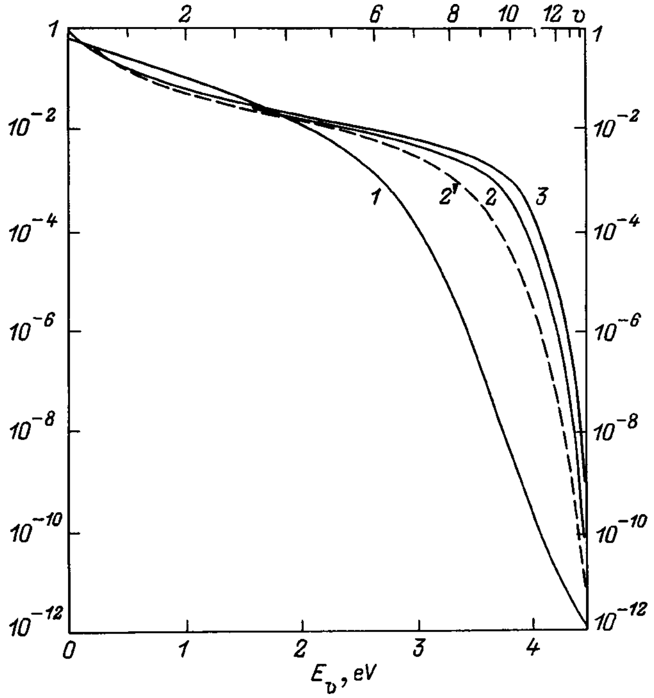


FIG. 3. Vibrational distribution function of H_2 molecules in the initial discharge (*1*) and at the channel exit (*2, 2′, 3*): $L=0.3\,\mathrm{cm}$; $h=3.5$ (*2, 2′*) and 4.5 cm (*3*); the solid curves give the results of the present calculations and the dashed curve gives the result[6] obtained for $\Theta=0$. The discharge parameters are the same as in Fig. 1.

$$\langle K_{DA}(h,T_e')\rangle=\sum_v f_v(h)K_v(T_e'), \qquad (16)$$

where $K_v(T_e')$ is the rate constant for generation of H^- ions by attachment of electrons to molecules vibrationally excited to the level v (Refs. 33 and 34).

Figure 4 gives values of $f_v(h)K_v(T_e')$ which indicate the partial contribution of the different vibrational levels to the total rate constant for dissociative attachment. Curve *1* was calculated for $h=0$ using the vibrational distribution function $f_v(0)$ in the initial discharge. Curve *2* gives the result of the present calculations where $f_v(h)$ is the vibrational distribution function at the channel exit, determined allowing for the finite coverage $\Theta(x)$ of the channel walls with adsorbed hydrogen atoms. It can be seen that the values of $f_v(h)K_v(T_e')$ are increased substantially as a result of the modification of the vibrational distribution in the channel.

Figure 5 gives $\langle K_{DA}\rangle$ as a function of T_e' for various calculation variants. Curve *1* was calculated using the initial vibrational distribution function $f_v(0)$ formed in the discharge. Curve *2′* was calculated in Ref. 6 and corresponds to conditions when the vibrational distribution function $f_v(h)$ was formed in a channel of optimum length ($h=3.5\,\mathrm{cm}$) for $\Theta(x)=0$. Curves *2* and *3* give the results of the present calculations made allowing for a finite coverage of the channel walls with adsorbed hydrogen atoms. Curve *2*, as in Ref. 6, was calculated for $h=3.5\,\mathrm{cm}$ and curve *3* corresponds to

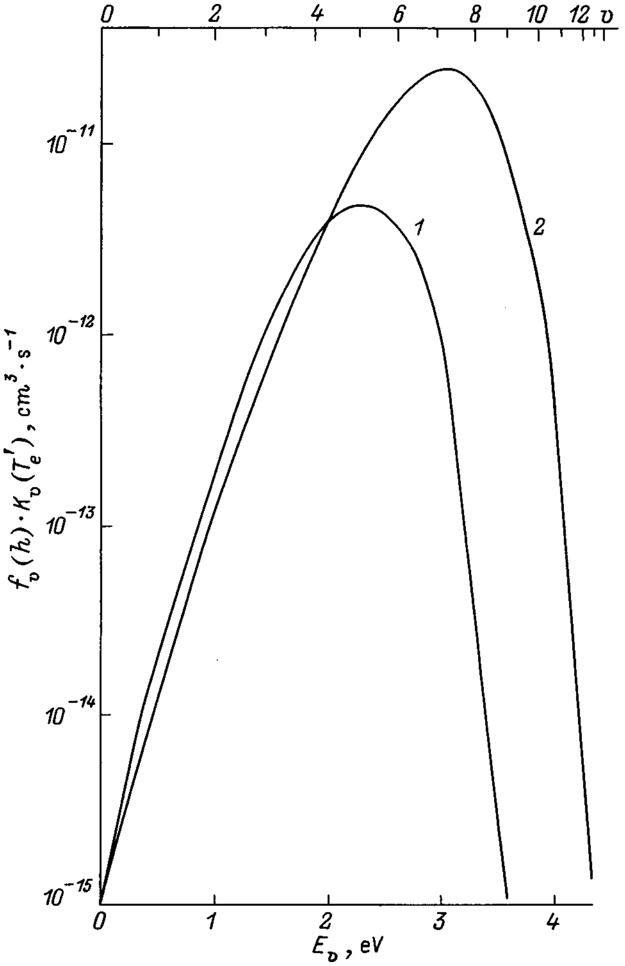


FIG. 4. Contributions of various vibrational levels to the effective rate constant for generation of H^- ions as a result of dissociative attachment. The discharge parameters are the same as in Fig. 1; $h=0$ (*1*), 3.5 cm (*2*); $L=0.3$ cm, and $T'_e=0.75$ eV.

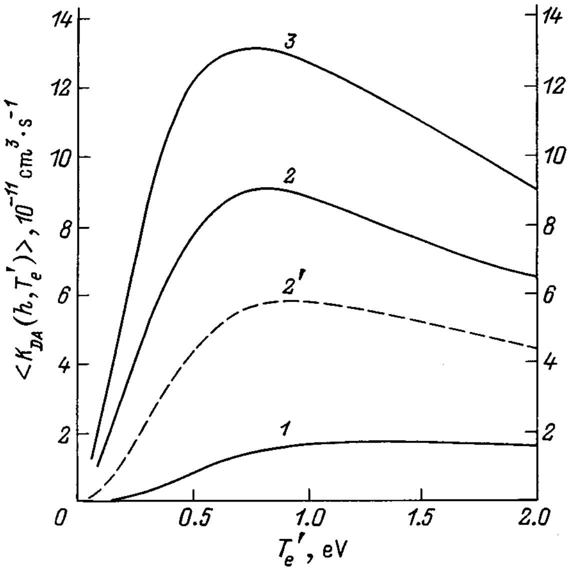


FIG. 5. Dependence of the effective constant $\langle K_{DA}(h,T'_e)\rangle$ for dissociative attachment on the electron temperature T'_e. The discharge parameters as the same as in Fig. 1; *1* — $\langle K_{DA}(0,T'_e)\rangle$; *2, 2'* and *3* — $\langle K_{DA}(h,T'_e)\rangle$, $L=0.3$ cm, $h=3.5$ (*2, 2'*), 4 and 4.5 cm (*3*); *2* and *3* — results of the present calculations, and *2'* — result[6] obtained for $\Theta=0$.

the optimum channel length $h=4$ or 4.5 cm (the results of the calculations for these two lengths are almost the same). It can be seen that allowance for a finite coverage $\Theta(x)$ leads to an increase in the effective rate constant for dissociative attachment $\langle K_{DA}\rangle$ (compare curves *2'* and *2*) and the optimum channel length corresponding to the maximum of $\langle K_{DA}\rangle$ (compare curves *2* and *3*).

8. To sum up, we have shown that when calculating the flow of vibrationally excited hydrogen in a channel and determining the vibrational distribution function of the molecules formed in the stream, it is important to allow for the coverage of the channel walls with adsorbed hydrogen atoms. Allowance for this coverage appreciably reduces the probabilities of vibrational deexcitation of H_2 molecules at the channel walls and increases the concentration of vibrationally excited molecules in the stream.

In particular, these effects may be of considerable interest for two-chamber sources of negative hydrogen ions in which the vibrational excitation of H_2 molecules and the generation of H^- ions are separated in space. In this case, a substantial increase in the rate of H^- generation in the discharge chamber can be achieved if these ions are generated by dissociative attachment of thermal electrons to H_2 molecules injected in the chamber, whose vibrational distribution function has been pre-enriched in excited molecules by suitably organizing the hydrogen flow in the channel.

The authors would like to thank M. Bakal and S. M. Shkol'nik for useful discussions.

This work was supported by INTAS Grant No. 94-316.

[1] A. A. Ivanov and T. K. Soboleva, *Nonequilibrium Plasma Chemistry* [in Russian], Atomizdat, Moscow (1978), 264 pp.

[2] V. D. Rusanov and A. A. Fridman, *Physics of a Chemically Active Plasma* [in Russian], Nauka, Moscow (1984), 415 pp.

[3] G. W. Hamilton and M. Bacal, IEEE Trans. Plasma Sci. **19**, 1143 (1991).

[4] F. G. Baksht and V. G. Ivanov, Zh. Tekh. Fiz **62**(2), 195 (1992) [Sov. Phys. Tech. Phys. **37**, 223 (1992)].

[5] F. G. Baksht and V. G. Ivanov, Pis'ma Zh. Tekh. Fiz **23**(1), 26 (1997) [Tech. Phys. Lett. **23**, 20 (1997)].

[6] F. G. Baksht and V. G. Ivanov, Zh. Tekh. Fiz **68**(10), 10 (1998) [Tech. Phys. **43**, 1145 (1998)].

[7] C. Gorse, M. Capitelli, M. Bacal *et al.*, Chem. Phys. **117**, 172 (1987).

[8] F. G. Baksht, G. A. Dyuzhev, L. I. Elizarov *et al.*, Pis'ma Zh. Tekh. Fiz **19**(22), 39 (1993) [Tech. Phys. Lett. **19**, 716 (1993)].

[9] F. G. Baksht, V. G. Ivanov, A. G. Nikitin, and S. M. Shkol'nik, Pis'ma Zh. Tekh. Fiz **20**(22), 83 (1994) [Tech. Phys. Lett. **20**, 927 (1994)].

[10] M. Cacciatore and G. D. Billing, Surf. Sci. **232**(1), 35 (1990).

[11] C. T. Rettner, D. J. Auerbach, and H. A. Michelsen, Phys. Rev. Lett. **68**, 1164 (1992).

[12] H. A. Michelsen and D. J. Auerbach, J. Chem. Phys. **94**, 7502 (1991).

[13] C. T. Rettner, H. A. Michelsen, and D. J. Auerbach, J. Chem. Phys. **102**, 4625 (1995).

[14] C. T. Rettner, Phys. Rev. Lett. **69**, 383 (1992).

[15] R. I. Hall, I. Cadez, M. Landau *et al.*, *Etude du Role des Surfaces dans la Population et la Relaxation de H_2 Vibrationnellement Excite*, Université Pierre et Marie Curie, Paris, 1995.

[16] U. Bischler, P. Sandl, and E. Bertel, Rev. Plasma Phys. **70**, 3603 (1993).

[17] B. Jackson and M. Persson, J. Chem. Phys. **96**, 2378 (1992).

[18] M. Persson and B. Jackson, J. Chem. Phys. **102**, 1078 (1995).

[19] B. Jackson and M. Persson, J. Chem. Phys. **103**, 6257 (1995).

[20] F. G. Baksht, G. A. Djuzhev, L. I. Elizatov *et al.*, Plasma Sources Sci. Technol. **3**(2), 88 (1994).

[21] F. G. Baksht, G. A. Dyuzhev, L. I. Elizarov *et al.*, Zh. Tekh. Fiz **62**(9), 148 (1992) [Sov. Phys. Tech. Phys. **37**, 959 (1992)].

[22] F. G. Baksht, V. G. Ivanov, A. A. Kostin *et al.*, Zh. Tekh. Fiz **65**(8), 186 (1995) [Tech. Phys. **40**, 851 (1995)].
[23] F. G. Baksht, L. I. Elizarov, V. G. Ivanov, and V. G. Yur'ev, Fiz. Plazmy **14**, 91 (1988) [Sov. J. Plasma Phys. **14**, 56 (1988)].
[24] Yu. Z. Ionikh, Opt. Spektrosk. **51**, 76 (1981) [Opt. Spectrosc. **51**, 39 (1981)].
[25] *Physical Quantities*, edited by I. S. Grigor'ev and E. Z. Meĭlikhov [in Russian], Atomizdat, Moscow (1991), 1230 pp.
[26] G. Blyth, J. Chem. Soc., Faraday Trans. 1 **83**, 2673 (1987).
[27] M. A. Pick and K. Sonnenberg, J. Nucl. Mater. **131**, 208 (1985).
[28] P. L. Andrew and A. A. Haasz, J. Appl. Phys. **72**, 2749 (1992).
[29] *Interaction between Hydrogen and Metals* edited by A. P. Zakharov [in Russian], Nauka, Moscow (1987), 295 pp.
[30] P. Kratzer and W. Brenig, Z. Phys. B **99**, 571 (1996).
[31] G. Anger, A. Winkler, and K. D. Rendulic, Surf. Sci. **220**, 1 (1989).
[32] M. Bacal and G. W. Hamilton, Phys. Rev. Lett. **42**, 1538 (1979).
[33] J. M. Wadehra, Phys. Rev. A **29**, 106 (1984).
[34] D. A. Skinner, A. M. Brunetau, P. Berlemont *et al.*, Phys. Rev. E **48**, 2122 (1993).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Nonlinear moment method for the isotropic Boltzmann equation and invariance of the collision integral

A. Ya. Énder

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

I. A. Énder

St. Petersburg State University, 199034 St. Petersburg, Russia

(Submitted February 20, 1998)
Zh. Tekh. Fiz. **69**, 22–29 (June 1999)

A new approach is proposed for the development of a nonlinear moment method of solving the Boltzmann equation. This approach is based on the principle of invariance of the collision integral with respect to the choice of basis functions. Sonine polynomials with a Maxwellian weighting function are taken as these basis functions for the velocity-isotropic Boltzmann equation. It is shown that for arbitrary interaction cross sections the matrix elements corresponding to the moments of the nonlinear collision integral are not independent but are coupled by simple recurrence formulas by means of which all the nonlinear matrix elements are expressed in terms of linear ones. As a result, a highly efficient numerical scheme is constructed for calculating the nonlinear matrix elements. The proposed approach opens up prospects for calculating relaxation processes at high velocities and also for solving more complex kinetic problems. © *1999 American Institute of Physics.* [S1063-7842(99)00406-7]

INTRODUCTION

The basic mathematical results of solving the Boltzmann equation include a linearized equation and slight deviations from equilibrium.[1–3] Analytic solutions of the nonlinear Boltzmann equation are known in very limited number of cases.[4] Existing numerical methods of solution (mainly various modifications of the Monte Carlo method) give only a rough idea of the behavior of the distribution function at high velocities.[5] At the same time, the behavior of the distribution function at high energies is a decisive factor in a whole range of physicochemical processes.

In 1982 Turchetti and Paolilli[6] published a study in which a nonlinear moment method was used for the first time to make systematic calculations of the distribution function at high velocities for the isotropic Boltzmann equation. The distribution function was represented as a truncated series of Sonine polynomials, and the tails of the distribution function were described using higher moments. The main difficulty in this method is calculating the interaction matrix corresponding to moments of the nonlinear collision integral. Even for the velocity-isotropic Boltzmann equation the problem is fairly complex. For instance, in this study where the power potentials are considered assuming that the scattering cross section is independent of the angles, the analytic formulas obtained for the matrix elements contain six nested sums. The calculation difficulties increase catastrophically as the number of moments taken into account increases so that it becomes impossible to go beyond the thirteenth moment.

The present authors previously developed a method of integral transformation of the nonlinear Boltzmann equation in which the distribution function and the collision integral are expressed as a superposition of Maxwellian distributions and an equation is constructed for the weighting function (α representation of the Boltzmann equation) equivalent to the Boltzmann equation in v-space.[7,8] In Ref. 9 we considered the same moment method as in Ref. 6 but by using the mathematical tools developed to construct the α-representation, we succeeded in obtaining formulas for the matrix elements for arbitrary power potentials, including those for Coulomb particle interaction. Moreover, the formulas obtained are considerably simpler (four nested sums) so that the calculations could be made as far as the thirtieth moment with the same accuracy.

In Ref. 10 we suggested that the invariance of the collision integral of the Maxwellian distribution function relative to the choice of basis functions could be used to analyze the matrix elements. These basis functions were taken to be Sonine polynomials with different temperatures of the Maxwellians characterizing the expansion. As a result, we obtained some relationships between the matrix elements which were used as criteria for the accuracy of the calculations. We particularly stressed that for arbitrary interaction cross sections, the nonlinear matrix elements are not independent but various relationships exist between them.

In the present paper the idea of invariance is generalized to the collision integral of an arbitrary distribution function. As a result, we obtained very detailed relations which could be used to express the nonlinear matrix elements in terms of linear ones. These relations can be applied to study how the basic properties of the linear elements influence the nonlinear ones and can also be used to check the accuracy of the calculations. Finally, these relations can be taken as recurrence formulas to calculate the nonlinear matrix elements in terms of linear ones.

INVARIANCE OF THE DESCRIPTION OF RELAXATION PROCESSES

When the moment method is applied to isotropic problems, the distribution function is expanded in terms of Sonine polynomials with a Maxwellian weighting

$$f(v,t)=n_0M(v,T)\sum_{r=0}^{\infty} C_r(T)S^r_{1/2}(mv^2/2kT). \tag{1}$$

Here $M(v,T)=(m/2kT\pi)^{3/2}\exp(-mv^2/2kT)$ is the Maxwellian and n_0 is the particle number density. We know that this series converges if the Grad criterion is satisfied (for further details see Ref. 8). Substituting Eq. (1) into the right-hand side of the Boltzmann equation, then multiplying both sides of the equation by $S^r_{1/2}(mv^2/2kT)$, and integrating with respect to v, we obtain

$$dC_r/dt'=\sum_{r_1,r_2} K'^{r}_{r_1,r_2}(T)C_{r_1}C_{r_2}, \quad t'=t/\tau. \tag{2}$$

The dimensionless matrix $K'^{r}_{r_1,r_2}$ is defined in terms of the collision integral $\hat{I}(f,f)$ as follows:

$$K'^{r}_{r_1,r_2}(T)=4\pi n_0\tau\left(\int_0^{\infty} S^r_{1/2}\hat{I}(MS^{r_1}_{1/2},MS^{r_2}_{1/2})v^2dv\right)\nu_r,$$

$$\nu_r=(2r+1)!!/(2r)!!, \tag{3}$$

where ν_r is the square of the norm of the Sonine polynomial. Subsequently, the primes will be omitted and the choice of τ will be specifically discussed.

In Ref. 10 we derived a formula to check the accuracy of the matrix element calculations. This formula was derived assuming that the collision integral of the Maxwellian distribution function is invariant relative to the choice of basis and comprised the following relation between the matrix elements:

$$\sum_{r_1=0}^{N} K^r_{r_1,N-r_1}=0, \quad r,N=0,\dots,\infty. \tag{4}$$

The equalities (4) are satisfied for arbitrary scattering cross sections. Quite clearly, not only the collision integral of a Maxwellian but also the collision integral of an arbitrary distribution function should possess properties of invariance.

When expanding in terms of Sonine polynomials, we must bear in mind that these are orthogonal, with a Maxwellian weighting characterized by a certain temperature T. The transition to a different temperature corresponds to a new unit of velocity measurement and essentially involves a transition to a new basis. In the kinetic theory of gases the temperature T is usually taken to be the equilibrium gas temperature.

For the transition from one basis to another it is convenient to use the α representation $(\alpha=m/(2kT))$ of the Boltzmann equation:[7,8]

$$n_0\frac{\partial\varphi}{\partial t}=n_0^2\int_0^{\infty} A(T,T_1,T_2)\varphi(T_1,t)\varphi(T_2,t)dT_1dT_2, \tag{5}$$

where the distribution function in v-space is related to $\varphi(T,t)$ by:

$$f(v,t)=\int_0^{\infty} M(v,T)\varphi(T,t)dT. \tag{6}$$

The kernel $A(T,T_1,T_2)$ is a mapping of the collision integral of the two Maxwellians $J^M(T_1,T_2,v)$ in α-space

$$J^M(T_1,T_2,v)=n_0^2\int_0^{\infty} M(v,T)A(T,T_1,T_2)dT. \tag{7}$$

The Maxwellian-weighted orthogonal system of Sonine polynomials $S^r_{1/2}(mv^2/T_*)$ corresponds to the biorthogonal system of functions s^r_L and s^r_R in α-space[8]

$$s^r_R(T,T_*)=(T_*)^r\delta^{(r)}(T-T_*)/r!, \tag{8}$$

$$s^r_L(T,T_*)=(1-T/T_*)^r. \tag{9}$$

Here $\delta^{(r)}(T-T_*)$ is an rth-order derivative of the δ-function. For conciseness we use the notation $M(v,T)S^r_{1/2}(mv^2/2kT)=S_r(v,T)$. Then we have

$$S_r(v,T_*)=\int_0^{\infty} M(v,T)s^r_R(T,T_*)dT, \tag{10}$$

and it follows from Eqs. (8) and (9) that

$$\int_0^{\infty} s^i_L(T,T_*)s^j_R(T,T_*)dT=\delta_{ij}. \tag{11}$$

We express the distribution function in v-space in two bases with the temperatures T_0 and T_1

$$f(v,t)=\sum_{k=0}^{\infty} C^0_k(t)S_k(v,T_0)=\sum_{r=0}^{\infty} C^1_r(t)S_r(v,T_1). \tag{12}$$

Quite clearly the vectors C^0 and C^1 do not change on transition from v- to α-space and the equalities (12) in the α-representation have the form

$$\varphi(T,t)=\sum_{k=0}^{\infty} C^0_k(t)s^k_R(T,T_0)=\sum_{r=0}^{\infty} C^1_r(t)s^r_R(T,T_1). \tag{13}$$

In order to find the relation between the vectors C^0 and C^1, we perform scalar multiplication of both sides of equation (13) by $s^r_L(T,T_1)$. Then, using Eqs. (8), (9), and (11), we obtain

$$C^1_r=\sum_{k=0}^{\infty} d_{r,k}(T_1,T_0)C^0_k, \tag{14}$$

where the matrix elements of the transition from one basis to the other $D(T_1,T_0)$ are expressed in terms of the scalar product

$$d_{r,k}(T_1,T_0)=\int_0^{\infty} s^r_L(T,T_1)s^k_R(T,T_0)dT$$

$$=(s^r_L(T,T_1),s^k_R(T,T_0)). \tag{15}$$

The Sonine polynomials are then transformed as follows:

$$s_R^r(T,T_1)=\sum_{k=0}^{\infty} d_{k,r}(T_0,T_1)s_R^k(T,T_0). \tag{16}$$

Using the simple expressions for s_L^r and s_R^k in the α-representation, Eqs. (8) and (9), we obtain

$$d_{r,k}(T_1,T_0)=\begin{cases}\begin{pmatrix} r\\ k\end{pmatrix}(T_1-T_0)^{r-k}T_0^k/T_1^r, & r\geqslant k,\\ 0, & r<k.\end{cases} \tag{17}$$

The matrix D is triangular. We also note that it is not unitary. For any linear transformation we have

$$d_{r,k}(T_1,T_0)=\sum_{p=0}^{\infty} d_{r,p}(T_1,T_*)d_{p,k}(T_*,T_0). \tag{18}$$

In our case, because the matrix is triangular, summation is performed between r and k. By directly substituting Eq. (17) into Eq. (18), we can confirm that this property is indeed found. In operator form Eq. (18) is given by

$$\hat{D}(T_1,T_0)=\hat{D}(T_1,T_*)\hat{D}(T_*,T_0). \tag{19}$$

Since $\hat{D}(T_0,T_0)$ is the unit operator $\hat{E}$, assuming $T_1=T_0$, we obtain

$$\hat{D}(T_*,T_0)=\hat{D}^{-1}(T_0,T_*). \tag{20}$$

Note that in the nonisotropic case, the distribution function is expanded in terms of Hermite polynomials where the weighting Maxwellian is characterized by the four-vector $W=(T,\mathbf{u})$. Here T and $\mathbf{u}$ are the temperature and velocity about which the expansion is performed. In this general case, we can determine the operator of the transition from one basis to another $\hat{D}(W_0,W_1)$, which will also have properties similar to Eqs. (19) and (20). This general matrix is best constructed using a representation of the Hermite polynomials in α–u space.[11,12]

We shall continue our analysis of the isotropic Boltzmann equation. Quite clearly, the transition to a new basis should not affect the result, i.e., the time derivatives of the distribution function should be the same in the bases T_1 and T_2,

$$\frac{d\varphi(T,t)}{dt}=\sum_{k=0}^{\infty}\frac{dC_k^0(t)}{dt}s_R^k(T,T_0)\sum_{r=0}^{\infty}\frac{dC_r^1(t)}{dt}s_R^r(T,T_1). \tag{21}$$

If both sides of Eq. (21) are multiplied scalarly by $s_L^r(T,T_1)$, then using the condition for orthogonality of the Sonine polynomials and Eq. (15), we obtain

$$\frac{dC_r^1(t)}{dt}=\sum_{k=0}^{\infty} d_{r,k}(T_1,T_0)\frac{dC_k^0(t)}{dt}. \tag{22}$$

We substitute Eq. (2) into Eq. (22) in the initial basis T_0 and express the vector C^0 in terms of C^1 using Eqs. (14) and (20): $C^0=\hat{D}(T_0,T_1)C^1$. Then, we have

$$\frac{dC_r^1}{dt}=\sum_{k=0}^{r} d_{r,k}(T_1,T_0)\sum_{k_1,k_2=0}^{\infty}K_{k_1,k_2}^k(T_0)\times\sum_{r_1=0}^{k_1} d_{k_1,r_1}(T_0,T_1)C_{r_1}^1\sum_{r_2=0}^{k_2} d_{k_2,r_2}(T_0,T_1)C_{r_2}^1. \tag{23}$$

If the distribution function is expanded directly in terms of Sonine polynomials with the temperature T_1, instead of the expression (23), the system of moment equations can be written as

$$\frac{dC_r^1}{dt}=\sum_{r_1,r_2=0}^{\infty}K_{r_1,r_2}^r(T_1)C_{r_1}^1C_{r_2}^1. \tag{24}$$

Since the expansion coefficients $C_{r_1}^1$ and $C_{r_2}^1$ are arbitrary, we obtain from expressions (23) and (24)

$$K_{r_1,r_2}^r(T_1)=\sum_{k=0}^{r} d_{r,k}(T_1,T_0)\times\sum_{k_1=r_1,k_2=r_2}^{\infty}K_{k_1,k_2}^k(T_0)d_{k_1,r_1}\times(T_0,T_1)d_{k_2,r_2}(T_0,T_1). \tag{25}$$

Hence, the invariance of the description of the relaxation process relative to the choice of basis yields a relation between the matrix elements in the bases T_1 and T_0. Note that for fixed T_0 formula (25) holds for any T_1. The form of the dependence $K_{r_1,r_2}^r(T_1)$ on the left-hand side of expression (25) is determined by the energy dependence of the cross section whereas the expression on the right only depends on T_1 via the matrix elements of the transition matrix D, which are the same for all cross sections. After substituting the matrix elements of D (17) and a series of simple transformations, expression (25) gives

$$K_{r_1,r_2}^r(T_1)=(1-z)^R(-1)^{N+q}\sum_{q=0}^{\infty}z^qB_{r_1,r_2}^r(q,T_0), \tag{26}$$

$$B_{r_1,r_2}^r(q,T_0)=\sum_{k=\max(q+R,N)}^{\infty}\begin{pmatrix} r\\ k-q-R\end{pmatrix}(-1)^k\times\sum_{k_1=r_1}^{k-r_2}\begin{pmatrix} k\\ r_1\end{pmatrix}\begin{pmatrix} k-k_1\\ r_2\end{pmatrix}K_{k_1,k-k_1}^{k-q-R}(T_0). \tag{27}$$

Here we have $z=1-T_1/T_0$, $R=N-r$, and $N=r_1+r_2$. The expressions (25) or (26), (27) impose certain relations on the matrix elements K in the initial reference frame T_0. This can be seen particularly clearly for Maxwellian molecules for which, as we know (see Ref. 9, for example), the elements $K_{r_1,r_2}^r(T)$ do not depend on temperature. For this model Eq. (26) has a constant on the left and a power series with respect to z on the right. Equating to zero the coefficients of all positive powers z, we obtain relations containing different $K_{r_1,r_2}^r(T_1)$ and coefficients which do not depend on T. For example, equating the coefficients of the first power z, we can obtain

$$RK^r_{r_1,r_2}+\sum_{k=\max(1+R,N)}^{N+1}\binom{r}{k-R-1}(-1)^{N+k}$$
$$\times\sum_{k_1=r_1}^{k-r_2}\binom{k_1}{r_1}\binom{k-k_1}{r_2}K^{k-R-1}_{k_1,k-k_1}=0. \tag{28}$$

This formula shows that some linear combinations of the matrix elements of the matrix K vanish. For arbitrary interaction models we can derive more general relations by successively differentiating both sides of expression (25) with respect to T_1 and then setting $T_1=T_0$. Formula (25) is written in the operator form

$$\hat{\hat{K}}(T_1)=\hat{D}(T_1,T_0)\hat{\hat{K}}(T_0)(\hat{D}(T_0,T_1),\hat{D}(T_0,T_1)). \tag{29}$$

The rules for the action of the bilinear operator $\hat{\hat{K}}$ on the vector are clear from Eq. (2) and the rules of action when the operator is expressed in the new basis are clear from expression (25). In general when Hermite polynomials are considered as the basis functions, relations similar to (19) and (20) are satisfied for the transition operator $\hat{\hat{D}}(W_0,W_1)$ and the following inequality should be satisfied instead of Eq. (29)

$$\hat{\hat{K}}(W_1)=\hat{D}(W_1,W_0)\hat{\hat{K}}(W_0)(\hat{D}(W_0,W_1),\hat{D}(W_0,W_1)). \tag{30}$$

COUPLING BETWEEN MATRIX ELEMENTS

We shall analyze arbitrary cross sections for particle interaction. We shall differentiate Eq. (29) with respect to T_1 and set $T_1=T_0$. We then assume that for $T_1=T_0$ the operator $\hat{D}$ becomes the unit operator $\hat{E}$

$$\left(\frac{d\hat{\hat{K}}(T_1)}{dT_1}\right)_{T_1=T_0}=\left(\frac{d\hat{D}(T_1,T_0)}{dT_1}\right)_{T_1=T_0}\hat{\hat{K}}(\hat{E},\hat{E})$$
$$+\hat{E}\hat{\hat{K}}\left(\left(\frac{d\hat{D}(T_0,T_1)}{dT_1}\right)_{T_1=T_0},\hat{E}\right)$$
$$+\hat{E}\hat{\hat{K}}\left(\hat{E},\left(\frac{d\hat{D}(T_0,T_1)}{dT_1}\right)_{T_1=T_0}\right). \tag{31}$$

In order to calculate the derivative of the matrix elements of the matrix D, we consider the explicit form $d_{r,k}(T_1,T_0)$ given by Eq. (17). Quite clearly, the derivative of this function for $T_1=T_0$ is only nonzero if $r=k+1$ or $r=k$,

$$\frac{d}{dT_1}(d_{r,k}(T_1,T_0))_{T_1=T_0}=r(\delta_{r-1,k}-\delta_{r,k})/T_0. \tag{32}$$

Similarly we have

$$\frac{d}{dT_1}(d_{k_1,r_1}(T_0,T_1))_{T_1=T_0}$$
$$=(-(r_1+1)\delta_{r_1+1,k_1}+r_1\delta_{r_1,k_1})/T_0, \tag{33}$$

$$\frac{d}{dT_1}(d_{k_2,r_2}(T_0,T_1))_{T_1=T_0}$$
$$=(-(r_2+1)\delta_{r_2+1,k_2}+r_2\delta_{r_2,k_2})/T_0. \tag{34}$$

Converting from the operator to the matrix form in Eq. (31) and substituting Eqs. (32)–(34), we obtain a fundamental relation linking the matrix elements $K^r_{r_1,r_2}$,

$$T\frac{dK^r_{r_1,r_2}(T)}{dT}-RK^r_{r_1,r_2}(T)$$
$$=rK^{r-1}_{r_1,r_2}(T)-(r_1+1)K^r_{r_1+1,r_2}(T)$$
$$-(r_2+1)K^r_{r_1,r_2+1}(T),$$
$$(R=r_1+r_2-r). \tag{35}$$

Here T_0 is replaced by T and we stress than Eq. (35) should be satisfied for arbitrary T. In the particular case of Maxwellian molecules, formula (35) is the same as (28).

In order to obtain relationships which appear under repeated differentiation, we divide both sides of equation (26) by $(1-z)^R$, differentiate twice with respect to T_1, setting $T_1=T_0$ $(z=0)$ and denoting $T_0=T$. As a result, we have

$$\left(T^2\frac{d^2}{dT^2}-2RT\frac{d}{dT}+R(R+1)\right)K^r_{r_1,r_2}(T)=B^r_{r_1,r_2}(2,T). \tag{36}$$

We shall now show that the relations (36) are a consequence of (35) and do not impose additional correlations on the matrix elements. Using the fact that the identity (35) is valid for any T, we apply the operator Td/dt to this. Expressing the derivatives on the right-hand side of the resulting equality in terms of the matrix elements using (35) and performing simple transformations, we obtain

$$\left(T^2\frac{d^2}{dT^2}-T(R-1)\frac{d}{dT}\right)K^r_{r_1,r_2}(T)-B^r_{r_1,r_2}(2,T)$$
$$=(R+1)(rK^{r-1}_{r_1,r_2}(T)-(r_1+1)K^r_{r_1+1,r_2}(T)$$
$$-(r_2+1)K^r_{r_1,r_2+1}(T)). \tag{37}$$

Now, on the right-hand side of Eq. (37) we again use equality (35)

$$\left(T^2\frac{d^2}{dT^2}-(R-1)T\frac{d}{dT}\right)K^r_{r_1,r_2}(T)$$
$$=(R+1)\left(T\frac{d}{dT}-R\right)K^r_{r_1,r_2}(T)+B^r_{r_1,r_2}(2,T). \tag{38}$$

On comparing expressions (36) and (38), we note that these equalities are the same. Similarly, we can show that the relations obtained for higher derivatives of Eq. (25) also impose no additional correlations on the matrix elements compared with relation (35). Hence we can confirm that in order to satisfy formula (25), it is necessary and sufficient that relation (35) is satisfied for any T.

TABLE I. Verification of formula (45).

	K^2_{22}	K^2_{32}	K^2_{23}	Right-hand side of Eq. (45)
From Eq. (39)	-1.6667×10^{-2}	3.9063×10^{-4}	-8.7240×10^{-3}	0
From Ref. 13	1.7578×10^{-2}	-3.4180×10^{-3}	-1.2533×10^{-2}	7.4218×10^{-2}

FUNDAMENTAL RELATIONS FOR VARIOUS PARTICLE INTERACTION LAWS

In Ref. 9, for arbitrary power potentials we made analytic calculations of the matrix elements in the form of four nested sums, having the form

$$
\begin{aligned}
K^r_{r_1,r_2}={}&n_0^2\left(\frac{4kT_0}{m}\right)^{\mu}\frac{(-1)^{r+r_1}r!\,\Gamma(\mu+5/2)}{2^{r_2}}\\
&\times\sum_{q=1}^{r}\frac{(\mu+5/2)_{q-1}}{q!}J_q\sum_{i=\max(0,q-r_1)}^{\min(r_2,q)}\binom{q}{i}\frac{2^i}{(r_2-i)!}\\
&\times\sum_{l=\max(q,r-r_1+q-i)}^{r}\frac{(2l-2q-1)!!(-1)^l}{\Gamma(l+3/2)(r-l)!}\\
&\times\sum_{j=\max(0,m_2)}^{m_1}\frac{(-1)^j(-\mu)_{j+m_3}}{j!2^j(m_1-j)!(j-m_2)!},
\end{aligned}
\tag{39}
$$

where $\mu=\gamma/2$, $m_1=r_1-r+l-q+i$, $m_2=r_1-r-l+q+i$, and $m_3=l-q+r_2-i$.

Here we used the following notation

$$a_k=a(a+1)\dots(a+k-1)=\Gamma(a+k)/\Gamma(a),$$

$$J_q=4\pi\int_0^1F(z)z^q dz. \tag{40}$$

In this case, the scattering cross section is given in the form

$$g\sigma(g,z)=g^{\gamma}F(z). \tag{41}$$

Here $F(z)$ is the angular part of the cross section, $z=\sin^2\Theta/2$, and Θ is the scattering angle. It can be seen from Eqs. (39) and (40) that when the scattering cross section depends on the angle and velocity as given by Eq. (41), the matrix elements $K^r_{r_1,r_2}$ depend quite specifically on the basis temperature , i.e., a general coefficient—the collision frequency $\tau(T)=bT^{\mu}$ depends on temperature. Relation (35) now has the form

$$
\begin{aligned}
(\mu-R)K^r_{r_1,N-r_1}={}&rK^{r-1}_{r_1,N-r_1}-(r_1+1)K^r_{r_1+1,N-r_1}\\
&-(N+1-r_1)K^r_{r_1,N+1-r_1}.
\end{aligned}
\tag{42}
$$

It is interesting to note that the relations between the matrix elements (42) do not depend on the form of the function $F(z)$. It is well-known that for pseudo-Maxwellian molecules $(\mu=0)$ a simple analytic formula exists which, correct to within the normalization constant, has the form

$$K^r_{r_1,r-r_1}=\frac{1}{r+1}-\delta_{r,r_1}. \tag{43}$$

For nonzero matrix elements in cases of Maxwellian molecules Eq. (42) gives

$$
\begin{aligned}
&(r+1)K^r_{r_1,r-r_1}-(r_1+1)K^{r+1}_{r_1+1,r-r_1}\\
&\quad-(r-r_1+1)K^{r+1}_{r_1,r-r_1+1}=0.
\end{aligned}
\tag{44}
$$

It is easily verified that the matrix elements (43) satisfy the identity (44).

In Ref. 10 we used the relations (4) as the criterion for the accuracy of the calculations which, as can be shown, are a consequence of the relations (35) which are more detailed criteria for the accuracy of the calculations. Assuming that r_1 and r_2 vary in the range $0\leqslant r_1+r_2\leqslant N_0$, and r varies in the range $0\leqslant r\leqslant N_0$, then the total number of matrix elements is $(N_0+1)^2(N_0+2)/2$, and the total number of new, more detailed relations of the type (35) or (42) is $(N_0+1)^2N_0/2$, whereas the total number of relations obtained from formula (4) is only $(N_0+1)^2$. Thus, for large N_0 the number of relations (42) far exceeds the number of relations (4).

For other laws of interaction the relations (42) were checked by calculating the matrix elements using formulas (39)–(41) in the range $r,N\leqslant20$. We showed in Ref. 9 that small numerical errors are known to occur in this range. A check was made for various values of μ and it was established that all the matrix elements satisfy the equalities (42).

Schürrer and Kügerl[13] gave formulas for the matrix elements for the hard sphere model $(\mu=0.5)$. Separate formulas are given to calculate the linear elements which after various simplifications agree with Eqs. (39)–(41). Special formulas are given for the nonlinear elements which do not include the matrix elements for which one of the subscripts is 1. It is therefore difficult to verify the identities (42) when these relate nonlinear and linear elements. We shall therefore analyze the identity (42) for $N=4$ and $r=r_1=2$

$$1.5K^2_{2,2}-3K^2_{3,2}-3K^2_{2,3}=-2K^1_{2,2}. \tag{45}$$

The values of these elements calculated using the formulas (39) and the formulas from Ref. 13 in the same units of measurement are given in Table I. Despite the substantial differences between the corresponding elements, the relations (45) are satisfied in both cases. Moreover, other sets of elements from Ref. 13 with subscripts not including 0 or 1, also satisfy the relations (42).

It follows from the energy conservation law that $K^1_{2,2}=0$. The main error in Ref. 13 is that the formulas proposed there yield $K^1_{2,2}\neq0$ and the identity (45) is only satisfied for this inaccurate value of $K^1_{2,2}$. If a zero value for $K^1_{2,2}$ is substituted into identity (45) and all the other elements therein are taken from Ref. 13, criterion (42) is satisfied. This implies that there is an error in the determination of the other

TABLE II. $K^r_{r_1,r_2}$ for $r_1+r_2=4$ for Coulomb particle interaction.

	$r_1=0$	1	2	3	4
r	$r_2=4$	3	2	1	0
1	0.41139	0.82278	0.00000	−0.82278	−0.41139
	0.41139	0.82278	0.00000	−0.82278	−0.41139
2	0.74050	0.91681	0.56419	−0.91681	−1.30469
	0.76871	0.91681	0.50777	−0.91681	−1.27648
3	0.87273	0.89918	0.84628	−0.12543	−2.49276
	0.92794	0.91207	0.76166	−0.21570	−2.38597
4	0.84041	0.83453	0.84628	0.94222	−3.46344
	0.90469	0.86414	0.80751	0.67380	−3.25014
5	0.71626	0.72728	0.70524	0.58184	−2.73061
	0.76813	0.74909	0.66598	0.38743	−2.57062
6	0.56199	0.59504	0.52893	0.23912	−1.92508
	0.59933	0.60895	0.49407	0.11776	−1.82010
7	0.41525	0.46024	0.37025	0.03995	−1.28568
	0.44046	0.46858	0.34304	−0.03150	−1.22058
8	0.29311	0.33940	0.24683	−0.04957	−0.82977
	0.30942	0.34423	0.22735	−0.09032	−0.79068
9	0.19959	0.24050	0.15868	−0.07628	−0.52248
	0.20981	0.24323	0.14552	−0.09905	−0.49951
10	0.13200	0.16483	0.09917	−0.07325	−0.32275
	0.13825	0.16636	0.09065	−0.08580	−0.30947

TABLE III. $K^r_{r_1,r_2}$ for $r_1+r_2=5$ for Coulomb particle interaction.

	$r_1=0$	1	2	3	4	5
r	$r_2=5$	4	3	2	1	0
1	.23141	.69422	.46281	−.46281	−.69422	−.23141
	.23141	.69422	.46281	−.46281	−.69422	−.23141
2	.50909	.86906	.68908	−.03085	−1.19817	−.83820
	.52555	.88551	.65616	−.06377	−1.18171	−.82175
3	.70909	.85785	.78347	.48595	−1.00898	−1.82740
	.74889	.88153	.75224	.39024	−1.01754	−1.75536
4	.78788	.81212	.80000	.75152	−.14786	−3.00367
	.84494	.84411	.78404	.63525	−.28066	−2.82768
5	.76102	.75505	.75804	.76998	.86552	−3.90961
	.82273	.79389	.75985	.69699	.54218	−3.61564
6	.66736	.67975	.67356	.64876	.51849	−3.18792
	.71941	.71203	.67098	.57863	.27711	−2.95816
7	.54501	.58599	.56550	.48354	.16171	−2.34175
	.58483	.61047	.55893	.42238	.00670	−2.18331
8	.42135	.48242	.45188	.32975	−.05400	−1.63140
	.45001	.50007	.44388	.28145	−.14728	−1.52813
9	.31186	.38019	.34602	.20935	−.14947	−1.09795
	.33160	.39247	.33842	.17385	−.20347	−1.03287
10	.22275	.28810	.25542	.12474	−.17066	−.72036
	.23591	.29638	.24910	0.9999	−.20109	−.68029

three elements in identity (45). If these inaccurate elements are changed so that identity (45) is satisfied, criterion (42) is violated for subsequent values of the indices, and so on. In Ref. 10 we analyzed how this inaccurate determination of the nonlinear matrix elements gives such erroneous results in the course of a relaxation process.

Formulas (39)–(41) can also be applied to Coulomb particle interaction. In this case, in order to eliminate divergence, we can use the Coulomb potential with the Debye cutoff radius λ_D following the usual procedure.[9] This corresponds to the fact that when J_q is determined using formula (40), integration over angles is not performed from zero but from some small angle θ_0 which is assumed to be $2e^2/(\lambda_D m g^2)$. It can be shown that the first term in the summation over q in formula (39) corresponds to the Landau approximation.[11] For the complete Boltzmann equation summation over q is performed as far as r and the number of correction terms increases with increasing r. Tables II and III give the dimensionless matrix elements calculated using formulas (40) and (41) for $N=4$ and 5. For the calculations of the matrix elements the unit of measurement of time in the Boltzmann equation was taken to be $\tau=16(kT_0/m)^{3/2}/(J_1 n_0)$. In this case we have

$$J_1=8\pi\ln\Lambda e^4/m^2.$$

In our calculations the Coulomb logarithm $\ln\Lambda$ was taken to be 5. The upper number in each line in the table corresponds to $q=1$ (Landau approximation) and the lower line gives the total sum over q (Boltzmann equation for Coulomb particle interaction). Substituting the values from the tables into relation (42) shows that these identities are satisfied for the Landau approximation and for the complete Boltzmann equation. In addition, if the summation over q in Eq. (39) is artificially truncated at $r=2,3,\ldots,$ these sums also satisfy the relations (42). This was checked as far as $N_0=20$.

RECURRENCE FORMULAS

The relations (42) can be used as recurrence formulas to calculate the matrix elements. This is particularly important for high values of the indices. We shall show that in order to determine the nonlinear matrix elements, it is sufficient to define the linear elements of just one of the two types $K^r_{r_1,0}$ or K^r_{0,r_1}. Then, one half of the linear elements is defined in terms of the other. Let us assume that all linear elements of the type $K^r_{q,0}(r,q\leqslant N_0)$ are defined. Formula (42) can then be written as:

$$(q+1)K^r_{N-q.q+1}=(R-\mu)K^r_{N-q,q}+rK^{r-1}_{N-q,q} -(N-q+1)K^r_{N-q+1,q}. \tag{46}$$

We shall first analyze this formula for $N=0$

$$K^r_{0,1}=(-\mu-r)K^r_{0,0}+rK^{r-1}_{0,0}-K^r_{1,0}. \tag{47}$$

We then find all $K^r_{0,1}$ from Eq. (47). Converting to $N=1$ in Eq. (46) we obtain two relations

$$K^r_{1,1}=(1-\mu-r)K^r_{1,0}+rK^{r-1}_{1,0}-2K^r_{2,0}, \tag{48}$$

$$2K^r_{0,2}=(1-\mu-r)K^r_{0,1}+rK^{r-1}_{0,1}-K^r_{1,1}. \tag{49}$$

For fixed r we find $K^r_{1,1}$ from Eq. (48) and then determine $K^r_{0,2}$ in terms of the elements from Eq. (49) $K^r_{1,1}$ just determined and the linear elements with subscripts 0, 1 determined earlier using formula (47). In this way all elements with $N=2$ are determined.

For $N=2$ we obtain three relations from Eq. (46) from which all $N=3$ elements $K^r_{2,1}$, $K^r_{1,2}$, and $K^r_{0,3}$ can be systematically determined in terms of the $N=2$ elements: $K^r_{2,0}$, $K^r_{1,1}$, and $K^r_{0,2}$.

We shall assume that all matrix elements for which the sum of the subscripts is equal to N are known and we write the corresponding equalities (46) for fixed r and $q=0,1,\dots N$

$$K^r_{N,1}=(N-r-\mu)K^r_{N,0}+rK^{r-1}_{N,0}-(N+1)K^r_{N+1,0}, \tag{50}$$

$$2K^r_{N-1,2}=(N-r-\mu)K^r_{N-1,1}+rK^{r-1}_{N-1,1}-NK^r_{N,1}, \tag{51}$$

$$\begin{aligned}3K^r_{N-2,3}=&(N-r-\mu)K^r_{N-2,2}+rK^{r-1}_{N-2,2}\\&-(N-1)K^r_{N-1,2},\end{aligned} \tag{52}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots,$$

$$(N+1)K^r_{0,N+1}=(N-r-\mu)K^r_{0,N}+rK^{r-1}_{0,N}-K^r_{1,N}. \tag{53}$$

Quite clearly, moving from top to bottom, we can use Eqs. (53)–(55) to systematically determine all elements for which the sum of the subscripts is $N+1$ in terms of already known elements. Thus, all the nonlinear elements can be determined for given $K^r_{q,0}$. In addition, the linear elements of the other type $K^r_{0,q}$ can also be determined in terms of these. Conversely we can assume that linear elements of the type $K^r_{0,q}$ are known and by reducing q, we can determine all nonlinear elements and linear ones of the type $K^r_{q,0}$.

As a result, we can construct such a recurrence procedure. Initially, N is increased from $N=0$ and then r is increased, also from zero. Then, for given r and N the values of q are varied between 0 and N and the elements $K^r_{N-q,q+1}$ are determined from formula (46). Assuming that linear elements of the type $K^r_{0,q}$ are known, we obtain a similar recurrence procedure by replacing $N-q$ with q in formula (46) and then determining the extreme right element $K^r_{q+1,N-q}$.

Note that when a gas of the one species is being studied, a knowledge of the symmetrized matrix elements is sufficient

$$\tilde{K}^r_{r_1,r_2}=(K^r_{r_1,r_2}+K^r_{r_2,r_1})/2, \quad r_1\geqslant r_2. \tag{54}$$

In this case, instead of formulas (50)–(53) we have

$$\begin{aligned}(q+1)\tilde{K}^r_{N-q,q+1}=&(N-r-\mu)\tilde{K}^r_{N-q,q}+r\tilde{K}^{r-1}_{N-q,q}\\&-(N+1-q)\tilde{K}^r_{N+1-q,q}\\q=0,\dots,&[(N-1)/2].\end{aligned} \tag{55}$$

Here the square brackets denote the integer part of the number.

Understandably, when the recurrence formulas (46) are used, we need to have simple formulas to calculate the linear elements. We now go over to the linear case in formula (39). Assuming that $r_1=0$, we then have $i=q$, $m_1=l-r\geqslant 0$, from which it follows that $r=l$, and then $m_1=0$ and $j=0$. As a result we obtain

$$\begin{aligned}K^r_{0,r_2}=&\left(\frac{4kT_0}{m}\right)^{\mu}\frac{r!n_0^2}{\Gamma(r+3/2)2^{r+r_2}}\\&\times\sum_{q=1}^{\min(r_2,r)}2^{2q}W^{r_2}_{r,q}J_q\Gamma(q+\mu+3/2)/q!,\end{aligned} \tag{56}$$

where

$$\begin{aligned}W^{r_2}_{r,q}=&\frac{-\mu(-\mu+1)\dots(-\mu+r_2+r-2q-1)}{(r_2-q)!(r-q)!}\\=&\frac{\Gamma(-\mu+r_2+r-2q)}{\Gamma(-\mu)(r_2-q)!(r-q)!}.\end{aligned} \tag{57}$$

The similar formula for $K^r_{r_1,0}$ has the form

$$\begin{aligned}K^r_{r_1,0}=&\left(\frac{4kT_0}{m}\right)^{\mu}\frac{r!n_0^2}{\Gamma(r+3/2)2^{r+r_1}}\\&\times\sum_{q=1}^{\min(r_1,r)}2^{2q}W^{r_1}_{r,q}\tilde{J}_q\Gamma(q+\mu+3/2)/q!.\end{aligned} \tag{58}$$

Here we have

$$\tilde{J}_q=4\pi\int_0^1 F(z)((1-z)^q-1)dz. \tag{59}$$

In practice, an approximation in the form of a finite sum as far as some N_0 rather than the series (1) must be used to represent the distribution function. The main errors incurred when using the moment method are caused by this truncation. In order to calculate the distribution function at high velocities, we need to strive to increase N_0. However, when formulas with multiple summation are used for the calculations, both the errors and the computation time begin to increase catastrophically with increasing N_0. This is quite understandable since calculations of N_0^3 elements involving six sums, as in Ref. 6, requires a number of summations proportional to N_0^9. For four sums, as in Ref. 9, this number of summations is proportional to N_0^7. In this sense, calculations using recurrence formulas are advantageous since the number of summations for calculating any element is three and the computation time only increases proportionately as N_0^3.

The calculations were made using a 66 MHz PC-486. The calculations using formula (39) allowing for all the simplifications considered in Ref. 9 for $N_0=30$ took 3.5 h whereas calculations of all the nonlinear elements to double precision using the recurrence formula (46) for $N_0=30$ only required 0.23 s. Quite clearly if the number of summations is N_0^9, a time of 3.5 h is required for $N_0=14$ and if calculations up to $N_0=30$ could be made using the formulas from Ref. 6 with sixfold summation and no significant errors, nine months would be required.

CONCLUSIONS

It has been shown that the conditions of invariance in the description of the relaxation process relative to the choice of basis impose certain relations on the matrix elements which are a mapping of the collision integral in the moment method of solving the Boltzmann equation. These relations exist for arbitrary particle interaction laws. Using these relations as recurrence formulas can reduce the time required to calculate nonlinear matrix elements by many orders of magnitude.

This opens up the possibility of calculating matrix elements with large indices. As a result, the distribution function for high velocities can be calculated ''almost analytically.''

The principle of invariance of the collision integral relative to the choice of basis functions may prove highly effective for calculating the matrix elements in cases of nonisotropic relaxation. Therefore, the distribution function must be expanded in terms of Hermite polynomials $H_{r,l,m}$, which are the product of spherical harmonics and Sonine polynomials with a weighting Maxwellian, which depends on the temperature and average temperature. In this case, the relationship between the matrix elements will be determined by going over to a basis with a Maxwellian distribution not only with a different temperature but also with a different shear velocity.

So far, the moment method has been little developed in the nonisotropic case because of the major difficulties involved in deriving the formulas and calculating the nonlinear matrix elements. The formulas known from transport theory[1,4] have only been derived for linear elements and $l \leqslant 2$. However, in nonisotropic problems the apparatus of $\alpha-u$ transformations developed in Refs. 11 and 12 can be used to obtain simple relations between the matrix elements similar to those obtained here. As a result, even in nonisotropic problems, progress can be made toward high velocities when constructing the distribution function. This is particularly important for problems in physicochemical kinetics and for the description of highly nonequilibrium transport processes. Particular mention should be made of the need to use high moments and electron transport matrix elements in a thermonuclear plasma.

This work was partly supported by the Russian Fund for Fundamental Research (Project No. 97-02-18080) and partly by the Federal Target Program ''State Support for Integration of Higher Education and Fundamental Science'' (Contract No. 326.53).

[1] S. Chapman and T. G. Cowling, *Mathematical Theory of Non-Uniform Gases*, 2nd ed. [Cambridge University Press, Cambridge, 1952; IL, Moscow 1960].

[2] D. Burnett, Proc. London Math. Soc. No. 2, 39 (1933).

[3] H. Grad, Commun. Pure Appl. Math. 331 (1940).

[4] M. H. Ernst, Phys. Rep. **78**, 1 (1981).

[5] G. A. Bird, *Molecular Gas Dynamics* [Clarendon Press, Oxford, 1976; Mir, Moscow, 1981].

[6] G. Turchetti and M. Paolilli, Phys. Lett. A **90**, 123 (1982).

[7] I. A. Énder and A. Ya. Énder, Dokl. Akad. Nauk SSSR **193**, 61 (1970) [Sov. Phys. Dokl. **15**, 633 (1970)].

[8] I. N. Kolyshkin, A. Ya. Énder, and I. A. Énder, Zh. Vychisl. Mat. Mat. Fiz. **28**, 901 (1988).

[9] A. Ya. Énder and I. A. Énder, Zh. Tekh. Fiz **64**(10), 38 (1994) [Tech. Phys. **39**, 997 (1994)].

[10] A. Ya. Énder and I. A. Énder, Zh. Tekh. Fiz **68**(5), 18 (1998) [Tech. Phys. **43**, 493 (1998)].

[11] A. Ya. Énder and I. A. Énder, *Integral Transformation of the Boltzmann Equation for Different Particle Interaction Laws*, Preprint No. 605 [in Russian], Physicotechnical Institute, Academy of Sciences of the USSR, Leningrad (1979).

[12] I. N. Kolyshkin, A. Ya. Énder, and I. A. Énder, *Modeling in Mechanics* [in Russian], Novosibirsk, (1990), pp. 54–64.

[13] F. Schürrer and G. Kügerl, Phys. Fluids A **2**, 609 (1990).

[14] J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport in Gases* [North-Holland, Amsterdam, 1972; Mir, Moscow, 1976].

Translated by R. M. Durham

Investigation of the charge distribution in the insulating envelope of a high-voltage vacuum device

V. D. Bochkov

''Plasma'' Research Institute of Gas-Discharge Devices, 390023 Ryazan, Russia

M. M. Pogorel'skiĭ

Ryazan State Radio Engineering Academy, 391000 Ryazan, Russia

(Submitted October 3, 1997; resubmitted May 13, 1998)
Zh. Tekh. Fiz. **69**, 30–35 (June 1999)

Methods of measuring charges in the bulk and at the surface of the insulating envelope of a high-voltage vacuum device are developed as part of studies of a complex range of dielectric strength aspects. These methods were used to measure the charge distribution over the length of the envelope and to study how this distribution is influenced by the operating regimes of the device. Laws governing the formation of charges were identified and its correlation with the appearance of physical changes in the structure of the surface layers of the insulator was determined. Processes of structural change were studied and simulated and the change in the dielectric strength of the device was compared with changes in the magnitude and polarity of the surface and bulk charges. © *1999 American Institute of Physics.* [S1063-7842(99)00506-1]

INTRODUCTION

The insulating envelope is one of the main components of any vacuum device. However, such an important parameter as the dielectric strength of the vacuum device is not generally related to the presence and properties of the insulating envelope. For instance, among the numerous studies of vacuum breakdown,[1] no consideration is given to the influence of the insulating envelope on the initiation of breakdown. However, a few authors do draw attention to the fact that effects associated with secondary electron emission from insulating components[2,3] and luminescence at the envelope[4] may affect the operating reliability of vacuum devices. In x-ray tubes and hydrogen thyratrons luminescence and breakdown of the vacuum interelectrode gap is accompanied by the observation of catastrophic damage to the glass or metal ceramic envelope.[5,6] All these processes are caused by exposure of the insulating elements to ion and electron beams.

The aim of the present paper is to study the charges formed in the insulating envelope of high-voltage vacuum devices, especially x-ray tubes, to identify the charge distribution pattern over the length of the envelope, and to determine the processes promoting electrical breakdown of the envelope and lowering the dielectric strength of the devices. It should be noted that studies of the interaction between charged particle beams and insulators are important for various practical applications concerned with the reliability not only of high-voltage vacuum devices[7–10] but also with larger objects using insulators in a vacuum, including accelerators, nuclear reactors, and space station equipment.[11,12]

DESCRIPTION OF METHODS OF MEASUREMENT

The glass envelope of x-ray tubes is exposed to the action of a complex range of factors: strong electric fields, and irradiation by hard x-rays, ions, and high-energy electrons. Under these conditions, polarization should occur in the glass envelope, similar to that in radio- or electro-electrets.[13–15]

Numerous methods are available for measuring the charges of insulators.[16–18] Three methods suitable for the real operating conditions of sealed high-voltage electric-vacuum devices proved the most convenient for measuring the charges in the glass envelope of x-ray tubes. For instance, during operation of the tubes the effective total surface charge was investigated by measuring the electrostatic force created by the charge field.[19] The charge formed in the bulk and at the surface of the envelope and its distribution over the envelope length away from the device were estimated by measuring the thermostimulated depolarization (TSDP) currents and by an electrostatic indication (ESI) method.[14,20,21]

Experimental prototypes were fabricated with S40-1 and S49-1 borosilicate glass envelopes (similar to Corning 7052) 1.7 ± 0.3 mm thick, 35–60 mm in diameter, and 12–18 cm long. The design was essentially that of small x-ray tubes. In order to measure the electrostatic force directly during operation under voltage, we used a prototype x-ray tube with a metal probe inserted in one of the high-voltage electrodes (Fig. 1). Its displacement under the action of the charge field was recorded by a special vacuum ''diode'' transducer having two planar anodes connected mechanically to the probe with a filamentary cathode inserted between them. The operating principle of the transducer is based on the change in the current distribution between the anodes of the diode as the probe is displaced. The displacements of the probe were calibrated to obtain direct measurements of the envelope potential (Fig. 2).

In order to measure the charge at the envelope, the high-voltage electrode in which the probe was inserted was con-

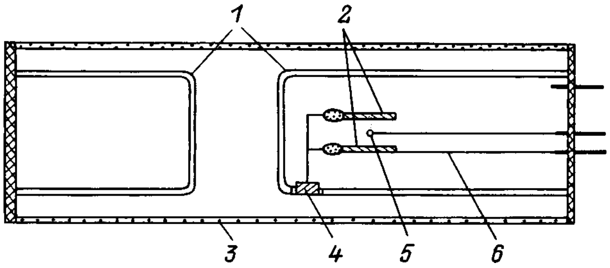


FIG. 1. Schematic of x-ray tube prototype with diode transducer to measure probe displacement: *1* — high-voltage devices, *2* — transducer anodes, *3* — glass envelope, *4* — probe, *5* — transducer cathode, and *6* — spring.

nected to it electrically and could be connected either as the cathode or as the anode of the tube. Figure 2 shows that the displacement of the probe depends linearly on the voltage in the range 0–15 V. As the voltage increases further, the deflection increases abruptly and the probe then ''attaches'' to the envelope. This method was used to measure small (up to 15 kV) potential differences between the envelope and the high-voltage electrode of the prototype.

Measurements of the charge using the ESI and TSDP methods were made using probes with an active area of 1 cm^2 inserted in the high-voltage electrodes of x-ray tube prototypes. The following types of probes were used: a flexible thin niobium foil probe which ensured good contact with the surface of the cylindrical envelope; shielded probes, having a shield to protect against the charge built up outside the surface of the probe; and insulated probes where the probe is insulated from the envelope surface by a thin layer of glass. We used stationary and moving probes which could be inserted at any point in the envelope.

The prototype being tested was held under conditions close to real operation, at high voltage in an oil bath. After the voltage had been switched off, the prototype was removed from the bath and an electrostatic voltmeter was connected to the probe output. The charge was first measured by the ESI method. A grounded electrode was brought into external contact with the envelope near the probe. The probe was then brought in contact with the glass. The envelope charge was determined from the voltage and the capacitance of the voltmeter and an additional capacitor connected to it. The charge thus recorded frequently passed over to the probe accompanied by breakdown. The charge density was several orders of magnitude higher than the induced charge typical of electrets. The prototype was then placed in a shielded chamber, the probe was mounted on the surface of the envelope at the point being studied, and its output was connected to a U5-6 electrometer amplifier to measure the TSDP currents. A grounded copper electrode with a heater to heat the area of the glass being studied was mounted externally, facing the probe. The rate of heating was varied between 3.7 and 16 deg/min. The inside of the envelope was heated to 150–180 °C while the outside was heated to 210–240 °C.

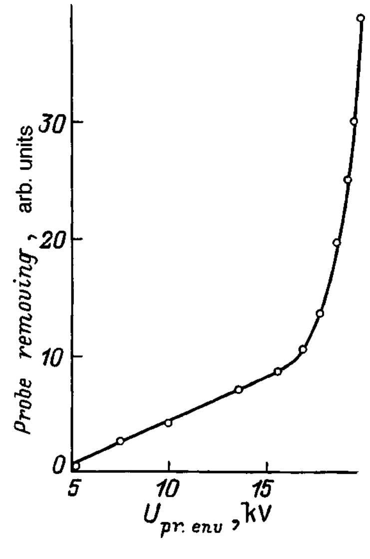


FIG. 2. Diode transducer probe deflection as a function of voltage between envelope and probe.

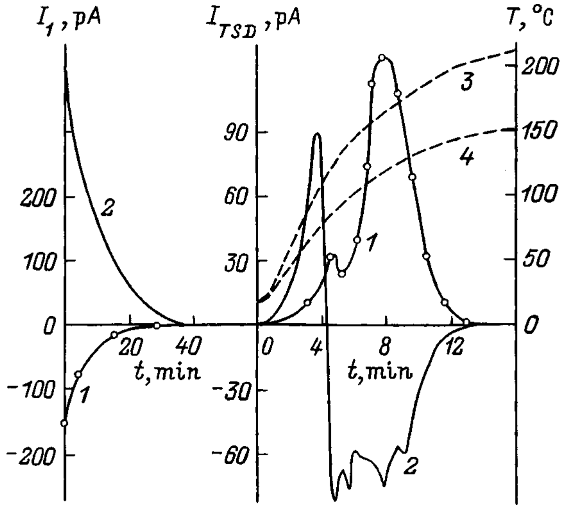


FIG. 3. Currents recorded by flexible probe in the cathode (*1*) and anode (*2*) regions of the envelope, without (I_1) and with heating (I_{TSD}): *3* — temperature of outer envelope surface, *4* — temperature of inner envelope surface, anode voltage 120 kV, and operating regime — with discharges in interelectrode gap.

RESULTS AND DISCUSSION

Investigation of the residual polarization of the envelope

The experimental dependences of the currents produced by relaxation of the charges in the envelope measured using the U5-6 amplifier have two characteristic sections (Fig. 3). Initially, after the irradiated prototype has been inserted in the chamber, we observe a discharge current I_1 which decreases over 10–60 min reaching low values, without any heating of the envelope. The polarity of this current depends on the measurement point. The positive direction was taken to be the direction of the current from the irradiated side of the envelope to the outside. After switching on the heater, we observe a second type of current, a thermostimulated depolarization current I_{TSD}.

In order to analyze the relaxation processes in the envelope of x-ray tube prototypes, it is important to know the direction of current flow in the measuring circuit and the factors determining this. This allows us to assess how the residual polarization corresponds to the particular process.

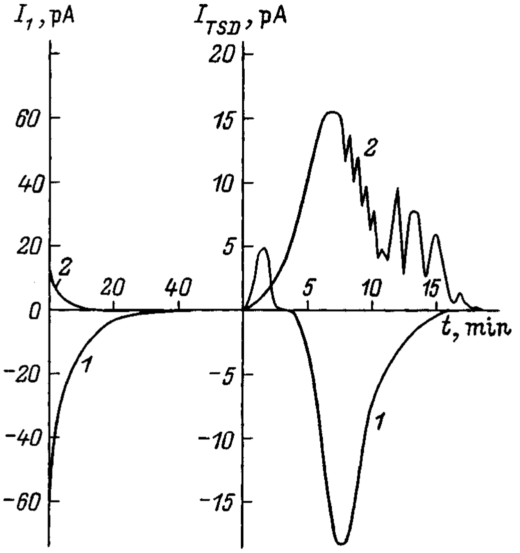


FIG. 4. Currents recorded by insulated probe in the anode region of the envelope, without (I_1) and with heating (I_{TSD}): anode voltage 100 (*1*) and 120 kV (*2*); operating regime – with (*1*) and without discharges (*2*).

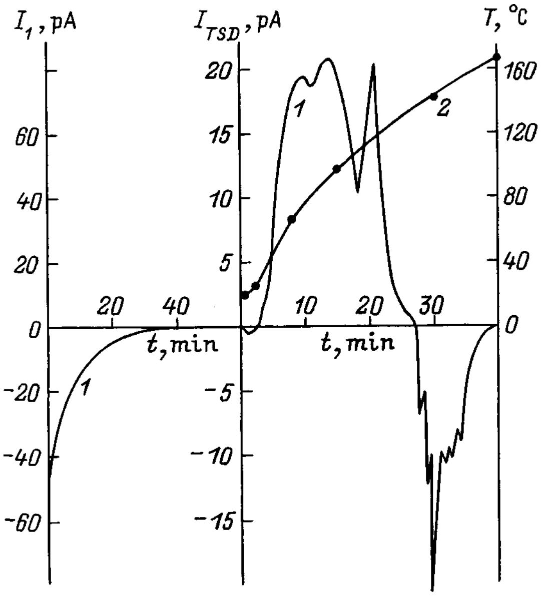


FIG. 5. Currents recorded by flexible probe in the anode part of the envelope (*1*), without (I_1) and with heating (I_{TSD}): anode voltage 120 kV, operating regime — without discharges, *2* — temperature of inner surface of envelope.

The TSDP current has different directions when the ''internal'' residual polarization (heterocharge) and the ''external'' residual polarization (coupled homocharge) decay.[14] The polarity of the displacement current observed when the heterocharge decays is opposite to that of the charged insulator. The homocharge usually discharges through the bulk of the sample. In this case, the TSDP current flows in the opposite direction. If the homocharge reaches the electrodes, the TSDP current has the same direction as the current caused by the decay of the heterocharge. However, the direction of the current can be altered by using insulating spacers in the insulator–electrode gap. The experimental characteristics show that when insulation is inserted near the probe, only the direction of the current I_1 changes while the direction of the TSDP current remains as before (Fig. 4).

Consequently, the discharge current I_1 in the cathode and anode regions of the envelope is caused by the release of weakly bound homocharge from surface traps, which relaxes without heating. An analysis of a large number of TSDP current measurements also yielded the following conclusions. The thermostimulated current is always positive in the cathode region and is solely due to homocharge discharging through the bulk of the envelope. In the anode region the TSDP currents produced by the buildup of different charges can be either positive or negative. For example, when discharges and breakdown occur in the vacuum high-voltage gap of the prototype, a heterocharge appears in the anode region of the envelope, whose current is positive. The discharging heterocharge peak is followed by a peak caused by the decay of homocharge, which discharges through the bulk of the envelope as in the cathode region. This homocharge is activated at temperatures of 80–120 °C which indicates that its is distributed in deep traps. In the envelopes of devices operating at voltages above 100 kV without any discharges, we observe negative TSDP currents which appear after the positive current when the temperature of the internal envelope surface is around 120 °C, regardless of the rate of heating (Fig. 5).

No such change in polarity was observed for the TSDP currents measured in the anode region using insulated probes. From this we can conclude that the negative TSDP currents in the anode part of the envelope for tubes operating without discharges are currents from negative charges which reach the surface of the envelope during heating, i.e., they are caused by the electron space charge. The TSDP currents are caused by the relaxation of charges trapped in comparatively deep traps so that these charges may be described as bulk charges, in contrast to those recorded by the ESI method, which are mainly attributable to the relaxation of surface charge and are therefore subsequently called surface charges.

Charge distribution over the tube envelope

In order to identify the processes responsible for the appearance of different charges in the anode region of the envelope when the tube is operated under various conditions, we studied the influence of electric fields on the formation of these charges. For this we used an x-ray tube prototype having a probe with a ''diode'' transducer. Measurements were made of the potentials between the high-voltage electrodes and the envelope under conditions with normal switching and with current flow between the anode and the cathode eliminated. In this case, only a high voltage of different polarity was connected to the mutually shorted cathode and anode and the envelope was exposed to the action of an electric field between the grounded external circuit and the high-voltage electrodes.

The results of the measurements show that the electric fields in the anode and cathode regions differ substantially (Fig. 6). In the cathode region, for instance, the small deflec-

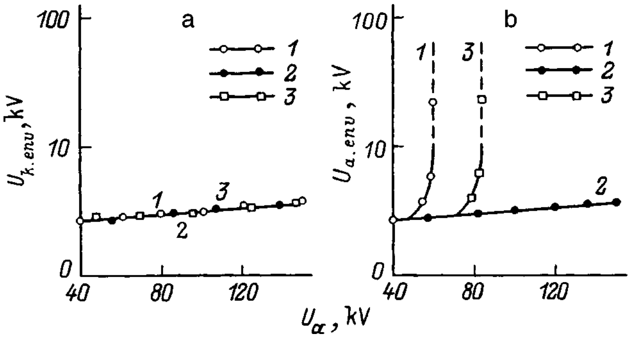


FIG. 6. Voltage between cathode and envelope (a) and between anode and envelope (b) during operation of prototype: *1* — with normal switching and no discharges, *2* — with normal switching after the appearance of discharges, and *3* — with no current flowing.

tion of the probe indicates that a potential close to the probe potential, i.e., close to the cathode potential, is formed at the envelope. When the operating conditions are varied, for example, the leakage currents increase or discharges and breakdowns take place between the cathode and the anode, the envelope potential varies little in this region. In the anode region, however, the envelope potential depends very strongly on the operating conditions. When the leakage currents (field emission currents) exceed 0.05 μA, the electrons reaching the envelope near the anode lower its potential, causing an abrupt increase in the deflection of the probe at an anode voltage around 60 kV. The probe remains in this position until discharges appear in the device. The probe then returns to a position close to zero, producing a small potential difference between the envelope and the anode.

After the voltage had been switched off, the charges on the envelope were investigated by the ESI and TSDP methods under normal conditions and in the absence of current flow. Figure 7 gives the maximum densities of the surface charge of the envelope as a function of the anode voltage measured by the ESI method. It can be seen that the charge in the cathode region of the envelope depends comparatively weakly on the electrode switching circuit and therefore the formation of this charge is more strongly influenced by the electric fields acting in this region and to a lesser extent by the current flowing in the gap. In the anode region, conversely, the current flow regime has the strongest influence on the envelope charge. The polarity of the charge indicates that it is attributable to the formation of homocharge in the appropriate regions of the envelope, i.e., negative in the cathode and positive in the anode regions. The experimental results suggest that the ''external'' polarization of the glass in the electric field created by the tube electrodes plays an important role in forming the envelope charge.

These results are consistent with the experimental data and theoretical ideas[14] on the formation of homocharge when insulators are polarized in strong fields. However, the polarization of the envelope of an electric vacuum device has important differences compared with the polarization of ordinary electrets. The main difference is that a negative charge appears in the anode region of the envelope when the tubes are operating without discharges.

A comparison of Figs. 3 and 7 reveals differences (approximately two orders of magnitude) in the magnitudes of the bulk and surface charges. However, this does not imply that the large difference between them is real, mainly because of the differences in the measurement conditions. For instance, the times between irradiation and the measurements obtained by these methods differ: the surface charge is measured 1–2 min after irradiation while the bulk charge is measured 1 h after irradiation. Under operating conditions with discharges when an appreciable positive charge appears at the surface, allowance should be made for a substantial reduction in the negative bulk charge formed by electron irradiation of the envelope, immediately after the high voltage has been switched off. This factor also explains why this charge is only found after operation at voltages above 100 kV under discharge-free conditions (Fig. 5).

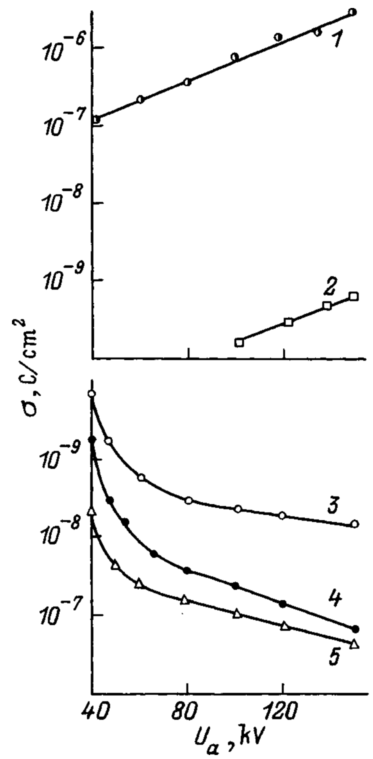


FIG. 7. Density of envelope surface charge as a function of anode voltage after exposure to discharges (*1*, *5*), without discharges (*3*), and without current (*2*, *4*) in the anode region (*1–3*) and in the cathode region (*4*, *5*).

CONCLUSIONS

To sum up, when a discharge-free operation of a vacuum device, the envelope potential in the anode region is slightly positive or close to a linear distribution law whereas the potential in the cathode region is close to the cathode potential. Under operating conditions when discharges appear in the interelectrode gap, in the regions adjacent to the anode and the cathode the envelope acquires potentials close to those of the corresponding electrodes. The potential distribution pattern over the envelope length is shown in Fig. 8.

The results show that the external polarization of the glass plays an important role in forming the envelope charge. Moreover, whereas the charge in the cathode region does not change significantly under various operating conditions of

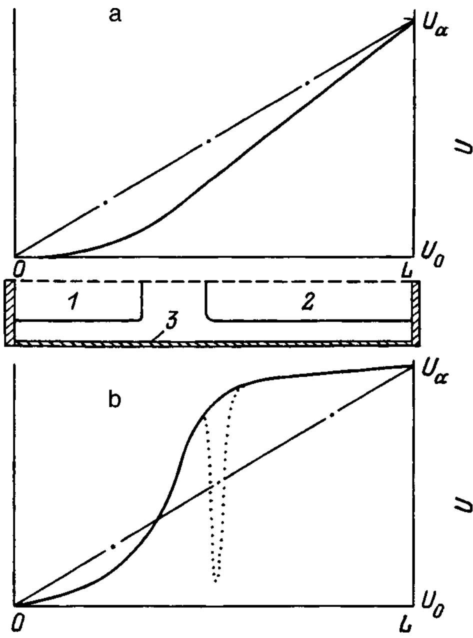


FIG. 8. Potential distribution over the length of the x-ray tube envelope without (a) and with discharges (b): the dashed line shows the change in the envelope potential at the site of bombardment by field emission electrons, *1* — cathode, *2* — anode, and *3* — insulating envelope.

the vacuum gap (homocharge is formed), homocharge only forms in the anode region when discharges or breakdowns occur in the vacuum interelectrode gap. The magnitude of this charge reaches 5×10^{-6} cm^2 and the envelope potential in the anode region becomes similar to the anode potential. Positive charge predominates in this region. As a result of its appearance, the operating conditions of the envelope in the anode region deteriorate abruptly: the intensity and energy of the bombarding electrons increase, which promotes the buildup of appreciable levels of electron space charge at the range depth (up to $\sim 60\,\mu$m at 150 keV). When the accumulated space charge reaches values of the order of 10^{-6} C, the strength of its field exceeds the dielectric strength of the glass, resulting in breakdown of the surface layers with the breakdown channel reaching the inner surface of the envelope. Thus, the breakdown of the envelope is closely related to the dielectric strength of the vacuum gap. During the operation of a device having a high dielectric strength where no discharges occur, the negative charge of the envelope near the anode impedes the bombardment of the glass, reducing the field emission current from the anode and lowering the energy of the electrons as they approach the surface of the glass.

[1] I. N. Slivkov, *Electric Insulation and Discharges in Vacuum* [in Russian], Atomizdat, Moscow (1972).
[2] N. V. Cherepnin, *Sorption Effects in Vacuum Technology* [in Russian], Sovetskoe Radio, Moscow (1973).
[3] V. N. Batygin, *Vacuum Ceramics* [in Russian], Énergiya, Moscow (1973).
[4] H. D. Edelson and B. G. Marks, IEEE Trans. Broadcast TV Rec. **BTR-17**(4), 264 (1971).
[5] V. I. Rachkov, *Electron X-Ray Tubes* [in Russian], GÉI, Moscow (1952).
[6] V. D. Bochkov, P. V. Poshekhonov, and M. M. Pogorel'skiĭ, Élektronnaya Tekhnika, Ser. 4. ÉV i GRP No. 5, 50 (1977).
[7] V. D. Bochkov, G. N. Petrov, M. M. Pogorel'skiĭ, and P. V. Poshekhonov, Élektronnaya Tekhnika, Ser. 4. ÉV i GRP No. 2(85), 26 (1981).
[8] V. D. Bochkov, Élektronnaya Tekhnika, Ser. 4 ÉV i GRP No.2(85), 11 (1981).
[9] V. D. Bochkov and M. M. Pogorelskiy, in *Proceedings of the 17th International Symposium on Discharges and Electrical Insulation in Vacuum*, Berkeley, CA, 1996, Vol. 1, pp. 548–551.
[10] V. D. Bochkov and M. M. Pogorelskiy, *ibid.*, pp. 544–547.
[11] L. B. Gordon, *Vacuum Insulation on the Moon, Proceedings of the 16th International Symposium on Discharges and Electrical Insulation in Vacuum*, St. Petersburg, 1994, pp. 459–462.
[12] A. R. Frederickson, IEEE Trans. Nucl. Sci. **43**, 426 (1996).
[13] B. Hilczer and J. Malecki, *Electrets* (Elsevier, Amsterdam, 1986).
[14] A. N. Gubkin, *Electrets* [in Russian], Nauka, Moscow (1978), 192 pp.
[15] G. M. Sessler, J. Appl. Phys. **43**, 408 (1972).
[16] G. Damamme and C. Le Gresus, in *Proceedings of the 16th International Symposium on Discharges and Electrical Insulation in Vacuum*, Berkeley, CA, 1996, pp. 405–409.
[17] J. P. Brainard, in *Proceedings of the Sixth International Symposium on Discharges and Electrical Insulation in Vacuum*, Swansea, 1974, pp. 190–194.
[18] C. H. De Tourreil and K. D. Srivastava, in *Proceedings of the Fifth International Symposium on Discharges and Electrical Insulation in Vacuum*, Poznan, Poland, 1972, pp. 295–299.
[19] V. D. Bochkov,*Vacuum and Gas Discharge Electronics*, [in Russian], Inter-University Collection, Ryazan' (1981), pp. 77–81.
[20] B. Gross, Phys. Rev. **107**, 368 (1957).
[21] V. D. Bochkov, *Vacuum and Gas Discharge Electronics*, [in Russian], Inter-University Collection, Ryazan' (1982), pp. 70–73.

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Hollow cathode glow discharge in long tubes

S. P. Nikulin

Institute of Electrophysics, Urals Branch of the Russian Academy of Sciences,
620219 Ekaterinburg, Russia

(Submitted March 9, 1998)
Zh. Tekh. Fiz. **69**, 36–39 (June 1999)

An analysis is made of some burning characteristics of a hollow-cathode glow discharge with a long tube ($L \gg D$) used as the cathode. It is shown that, as in the case $L \sim D$, the main factor imposing a lower limit on the range of operating voltages is the drift of fast electrons through the aperture in the cavity. Assuming that the electrons move along the cavity as a result of diffusion, it was possible to calculate the critical pressure at which the discharge can no longer burn and to determine the optimum ratio L/D for which the discharge can be sustained at the lowest voltage. The calculations showed satisfactory agreement with the experiment.
[S1063-7842(99)00606-6]

In Ref. 1 the present author analyzed the processes accompanying the burning of a glow discharge with electron oscillation in a hollow cathode and developed a model which agreed satisfactorily with the results of experiments using cathodes whose length L and diameter D are of the same order of magnitude. However, a comparison between the results predicted by this model and the experimental data obtained using long cathode tubes ($L \gg D$) reveals an appreciable discrepancy when the optimum hollow-cathode geometry is determined. The present author showed[1] that the minimum burning voltage is reached when $S_a/S_c \sim \sqrt{m/M}$, where S_a and S_c are the areas of the anode and the cathode, and m and M are the electron and ion masses, respectively. When this condition is satisfied, the fast electrons can produce a sufficient number of ionizations and at the same time, there are no problems with the transport of plasma electrons to the anode. This condition predicts that in a long tube ($S_c \approx \pi DL$) with an end anode ($S_a \approx \pi D^2/4$), optimum conditions for discharge burning in the most commonly used gases such as argon, nitrogen, and oxygen will be achieved when the condition $L/D \approx 0.25\sqrt{M/m} \sim 10^2$ is satisfied, although in experiments the minimum voltage or maximum discharge current (when a given voltage is maintained) were observed for $L/D \sim 10$ (Ref. 2). The aim of the present paper is to analyze the discharge burning characteristics in long tubes and to identify the reasons for this discrepancy.

Kirichenko *et al.*[3] reported experiments using tubes of different length, and on the basis of their results, concluded that the drift of fast electrons to the anode (these losses were taken into account in the author's model[1]) only impedes burning of the discharge for short lengths L. In long tubes the main factor limiting the discharge burning process is the drift of emitted electrons to the opposite part of the cathode and the discharge becomes quenched when $\lambda \approx S$, where λ is the electron mean free path. This conclusion is dubious since if we assume that electrons which have undergone collisions on their first transit through the cavity and have not reached the opposite part of the cathode subsequently expend all their energy in ionization, it is easily found that a self-sustained discharge can be achieved if the following condition is satisfied:

$$U = \frac{W}{e\gamma(1-\exp(-D/\lambda(\varepsilon)))}, \tag{1}$$

where U is the discharge voltage which is almost completely concentrated in the cathode sheath, W is the average energy dissipated in forming a pair of charged particles when a fast particle is stopped in the gas, e is the electron charge, γ is the coefficient of ion–electron emission, $\exp(-D/\lambda)$ are the fractions of particles which have reached the cathode without undergoing collisions and do not participate in ionization processes, and $\varepsilon = eU$ is the energy of the fast particles on their first transit. For any dependence $\lambda(\varepsilon)$ the voltage will gradually increase as the pressure P decreases and will only go to infinity when $\lambda \to \infty$, i.e., when $P \to 0$, but the discharge is abruptly quenched for nonzero P. In addition, Metel'[4] and Bersenev *et al.*[5] reported a discharge under conditions where $\lambda \gg D$, which also contradicts the conclusion reached in Ref. 3, but nevertheless an experiment was carried out to check this and the results are presented below.

A discharge was initiated in a long ($L/D = 10$) cathode tube *1* (Fig. 1) by a Penning discharge excited in a cell formed by end cathodes *2* and *3*, an anode *4*, and a magnet *5*. A probe *6* was placed opposite the exit aperture of the cavity. The probe was either floating or connected to the cathode. In the first case, the fast electrons escaping from the cavity were captured by the probe, while in the second case they were reflected by the probe and had a high probability of returning to the cavity since their radial drift was impeded by the magnetic field. Figure 2 gives the discharge burning voltage as a function of the gas supply for both cases. The measurements were made with the gas supply gradually reduced to the critical value for which the discharge was quenched in the cavity. If the main mechanism limiting the discharge burning process were electron losses to the walls of the cavity, the difference in the probe potential should not significantly influence the results. However, it can be seen that in the second case, the gas supply could be reduced by almost an order of

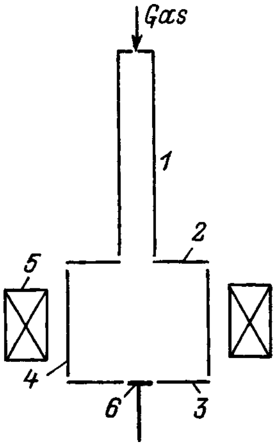


FIG. 1. Schematic of experiment.

magnitude compared with the first case. This result suggests that in long tubes the main mechanism impeding the discharge burning process at low pressures is the loss of fast particles through the cavity aperture.

Consequently, this discrepancy between theory and experiment cannot be attributed to the model neglecting the electron losses to the cavity walls. At the same time, the experimental results indicate that in long tubes, gas conditions are established under which the longitudinal motion of the fast particles is not a Knudsen process. This contradicts another assumption used in Ref. 1, that the electrons move freely over the entire volume of the cavity. We shall analyze the discharge conditions assuming that the transport of fast electrons along the cavity is a diffusion process and that the influence of the weak electric field in the plasma on the fast particle motion can be neglected. We shall write the equations of motion and continuity in the following form:

$$j_f=-D_f\frac{dn_f}{dx}, \tag{2}$$

$$\frac{dj_f}{dx}=\gamma\nu_i n_f-\frac{n_f}{\tau_r}, \tag{3}$$

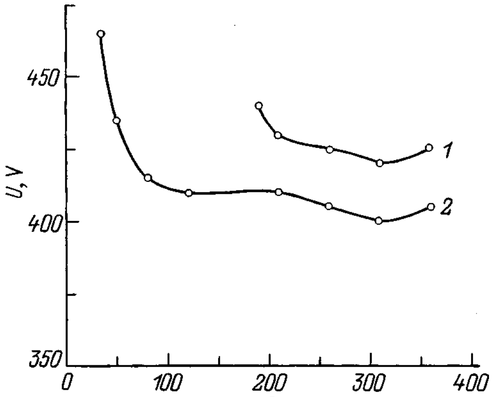


FIG. 2. Discharge burning voltage as a function of gas supply: *1* — probe at floating potential, *2* — probe at cathode potential.

where j_f and n_f are the flux density and concentration of fast particles, D_j and ν_i are the diffusion coefficient and the ionization frequency for the fast electrons, τ_r is the characteristic relaxation time of the fast electrons, for which we can write:

$$\tau_r=\frac{eU}{\nu_i W}. \tag{4}$$

Assuming that the square of the fast particle velocity varies between $2eU/m$ and 0 and has an average of $\sim eU/m$, we can write the following expression for D_f:

$$D_f=\frac{\langle v^2\rangle}{3\nu_f}=\frac{eU}{3m\nu_f}, \tag{5}$$

where ν_f is the effective fast-particle collision frequency.

The first term on the right-hand side of Eq. (3) describes the generation of new fast electrons by ion–electron emission. Here $\nu_i n_f$ is the density of the ions generated per unit time and it is assumed that the ions drift rapidly to the cathode in the radial direction without having time for any appreciable longitudinal displacement. The second term assumes that after the relaxation time has elapsed, a fast electron loses its capacity for ionization and is converted to a group of slow plasma electrons. Substituting Eq. (2) into Eq. (3), we obtain a second-order equation for n_f

$$\left(\gamma\nu_i-\frac{1}{\tau_r}\right)n_f+D_f\frac{d^2n_f}{dx^2}=0. \tag{6}$$

Assuming that near the fast-particle absorbing anode at the point $x=0$, the concentration of these particles is close to zero, we write the solution of the equation in the form

$$n_f=n_{f0}\sin\left(x\sqrt{\frac{\gamma\nu_i\tau_r-1}{D_f\tau_r}}\right). \tag{7}$$

The constant n_{f0} is related to the discharge current I_d and may be determined from the following equation:

$$I_d=(1+\gamma)e\pi\frac{D^2}{4}\int_0^L\nu_i n_f(x)dx. \tag{8}$$

At the opposite end of the tube, the end is at the cathode potential and electrons are reflected from it so that at the point $x=L$ the fast particle flux density, i.e., the derivative of the concentration, is close to zero. We then obtain the following equation:

$$L\sqrt{\frac{\gamma\nu_i\tau_r-1}{D_f\tau_r}}=\frac{\pi}{2}. \tag{9}$$

Note that when the second end electrode is at the anode or floating potential sometimes used, we need to impose the constraint that the derivative vanishes at the center of the system at the point $x=L/2$. Introducing the parameters

$$U_0=\frac{W}{e\gamma},\quad P_0=\frac{\pi}{\gamma L}\sqrt{\frac{W}{3m\nu_{f0}\nu_{i0}}}, \tag{10}$$

where ν_{f0} and ν_{i0} are the effective and ionization frequencies at 1 Torr pressure and the dimensionless variables are

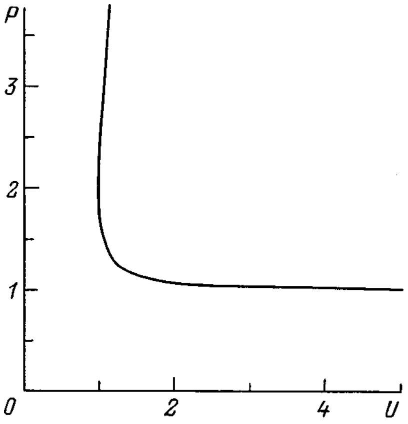


FIG. 3. Discharge burning voltage as a function of pressure.

$$u = U/U_0, \quad p = P/P_0, \tag{11}$$

we rewrite this equation in the following form:

$$2p\frac{\sqrt{u-1}}{u} = 1. \tag{12}$$

Figure 3 gives the discharge burning voltage as a function of pressure determined implicitly by this relationship. For $p > 1$ the dependence is two-valued. A similar result was obtained in Ref. 6 where the author analyzed the situation in a glow discharge with electron oscillation in a magnetic field and showed that states corresponding to the lower branch are stable and can be attained experimentally. As we move along this branch, the voltage drops rapidly with increasing pressure p and at the point $p = 1.74$ it is 1.1 so that conditions close to the optimum for maintaining a self-sustained discharge are achieved at this point (at the voltage $u = 1$ or $U = U_0$ where the fast electrons have time to expend all their energy before drifting to the anode). A further increase in pressure does not lead to any substantial increase in the number of ionizations accomplished by the fast electrons or any appreciable voltage drop.

For $p < 1$, i.e., when $P < P_0$ the discharge cannot be sustained at any voltage. Therefore the parameter P_0 is the critical parameter, below which the discharge cannot burn. Calculations of P_0 using the expression given above are of the same order of magnitude as the experimental results from Ref. 3. Thus, these results can be completely explained using concepts of electron losses through the cavity aperture, and the fact that the discharge was quenched at pressures consistent with the condition $\lambda \sim D$ is to some extent coincidental. In different cavity geometries the discharge can be quenched when $\lambda > D$ or $\lambda < D$.

Consequently the condition for a self-sustained discharge is satisfied for $P \leqslant P_0$. However, in order to ensure that a high-current discharge burns with a low discharge voltage, the constraint must also be imposed that at these pressures there are no problems with the transport of plasma electrons to the anode. A quasineutral state can be maintained over the entire length of the cavity and the electron current can be shorted at the anode in the absence of any negatively charged anode sheath if the time taken for the electrons to move along the plasma column is shorter than the time for radial ion drift to the cathode. We shall assume that the characteristic time for ion drift from the discharge is D/ν_b, where $\nu_b = \sqrt{kT_e/M}$ is the Bohm velocity, k is the Boltzmann constant, and T_e is the electron temperature. The conditions for sustaining a high-current discharge will be satisfied if a plasma electron can diffuse over a distance greater than the tube length within this time, i.e., if the condition

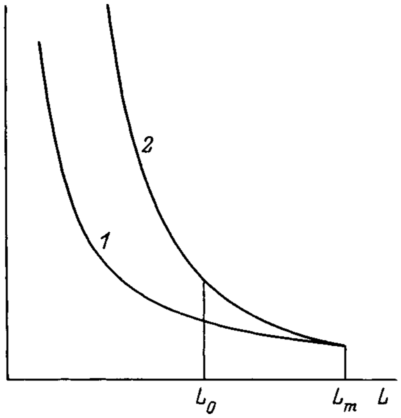


FIG. 4. Region in which a high-current discharge is achieved: *1* — $P_0(L)$, *2* — $P_1(L)$.

$$\sqrt{D_e\frac{D}{\nu_b}} > L, \tag{13}$$

is satisfied, where D_e is the diffusion coefficient of the plasma electrons.

By transforming this relationship, we can ascertain that a high-current discharge will be sustained if P is lower than a certain pressure P_1 given by:

$$P_1 = \sqrt{\frac{M}{m}}\sqrt{\frac{kT_e}{m}}\frac{D}{L^2\nu_{e0}}, \tag{14}$$

where ν_{e0} is the effective collision frequency for the plasma electrons at 1 Torr.

It should be noted that unlike fast particles, the transport of slow electrons along the cavity is strongly influenced by the electric field in the plasma. However, if we assume that the electrons are transported to the anode as a result of drift in the electric field rather than by diffusion, when the potential drop across the plasma is of the order of kT_e/e and the average electric field intensity is $\sim kT_e/eL$, we obtain exactly the same expression for P_1. On the whole, the electric field distribution pattern and the particle concentration in the plasma column of the hollow cathode are fairly complex and in some sections the electron transport is predominantly by diffusion while in others it is the result of drift.

The high-current discharge region is shown schematically in Fig. 4, this being the region above curve *1* and below curve *2*. Below curve *1* $(P = P_0(L))$ the discharge cannot burn in this form because the self-sustaining condition is not satisfied whereas above curve *2* $(P = P_1(L))$ this is not possible because of problems with shorting of the electron current at the anode. When the pressure increases above P_1, an electron sheath begins to form near the anode, across which the potential drop makes a significant contribution to the discharge voltage and the discharge goes over to a high-voltage form. If the anode is not positioned at the end of the cathode tube but some distance away, electrons may be transported in the region near the exit aperture as a result of the formation of a double layer. A requirement for its stable existence is that the external plasma should possess a sufficient ion-emission capacity.[1]

When the cathode length is increased to

$$L=L_m\approx\frac{\gamma D}{\pi}\sqrt{\frac{3kT_e}{W}}\sqrt{\frac{M}{m}}\,\frac{\sqrt{\nu_{f0}\nu_{i0}}}{\nu_{e0}} \tag{15}$$

the values of P_0 and P_1 become equal and the lowest working pressure P_m is achieved at this point, which for this particular configuration of gas-discharge system can be obtained using a cathode tube of diameter D

$$P_m\approx\frac{\pi}{\gamma L_m}\sqrt{\frac{W}{3m\nu_{f0}\nu_{i0}}}\approx\frac{\pi^2 W\nu_{e0}}{3\gamma^2 D\nu_{f0}\nu_{i0}\sqrt{3kT_eM}}. \tag{16}$$

For $L>L_m$ a high-current discharge cannot be sustained at any pressure.

Conditions close to the optimum for discharge burning at low pressures are evidently achieved for $L_0=L_m/1.74$ at the point $P=P_1(L_0)+1.74P_0(L_0)$ where, as was discussed above, the self-sustaining condition is satisfied at a low voltage close to U_0 and at the same time no anode electron sheath appears. Also bearing in mind that all the frequencies in Eq. (15) are of the same order and neglecting factors of the order of unity, we can derive the following expression for the optimum ratio L/D:

$$\left(\frac{L}{D}\right)_0\approx\gamma\sqrt{\frac{kT_e}{W}}\sqrt{\frac{M}{m}}\sim 10, \tag{17}$$

which agrees with the results of Ref. 2.

Broadly speaking, it should nevertheless be acknowledged that a discharge configuration in the form of a long cathode tube with an end anode even with $L/D\sim 10$ is not ideal from the point of view of achieving a high-current discharge at the lowest possible pressure, which is the situation usually most desired when hollow-cathode systems are used in charged particle sources. Substantially lower working pressures may be achieved by covering the exit aperture of the cathode tube with a diaphragm also at the cathode potential, as is convincingly evidenced by the results of Ref. 4. In this case, the area of the exit aperture in the diaphragm should obviously not be less than $\sim S_c\sqrt{m/M}$.

CONCLUSIONS

1. The main mechanism imposing a lower limit on the range of working pressures of a hollow-cathode glow discharge in the form of a long tube $(L\gg D)$, as in the case $L\sim D$, is the drift of fast electrons through the cavity aperture.

2. By allowing for electron losses through the cavity aperture and also taking into account that the electron motion along the cavity is not Knudsen motion, it is possible to calculate critical pressures and determine the optimum ratio L/D in satisfactory agreement with the experiment.

[1] S. P. Nikulin, Zh. Tekh. Fiz **67**(5), 43 (1997) [Tech. Phys. **42**, 495 (1997)].

[2] Yu. E. Kreĭndel', *Plasma Electron Sources* [in Russian], Atomizdat, Moscow (1977), 144 pp.

[3] V. I. Kirichenko, V. M. Tkachenko, and V. B. Tyutyunnik, Zh. Tekh. Fiz **46**, 1857 (1976) [Sov. Phys. Tech. Phys. **21**, 1080 (1976)].

[4] A. S. Metel', Zh. Tekh. Fiz **54**, 241 (1984) [Sov. Phys. Tech. Phys. **29**, 141 (1984)].

[5] V. V. Bersenev, N. V. Gavrilov, and S. P. Nikulin, in *Proceedings of the Conference ''Low-Temperature Plasma Physics''* Petrozavodsk, 1995, Part 2 [in Russian], pp. 251–253.

[6] S. P. Nikulin, Zh. Tekh. Fiz **68**(7), 56 (1998) [Tech. Phys. **43**, 795 (1998)].

Translated by R. M. Durham

Characteristics of the backscattering of moderate-energy electrons by solids having different atomic numbers

Yu. D. Kornyushkin

St. Petersburg State Institute of Precision Mechanics and Optics (Technical University), 197101 St. Petersburg, Russia

(Submitted January 5, 1997; resubmitted June 16, 1998)
Zh. Tekh. Fiz. **69**, 40–45 (June 1999)

The coefficient of backscattering of primary electrons is discussed as a function of their energy and atomic number *Z*. The amplitude of the wave function obtained in the first Born approximation and the Thomas–Fermi atom model are used to calculate the constant for screening of the electric field of the nucleus by atomic electrons. The theoretically calculated integral backscattering coefficients of primary electrons are compared with the experimental values in the range $12<Z<92$ for primary electron energies of 10 and 40 keV. Possible applications of these results are indicated. © *1999 American Institute of Physics.* [S1063-7842(99)00706-0]

INTRODUCTION

Studies of the laws governing the backscattering of primary electron fluxes are of particular interest for electron lithography[1] since the electron backscattering process primarily determines the resolution of the mask.[2] Electron backscattering laws are also used in electron scanning microscopy[3] to achieve the best contrast in the observed image of the surface of an object, and in local x-ray spectral analysis to calculate corrections in quantitative calculations.[4,5] Finally, studies of these relationships are important for problems relating to the first wall of a fusion reactor, since interaction between the hot plasma electrons and the wall produces a flux of backscattered electrons which have undergone appreciable energy losses, and this then leads to significant cooling of the entire plasma volume.[6,7] In the author's view, this aspect has not received sufficient attention. This effect can only be reduced by means of a suitable choice of first wall material and also by using a suitably selected surface profile and treatment characteristics. In the present paper an analysis is made of the backscattering of electrons having energies in the kiloelectronvolt range because electron fluxes at these energy levels are now the most widely used. The theoretical principles of the backscattering effect are fairly complex. Several approaches have been noted in the literature in which the backscattering process has been considered in some approximation.[8,9] However, no rigorous theory has yet been proposed to solve this problem. Here a possible variant is proposed to obtain an approximate solution of this problem.

THEORY

The penetration of a primary electron flux into a solid destroys the equilibrium distribution of free electrons. Consequently, this process can be considered on the basis of kinetic transport theory. The most acceptable approach to solving the problem would appear to be the following. A kinetic transport equation is constructed to describe the interaction between a flux of accelerated electrons and a solid within a microvolume dV at an arbitrary depth x taking into account the particle balance and energy. This equation only contains physical quantities averaged over a particular depth, such as the velocity, interaction cross section, and so on. This approach has proved quite justified in many cases.[10] The main difficulty associated with this approach involves determining the average energies and also the interaction parameters required for the calculations. This problem was solved by the present author in earlier studies.[11,12]

During the motion of primary electrons in a solid, each electron undergoes elastic and inelastic interaction for which the probability can be estimated using the mean free paths for elastic and inelastic scattering and absorption. If we introduce the average macroscopic interaction cross sections, the macroscopic cross section for the entire interaction will be

$$w=w_s+w_\gamma, \tag{1}$$

where w_s, w_γ are the average macroscopic cross sections for elastic and inelastic scattering and absorption, $w_s=\lambda_s^{-1}$, $w_\gamma=\lambda_\gamma^{-1}$, $w=\lambda^{-1}$, λ_s is the mean free path for elastic and inelastic scattering, λ_γ is the mean free path for absorption, and λ is the total mean free path.

We shall assume that during the interaction of primary particles with a solid the probability of new particles such as electrons being formed in the flux as a result of the ionization of atoms is fairly low compared with the probabilities of other processes. Jablonski[13] showed that the elastic mean free path of primary electrons in a solid is far shorter than the inelastic mean free path, i.e., $\lambda_e \ll \lambda_i$. Thus, we are sufficiently justified in considering that elastic scattering predominates in the collision process and we shall subsequently take w_s to be the macroscopic cross section for elastic scattering. It follows from Ref. 14 that the kinetic transport equation written in planar geometry allowing for energy losses and anisotropy of the primary particle scattering is given by:

$$-wf(x,\mu)+w_s\int_{-1}^{1}W_s(\mu\rightarrow\mu')f(x,\mu')d\mu'=\mu\frac{\partial f(x,\mu)}{\partial x}, \tag{2}$$

where

$$f(x,\mu)=\frac{dn}{v\,d\mu}, \tag{3}$$

$f(x,\mu)$ is the velocity distribution function of the primary electrons, *n* is the concentration of primary electrons at depth *x* in the range of angles between Θ and $\Theta+d\Theta$, v is their velocity, $\mu=\cos\Theta$, $\nu'=\cos\Theta'$; Θ and Θ' are the angles of incidence of primary electrons on the elementary volume *dV* and the angle of scattering from it, $W_s(\mu\rightarrow\mu')=dp_s/dv$ is the scattering indicatrix, and p_s is the elastic scattering probability.

Thus, the physical meaning of the scattering indicatrix is the probability that a primary electron will be scattered from the velocity range between $\mathbf{v}$ and $\mathbf{v}+d\mathbf{v}$ into the range between $\mathbf{v}'$ and $\mathbf{v}'+d\mathbf{v}'$, i.e., will be scattered taking into account a change in the direction of motion.

Since the time of arrival of a primary electron in the elementary volume *dV*, for which the kinetic transport equation (2) is constructed, and its time of departure from it are incompatible events, i.e., they always differ in time, in accordance with the ergodic theorem, the scattering indicatrix may be represented as a sum of two indicatrices, one depending only on the angle of incidence Θ and the other depending only on the scattering angles Θ',

$$W_s(\mu\rightarrow\mu')=\frac{1}{v}(1+W_s(\mu)+W_s(\mu')). \tag{4}$$

Equation (2) was solved by a method of separating the variables. As a result, after the necessary transformations we obtain the following particular solutions:

$$f_1(x,\mu)=C_1\frac{e^{ax}G(\mu)}{w/a+\mu},\quad f_2(x,\mu)=C_2\frac{e^{-ax}G(\mu)}{w/a+\mu},$$

$$f_3(x,\mu)=C_3\frac{e^{ax}G(\mu)}{w/a-\mu},\quad f_4(x,\mu)=C_4\frac{e^{-ax}G(\mu)}{w/a-\mu}, \tag{5}$$

where C_1, C_2, C_3, and C_4 are the integration constants, *a* is the interaction parameter, $a=R_p^{-1}$, R_p is the average longitudinal range of a primary electron in the material;

$$G(\mu)=\frac{w_s}{2}\{[1+W_s(\mu)]A_0+A_1\},$$

A_0, and A_1 are constants.

If the distribution function $f(x,\mu)$ is real, smooth, and finite, these particular solutions are unique, as was shown by the author in Ref. 14

The physical formulation of the problem indicates that two primary electron fluxes may be identified in a solid, propagating simultaneously in opposite directions. Hence, the resulting distribution of the primary electrons in the solid at depth *x* may be represented in the following form: for the forward-propagating flux,

$$f_\eta(x,\mu)=\left(C_1\frac{e^{ax}}{w/a+\mu}+C_4\frac{e^{-ax}}{w/a-\mu}\right)G(\mu), \tag{6}$$

and for the backward-propagating flux

$$f_r(x,\mu)=(C_3e^{ax}+C_4e^{-ax})\frac{G(\mu)}{w/a-\mu}. \tag{7}$$

Then, taking account of Eq. (3), the flux density of primary electrons in the range of angles between Θ and $\Theta+d\Theta$ at depth *x* will be

$$dj=v\,dn=v^2f(x,\mu)d\mu.$$

In the microvolume *dV*, the integral flux density of primary electrons propagating in the direction of the *X* coordinate at depth *x* will be

$$j(x)=2\pi v^2\int_{-1}^{1}\mu f(x,\mu')d\mu'. \tag{8}$$

After substituting the distribution function $f_\eta(x,\mu)$ from Eq. (6) into Eq. (8) and integrating, we obtain an expression for the integral flux density of the primary electrons propagating in the direction of increasing *x*,

$$d_\eta(x)=v^2(C_1B_{11}e^{ax}+C_4B_{12}e^{-ax}), \tag{9}$$

where

$$B_{11}=\int_0^\pi\frac{G(\Theta)\sin 2\Theta\,d\Theta}{w/a+\cos\Theta},\quad B_{12}=\int_0^\pi\frac{G(\Theta)\sin 2\Theta\,d\Theta}{w/a-\cos\Theta}.$$

However, since the function $G(\Theta)$ is even for normal incidence, we find $B_{11}=-B_{12}$. Taking account of Eq. (7), the integral flux density in the opposite direction at depth *x* will be

$$j_r(x)=v^2B_{21}(C_3e^{ax}+C_4e^{-ax}), \tag{10}$$

where

$$B_{21}=\int_{\pi/2}^{\pi}\frac{G(\Theta)\sin 2\Theta\,d\Theta}{w/a-\cos\Theta}. \tag{11}$$

The following boundary conditions were used to determine the integration constants:

$$j_\eta(0)=j_0-j_r(0),\quad j'_\eta(x)|_{x\geqslant h}=0,\quad j_r(h)=0, \tag{12}$$

where *h* is the thickness of the free layer.

For a layer of infinitely large thickness $h=R_p$, j_0 is the integral flux density of primary particles incident on the surface of a free layer of thickness *h*.

After substituting the values from Eqs. (9) and (10) into Eq. (12), we obtain a system of three algebraic equations which, when solved jointly, can determine the integration constants C_1, C_3, C_4:

$$C_1=\frac{(1-r)e^{-ah}j_0}{2v^2B_{11}\cosh(ah)},$$

TABLE I. Values of the coefficients, used in Eqs. (18) and (19).

E_p, keV	B_0	g	b_0	b_1	b_2
10	77.50	-0.843	-13.27	0.4476	-3.20×10^{-3}
40	83.23	-0.832	-12.63	0.4395	-3.10×10^{-3}

$$C_3=\frac{(r-1)e^{-ah}j_0}{2v^2B_{12}\cosh(ah)},\quad C_4=\frac{(1-r)e^{ah}j_0}{2v^2B_{12}\cosh(ah)},\tag{13}$$

where $r=j_r(0)/j_0$ and r is the integral backscattering coefficient.

After substituting the values of the integration constants C_3 and C_4 from Eq. (13) into Eq. (10) for $x=0$, we obtain[15]

$$2h\leqslant R_p,\quad r=\left(1+\frac{B_{12}}{B_{21}}\coth(2ah)\right)^{-1},$$

$$2h\leqslant R_p,\quad r=\left(1+\frac{B_{12}}{B_{21}}\coth(1)\right)^{-1}.\tag{14}$$

In order to calculate the integral backscattering coefficient r, we need to know the value of $a=R_p^{-1}$, h and the ratio B_{12}/B_{21}. The value of a can be calculated from[15]

$$a=a_0(\rho Z)^{0.69}E_p^{-\frac{2}{n}}F\left(\frac{b_m}{n}\right),\tag{15}$$

where $n=n_0+cZ$, $a_0=2.24\times10^{-4}$, $n_0=1.287$, $c=-2.05\times10^{-3}$, ρ (g/cm^3) is the density of the target material, and $[E_p]=$ keV.

Since the coefficients B_{12} and B_{21} have a fairly complex analytic expression, and their ratio is even more complex, in order to simplify the following calculations the ratio of these coefficients is replaced by the fairly simple analytic function

$$\frac{B_{12}}{B_{21}}=\frac{B_1}{B_2}B(Z),\tag{16}$$

where

$$B_1=\frac{w}{a}\ln\frac{w/a+1}{w/a-1}-2,\quad B_2=\frac{w}{a}\ln\frac{w/a}{w/a-1}-1.\tag{17}$$

An analysis shows that in the energy range 10–40 keV the function $B(Z)$ may be expressed as follows:

$$12<Z<50\quad B(Z)=B_0Z^g,\tag{18}$$

$$56<Z<90\quad B(Z)=b_0+b_1Z+b_2Z^2.\tag{19}$$

The values of the coefficients B_0, g, b_0, b_1, and b_2 are given in Table I. It can be shown[16] that the average longitudinal range R_p is proportional to the total range R whose value can be determined using Bethe theory.[17] It follows from this theory that the total electron range in a material is determined by the Bethe stopping law which only holds for electron energies $E\gg J$ where J is the average ionization potential of an atom.

The present author showed[15] that the Bethe law of continuous energy losses can be applied to obtain the following expression for the total macroscopic interaction cross section:

$$w=w_0\rho A^{-1}(Z-b)E_p^{-\frac{2}{n}}F\left(\frac{b_m}{n}\right),\tag{20}$$

where $w_0=6.3\times10^{-4}$, b is the screening constant, $b_m=\langle\Delta E\rangle/E_p$ is the average relative energy loss of a primary electron after passing through a layer of thickness h whose value can be calculated using the results of Ref. 18, n is the exponent in the power potential of the interaction between an impinging primary electron and the atomic core, $F(b_m/n)$ is a function which takes into account the energy losses of primary electrons in a solid,

$$F\left(\frac{b_m}{n}\right)=\sum_{i=1}^{Z}A_i\left(\frac{b_m}{n}\right)^i;\tag{21}$$

$A_0=A_1=1$, $A_2=(2+n)/3$, $A_3=(1+n)/2A_2$, $A_4=(2+3n)/5A_3$, $A_5=(1+2n)/3A_4$, $A_6=(2+5n)/42A_5$, and $A_7=(1+3n)/60A_6$.

Here the screening constant b was calculated theoretically using the Thomas–Fermi model taking into account the binding energy of atomic electrons in various electron shells of the atom whose values were given in Ref. 19. Then using Eqs. (15) and (20) we have for a semiinfinite layer

$$\frac{w}{a}=2.81\rho^{0.31}A^{-1}Z^{-0.69}(Z-b)^2.\tag{22}$$

Expressions (20) and (22) contain the screening constant b which has the following physical meaning. As a fast primary electron moves in a solid, it undergoes continuous collisions with atoms of the material, which are accompanied by elastic and inelastic scattering. If small-angle deflections predominate, this means that the scattering is mainly elastic where a primary electron penetrates fairly deeply into the electron shell of the atom. As it moves inside the atom, the primary electron is mainly scattered in the electric field of the nucleus and the atomic electrons bound strongly to it. These strongly bound electrons together with the nucleus create a resulting electric field under whose action the primary electron moves within the atom. Thus, the screening constant b in formulas (20) and (22) is taken to be the number of atomic electrons involved in screening the electric field of the atomic nucleus. It can therefore be postulated that in a solid the scattering of fast electrons can be considered as scattering by a static force center, in whose electric field the primary electron is located. The potential energy of the primary electron in the resultant electric field of the nucleus and the atomic electrons screening the nucleus will then be

$$U(r)=\frac{Z^*e^2}{4\pi\varepsilon_0r}e^{-\frac{r}{r_0}},\tag{23}$$

where $Z^*=Z-b$ and r_0 is the screening radius.

It was shown[15] that the total microscopic cross section for elastic scattering of a primary electron at the screened field of an atomic nucleus is given by

$$\sigma=\left(\frac{Z^*e^2}{4\pi\varepsilon_0mv^2}\right)^2\frac{\pi}{\alpha^2(1+\alpha^2)},\tag{24}$$

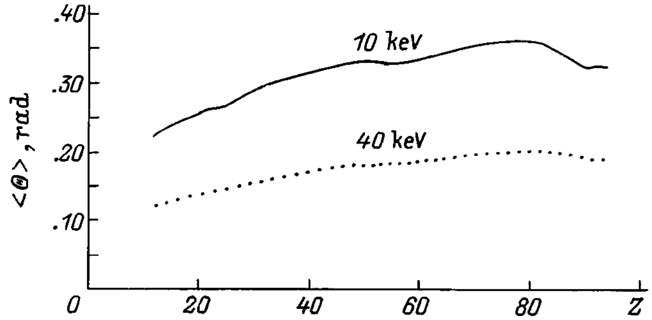


FIG. 1. Average angle of scattering of a primary electron by target atoms as a function of the atomic number of the element.

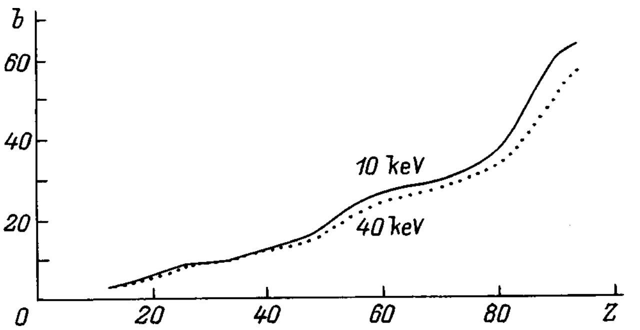


FIG. 2. Screening constant b as a function of atomic number of the element.

$$\alpha^2=7.46\times10^{-6}(Z^*)^{2/3}\left(\frac{c}{v}\right)^2, \qquad (25)$$

where c is the velocity of light in vacuum.

Since the scattering of a primary electron by individual atoms in a solid is a statistical process, we can use Eq. (24) to calculate the average angle for scattering of a primary electron in a solid in the small-angle approximation

$$\langle\Theta\rangle=\frac{4\alpha^2(1+\alpha^2)}{\pi}\int_0^{\pi}\frac{\Theta\sin\Theta\, d\Theta}{\left(\sin^2\frac{\Theta}{2}+\alpha^2\right)^2}. \qquad (26)$$

It follows from Eq. (25) that α^2 depends on the screening constant. In turn, b depends on the primary electron energy and the structural characteristics of the solid atoms with which the primary electrons interact. We know that as a result of a collision between a primary electron and an atom, the momentum of the primary electron changes by

$$\hbar q=2\hbar k\sin\frac{\Theta}{2}, \qquad (27)$$

where k is the wave number of the primary electron.

Consequently, the atom at which the primary electron is scattered receives the same momentum. This means that the elastic scattering condition for which the primary electron momentum remains the same and only the direction of motion changes, should correspond to the momentum (27) imparted to the atom, which we can then write in terms of the energy

$$\Delta E=4E_p\sin^2\frac{\Theta}{2}. \qquad (28)$$

Then, from the physical meaning of the formulation of the problem, the condition for elastic scattering at strongly bound atomic electrons should correspond to the inequality $\Delta E\leqslant E_i$, where E_i is the binding energy of the atomic electrons completely involved in screening the electric field of the nucleus. If $\Delta E>E_i$, atomic electrons with this binding energy will only participate partially in the screening. Thus, the concept of an effective number of electrons involved in the screening should be introduced for these electrons. This applies to the spherically symmetric electron shell of an atom. If the appropriate electron shell is not spherically symmetric, allowance should also be made for the departure from spherical symmetry by introducing a suitable coefficient, as in Ref. 15.

The angle $\langle\Theta\rangle$ was calculated by the method of successive approximations of the integral (26) until $\langle\Theta\rangle$ reached a limiting value. The value of b for which $\langle\Theta\rangle$ reaches a maximum was taken as the screening constant in this study.

Figures 1 and 2 give calculated values of $\langle\Theta\rangle$ and b for $E_p=10$ and 40 keV. These screening constants were then used to calculate w/a using formula (22). Then, taking account of Eqs. (15), (17), and (18) we can calculate the integral backscattering coefficient from Eq. (14) allowing for the macroscopic anisotropy of the properties of the material for various primary electron energies and various atomic numbers. For a free layer of considerable thickness the primary electrons penetrate to depths not exceeding the average longitudinal range R_p. Since during backscattering primary electrons initially move forward and then backward, their effective penetration depth is $R_p/2$, but the total longitudinal range remains the same as before, R_p. Therefore $2h$ in Eq. (14) should be taken to be the total longitudinal range R_p. Moreover, since the physical meaning of the interaction parameter is $a=R_p^{-1}$, in Eq. (14) the product is $2ah=1$.

Figure 3 gives the values of w/a calculated using formulas (14) and (22) and Figs. 4 and 5 give the theoretical and experimental values of r (Refs. 20 and 21). A comparison reveals good agreement over a wide range of Z. In addition, the theory predicts that the integral backscattering coefficient r should not vary monotonically as a function of the atomic number Z, as follows from Ref. 20, but some quasiperiodicity should be observed. This is caused by the gradual filling of the electron shells with atoms as Z increases. Some dis-

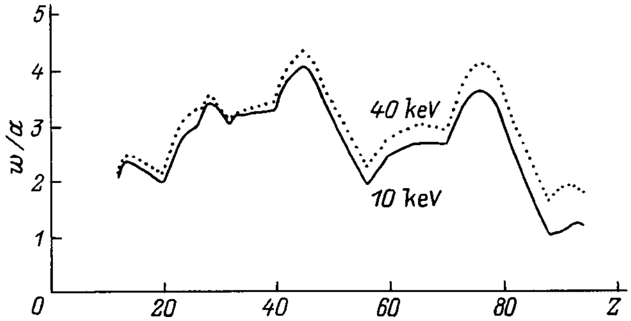


FIG. 3. Dependence of the ratio w/a on the atomic number of the element.

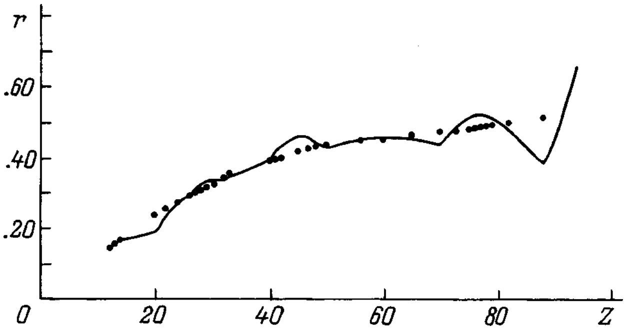


FIG. 4. Theoretical (solid curve) and experimental values[20,21] of integral backscattering coefficients for normal incidence of an accelerated electron flux on the surface of a target ($E_p = 10$ keV).

agreement between theory and experiment is evidently caused by the insufficiently accurate mapping of the experimental results with the empirical formulas used for a comparison between theory and experiment.[20]

The exponent n can be calculated from formula (15). It can be seen that as Z increases, n decreases, tending to unity. This indicates that the elastic scattering of primary electrons by heavy atoms is close to a Coulomb process whereas the scattering by light atoms may differ appreciably from this. This behavior is mainly attributable to some characteristic features of the energy loss law: scattering by light atoms is accompanied by higher energy losses than scattering by heavy atoms.

CONCLUSIONS

The relationships established can be used to calculate the integral backscattering coefficient for any solids in the energy range 10–100 and also to predict possible values of r for those materials not naturally observed in the solid state and also for those whose emission properties have not yet been sufficiently well-studied (actinides). All these factors indicate that the results are of major practical value.

The values of w/a plotted in Fig. 3 using results from Refs. 22 and 23 can be used to predict qualitatively which solids have a narrow angular distribution of backscattered primary electrons and which have a broad angular distribution. At this point, we can predict that all materials with a low ratio w/a have an appreciable angular half-width. Thus, we can affirm that the half-width of the angular distribution of the backscattered primary electron flux from a solid of infinitely large thickness, such as Mg, Al, or Si, is approximately the same and small whereas that for Pt, Au, and Pb should be appreciable. Therefore, light materials and coatings should be used to obtain the best resolution for masks used in electron lithography and to obtain images with the highest contrast in scanning electron microscopy, and in other similar cases.

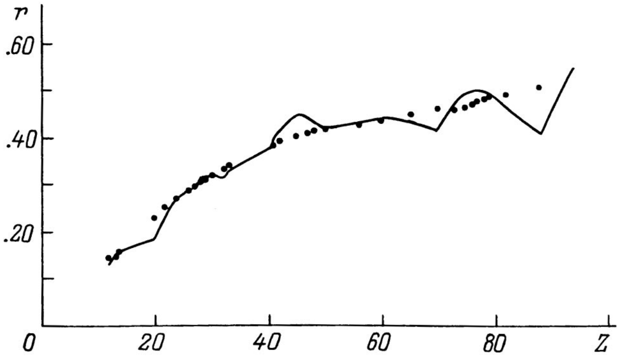


FIG. 5. As Fig. 4 for $E_p = 40$ keV.

This approach can also be used to analyze the emission properties of various chemical compounds in the solid state. It can also be applied to an obliquely incident primary electron flux on a free layer of solid of finite thickness and also on multilayer coatings.[24]

This work was supported financially by ''IKA'' (31–33 Dnepropetrovskaya Ul., 191119, St. Petersburg) for whose guidance the author is deeply grateful.

[1] K. A. Valiev and A. V. Rakov, *Physical Principles of Submicron Lithography in Microelectronics* [in Russian], Moscow (1984), 352 pp.
[2] I. Broudaĭ and D. Mereĭ, *Physical Principles of Microtechnology* [in Russian], Moscow (1985), 496 pp.
[3] *Practical Scanning Electron Microscopy: Electron and Ion Microprobe Analysis*, edited by J. I. Goldstein and H. Yakowitz [Plenum Press, New York (1975); Mir, Moscow (1978), 275 pp.].
[4] *Physical Principles of Local X-Ray Spectral Analysis*, edited by I. B. Borovskiĭ [in Russian], Moscow (1973), 312 pp.
[5] M. Cox, G. Love, and V. Scott, J. Phys. D **12**, 1441 (1978).
[6] R. K. Janev, M. F. A. Harrison, and H. W. Drawin, Nucl. Fusion **29**, 109 (1989).
[7] M. E. Woods, B. J. Hopkins, G. F. Matthews *et al.*, J. Phys. D **20**, 1136 (1987).
[8] H. J. August and J. Wernisch, J. Microsc. **157**, 247 (1990).
[9] W. Werner, I. Tilinin, and M. Hayek, Phys. Rev. B **50**, 4819 (1994).
[10] Yu. D. Kornyashkin, Thin Solid Films **190**, 373 (1990).
[11] Yu. D. Kornyushkin, Fiz. Tverd. Tela (Leningrad) **20**, 1175 (1978) [Sov. Phys. Solid State **20**, 676 (1978)].
[12] Yu. D. Kornyushkin, Izv. Akad. Nauk SSSR. Ser. Fiz. No. 12, 2288 (1982).
[13] A. Jablonski, Surf. Interface Anal. **6**(6), 291 (1984).
[14] Yu. D. Kornyushkin, Poverkhnost' No. 12, 13 (1992).
[15] Yu. D. Kornyushkin, Poverkhnost' No. 6, 24 (1993).
[16] A. P. Komar, S. P. Kruglov, and I. V. Lopatin, *Measurement of the Total Energy of Bremsstrahlung Beams from Electron Accelerators* [in Russian], Leningrad (1972), 236 pp.
[17] L. D. Spenser, Phys. Rev. **98**, 1597 (1955).
[18] T. Tabata, R. Ito, and S. Okabe, Jpn. J. Appl. Phys. **11**, 1220 (1972).
[19] A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* [Springer-Verlag, Berlin (1985); Énergoatomizdat, Moscow (1986), 344 pp.].
[20] A. Ya. Vyatskin, A. N. Kabanov, and V. V. Trunev, Radiotekh. Elektron. **17**, 1983 (1972).
[21] A. Ya. Vyatskin, A. N. Kabanov, B. N. Smirnov, and V. V. Trunev, Radiotekh. Elektron. **21**, 895 (1976).
[22] Yu. D. Kornyushkin, Fiz. Tverd. Tela (Leningrad) **21**, 627 (1979) [Sov. Phys. Solid State **21**, 372 (1979)].
[23] Yu. D. Kornyushkin, Fiz. Tverd. Tela (Leningrad) **23**, 2842 (1981) [Sov. Phys. Solid State **23**, 1660 (1981)].
[24] Ju. D. Kornyashkin, Phys. Status Solidi A **95**, K81, 247 (1986).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Distribution of 90° domain reorientations in lead titanate zirconate piezoceramic under longitudinal compression

M. G. Minchina and O. I. Yankovskiĭ

Rostov State University, 344090 Rostov-on-Don, Russia

(Submitted September 24, 1997; resubmitted March 18, 1998)
Zh. Tekh. Fiz. **69**, 46–49 (June 1999)

A static model of a critically polarized ceramic is used to analyze the behavior of the piezocoefficient d_{33} of lead titanate zirconate ceramic exposed to a longitudinal compressive stress σ and a quantitative analysis is made of the 90° and 180° domain reorientations. It is shown that for TsTG-83G ceramic a 180° antiparallel domain structure forms at compressive stresses $\sigma > 10^8\,\mathrm{N/m}^2$. © *1999 American Institute of Physics.* [S1063-7842(99)00806-5]

INTRODUCTION

It is known that the piezocoefficient d_{33} of a lead zirconate titanate (PZT) piezoceramic plate is reduced as a result of 90° domain reorientations which take place under the action of a longitudinal mechanical stress σ (Refs. 1 and 2). As the stress σ in lead titanate crystals increases, domains with the **c** polar axis perpendicular to the direction of compression increase in volume, annihilating domains with the **c** polar axis parallel to this direction as a result of the domain walls being displaced along the normal to their intrinsic planes. When the stress is removed, the twinning structure is partially restored to the initial state.[3,4] Detwinning of lead titanate crystals showed that 90° reorientation of the **c** polar axes is accompanied by 180° polarization reversal processes which occur when a 90° domain wall is displaced rapidly ($V_d > 1\times10^{-5}$ m/s) (Ref. 5). An antiparallel 180° domain structure also forms in polarized PZT ceramic exposed to compressive stresses $\sigma > 90\times10^6\,\mathrm{N/m}^2$ (Ref. 6). In lead titanate crystals, the domains only become reoriented when the mechanical stress σ exceeds the coercive stress σ_c of the domain.[7,8] In Ref. 9 the domain distribution over coercive stresses σ_c in the ceramic was expressed using the distribution function $f(\sigma)$ which is equal to the probability density of the **c** polar axis of the domain being rotated under the longitudinal compression σ. Here we propose to analyze the coercive stress σ_c averaged over all domains, assuming that the **c** polar axes of the domains in TsTS-83G ceramic are distributed nonuniformly under the action of a longitudinal stress σ.

The aim of the present study is to analyze the behavior of the piezocoefficient $d_{33}=d_{33}(\sigma)$ and the distribution of 90° domain reorientations in TsTS-83G ceramic exposed to a longitudinal compression σ using a static model for a critically polarized ceramic.

EXPERIMENTAL METHOD

Ferroceramic having the composition TsTS-83G was prepared by hot pressing. A block 100 mm in diameter and 12 mm thick was sintered at $T=1100\,^\circ\mathrm{C}$, held for 5 h, and loaded to 90 kg/cm^2. The block was polarized in silicone at $T=120°$ for 1 h at $E=25$ kV/cm. The direction of the polarizing field **E** coincided with the principal Z crystal-physics axis of the ceramic. A diamond disk was used to cut $4\times4\times4$ mm samples with the orientation of the principal XYZ crystal-physics coordinate system and the direction of the polarizing field **E** being maintained. Electrodes were deposited on the oriented polarized samples by cathode deposition from Al + Cr at $T=80\,^\circ\mathrm{C}$ for 30 min. Measurements of the piezocoefficients d_{33} and $d_{33}=d_{33}(\sigma)$ were made under quasistatic conditions using a ''Piezomodul''' device.

RESULTS AND DISCUSSION

We shall analyze a class $4mm$ piezoceramic plate in the principal XYZ crystal-physics coordinate system (the polarization vector **P** is directed along the Z axis). We shall apply a uniaxial mechanical stress

$$\begin{bmatrix} 0\,0\,0 \\ 0\,0\,0 \\ 0\,0\,\sigma \end{bmatrix}$$

to the plate as a result of which an electric charge will be induced on the hatched faces of the plate (Fig. 1), its magnitude being proportional to the piezocoefficient d_{33}.

The equation for the piezocoefficient d_{33} of a group $4mm$ ferroceramic on transition from the XYZ crystal-physics coordinate system to the $X'Y'Z'$ coordinate system using the laws for transformation of a third rank tensor has the form[9]

$$d_{33}=(d'_{15}+d'_{31})\cos\Theta\,\sin^2\Theta+d'_{33}\cos^3\Theta, \qquad (1)$$

where d_{33} is the piezocoefficient in the XYZ coordinate system, and d'_{33}, d'_{31}, and d'_{15} are the piezocoefficients in the $X'Y'Z'$ coordinate system.

Equation (1) can take into account the relationship between the piezocoefficient d_{33} of the polarized ceramic and the piezocoefficients d'_{33}, d'_{31}, and d'_{15} of the crystallites (assuming that these are single-domain). The experimentally determined piezocoefficients of TsTS-83G ceramic are $d_{33}=360$ pC/N, $d_{31}=-180.9$ pC/N, and $d_{15}=508.2$ pC/N. According to published sources, the piezocoefficients of lead titanate ceramic and single crystals differ negligibly[10,11] so

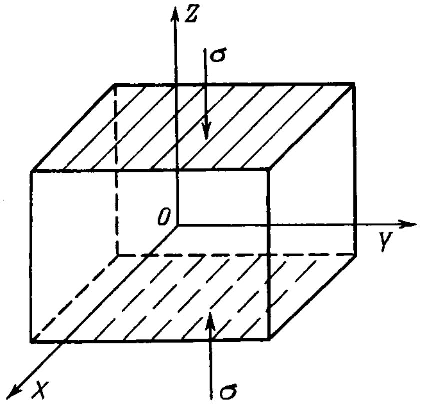


FIG. 1. Piezoceramic plate with longitudinal mechanical stress σ applied to hatched faces.

that the piezocoefficients of TsTS-83G ceramic were taken as the piezocoefficients d'_{33}, d'_{31}, and d'_{15} of the crystallite (domain).

A static model for a critically polarized ceramic was used to construct an equation for the piezocoefficient d_{33} of a ceramic exposed to longitudinal compression σ (Ref. 12). This model neglected interaction between the crystallites (domains) and the motion of the domain walls. The authors assumed that the internal mechanical stresses created in the ceramic during sintering remain constant under external influences and only the field applied to the ceramic acts on each crystallite.

The polarized ceramic was represented as a system of N single-domain crystallites polarized to saturation, whose **c** polar axes have the preferred orientation and are located in the upper part of the orientation sphere in the solid angle $\Theta=54°44'$. According to the model, the distribution of the **c** polar axes in zones of total and partial stability, taking into account all possible 180° and 90° reorientations, is inhomogeneous ($\rho_c^I=3N/2\pi$ and $\rho_c^{II}=3N/2\pi(1-(4\arccos(\coth\Theta))/\pi)$ are the densities of the **c** polar axes in zones of total and partial stability, respectively). The condition for total stability of the **c** polar axes in a ceramic to which an external stress σ_{33} is applied has the form (Fig. 2)

$$\cos\gamma_c-\cos\gamma_a\geqslant\sigma_c/\sigma, \tag{2}$$

where γ_c is the angle between the direction of the field **E** and the **c** axis of the domain, γ_a is the angle between the **a** axis closest to the field and the field **E**, σ_c is the coercive stress ($\sigma_c=16.555\times10^6$ N/m^2 for PbTiO$_3$ crystals[4]), and σ is the external mechanical stress.

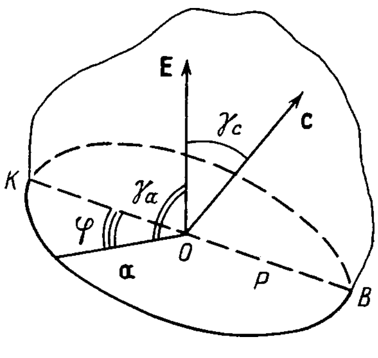


FIG. 2. Geometric relation between the angles γ_c, γ_a, and φ characterizing the position of the domains relative to the polarizing field **E**.

We shall use the geometric relation between the angles γ_c, γ_a, and φ characterizing the position of the domain relative to the polarizing field **E** (Ref. 13)

$$\cos\gamma_a=\sin\gamma_c\cos\varphi, \tag{3}$$

where φ is the angle between the **a** axis closest to the field **E** and the line formed by intersection of the plane passing through the **a** axis and the plane passing through the **c** axis and **E**.

We shall analyze the condition for total stability of the **c** polar axes (2) when the **a** axes of the domain (crystallite) are in the position closest to the polarizing field **E** ($\varphi=0°$),

$$\cos\gamma_c-\tan 45°\sin\gamma_c\geqslant\sigma_c/\sigma. \tag{4}$$

After simple transformations, Eq. (4) has the form

$$0\leqslant\gamma_c\leqslant(\arccos((\sigma_c/\sigma)\cdot(\sqrt{2}/2)))-45°. \tag{5}$$

Inequality (5) gives the angular boundaries of the zone of total stability of the **C** polar axes when the longitudinal stress σ is applied to the ceramic.

The condition for partial stability of the **c** polar axes when the **a** axes are further away from the direction of the polarizing field **E** ($\varphi=45°$), has the form

$$\cos\gamma_c-(\sqrt{2}/2)\sin\gamma_c\geqslant\sigma_c/\sigma. \tag{6}$$

Equation (6) can be reduced to the form

$$45°\leqslant\gamma_c\leqslant54.733°-\arcsin((\sigma_c/\sigma)\cdot(\sqrt{2/3})). \tag{7}$$

Inequality (7) gives the angular boundaries of the zone of partial stability of the **c** polar axes for a critically polarized ceramic.

In the zone of partial stability of the **c** polar axes we take into account that part of the **c** axes which after 90° switching, was converted into **a** axes and then the condition for stability of the **a** axes in this zone will have the form

$$\cos\gamma_a-\cos\gamma_c\geqslant\sigma_c/\sigma. \tag{8}$$

Substituting condition (3) into (8), we find the angle $|\varphi|$ containing the **c** axes close to the field **E**,

$$|\varphi|\leqslant\arccos(\coth\gamma_c+\sigma_c/(\sigma\sin\gamma_c)). \tag{9}$$

The fraction of crystallites β_c whose **c** axes ‘‘drift’’ out of the zone of partial stability of the **c** polar axes as a result of 90° switching will be

$$\beta_c=4|\varphi|/\pi=(4\arccos(\coth\gamma_c+\sigma_c/(\sigma\sin\gamma_c)))/\pi. \tag{10}$$

The fraction of crystallites β'_c whose **c** axes remain inside the zone of partial stability of the polar axes after switching will be

$$\beta'_c=1-(4\arccos(\coth\gamma_c+\sigma_c/(\sigma\sin\gamma_c)))/\pi. \tag{11}$$

The value of Δd_{33}^T by which the piezocoefficient d_{33} changed under the action of the longitudinal stress σ was calculated by integrating over the orientation sphere of the **c** polar axes in accordance with the angular boundaries of the

TABLE I. Experimental and theoretical values of the piezocoefficient $d_{33} = d_{33}(\sigma)$ and distribution of 90° and 180° domain reorientations in TsTS-83G ceramic.

$\sigma \times 10^{-6}$ N/m^2	d_{33}^{exp}, pC/N	d_{33}^{theor}, pC/N	$A^{90°}$, %	$A^{180°}$, %
24.5	360	36	–	–
42.9	340	338.8	17	–
61.3	300	298.9	37	–
91.9	200	198.5	88	–
128.7	140	139.2	100	23
165.5	100	99.5	100	53
202.3	84	83.7	100	62
220.7	68	67.5	100	75

zones of total and partial stability of the **c** axes and averaging the piezoelectric constants d'_{33}, d'_{31}, and d'_{15} over all the crystallites (domains) in the ceramic using the formula

$$\begin{aligned}\Delta d_{33} = {} & 1/N \int_{\varphi=0}^{2\pi} \int_{\Theta=0}^{\arccos((\sqrt{2}/2)\cdot(\sigma_c/\sigma))-45°} d_{33}(\Theta) \\ & \times (3N/2\pi) \sin\Theta \, d\Theta \, d\varphi \\ & + 1/N \int_{\varphi=0}^{2\pi} \int_{\Theta=0}^{54.733° - \arcsin((\sqrt{2/3})\cdot(\sigma_c/\sigma))} d_{33}(\Theta) \\ & \times (3N/2\pi)(1 - (4 \arccos(\coth\Theta \\ & + \sigma_c/(\sigma \sin\Theta)))/\pi) \sin\Theta \, d\Theta \, d\varphi, \end{aligned} \tag{12}$$

where $d_{33}(\Theta)$ is Eq. (1).

In order to simplify the calculations, in formula (12) we neglect the fraction of the crystallites β_c whose **c** axes drift out of the zone of partial stability as a result of 90° switching. Having separated the contributions of the 90° and 180° switching (from formula (12), for example, for the zone of total stability of the polar axes the density of the **c** axes for 180° switching is $\rho^{180°} = N/2\pi$, and that for 90° switching is $\rho^{90°} = N/\pi$), we determined the relative number of 90° and 180° domain switchings accomplished under the stresses σ (for a given σ in percent) $A^{90°}$ and $A^{180°}$ (%). The piezocoefficient d_{33}^T of a ceramic plate exposed to the stress σ can be finally calculated from

$$d_{33}^T = d_{33} - \Delta d_{33}^T, \tag{13}$$

where d_{33} is the piezocoefficient obtained using formula (1).

Table I give the experimental and theoretical values of the piezocoefficient $d_{33} = d_{33}(\sigma)$ and the relative number of 90° and 180° domain reorientations completed in the ceramic under the longitudinal stress σ. It can be seen that as the longitudinal stress σ increases from 25×10^6 to 221×10^6 N/m^2, the piezocoefficient d_{33}^{exp} decreases from 360 to 68 pC/N which confirms the data given in Refs. 1 and 2.The relative number of 90° rotations under the action of the longitudinal stress σ increases fairly rapidly for TsTS-83G ceramic (for $\sigma = 91.9 \times 10^6$ N/m^2, this number is $A^{90°} = 88\%$). X-ray structural analysis of BaTiO$_3$ ceramic exposed to the biaxial equilibrium compressive stress σ_{22} (TsTS-83G ceramic is ''ferrohard'' compared with BaTiO$_3$ ceramic) indirectly confirm the theoretical distribution of 90° domain reorientations in TsTS-83G ceramic (for BaTiO$_3$ ceramic under $\sigma_{22} = 4 \times 10^6$ N/m^2 we obtain $A^{90°} = 45\%$; for $\sigma_{22} = 30 \times 10^6$ N/m^2 this figure is $A^{90°} = 80\%$) (Ref. 14). The theoretical analysis indicates that above stresses $\sigma > 100 \times 10^6$ N/m^2 all possible 90° rotations (for given σ) take place in TsTS-83G ceramic and an antiparallel 180° domain structure forms where the relative number of 180° domain switchings $A^{180°}$ increases to 75% (for $\sigma = 221 \times 10^6$ N/m^2), which is consistent with the experimental results from Ref. 6. The contribution of the displacement σ is negligible so this was neglected in the estimates of $A^{90°}$ and $A^{180°}$ (in lead titanate ceramic the contribution of the displacements of the 90° domain boundaries is 0.14–0.61 pC/N) (Ref. 15).

CONCLUSIONS

1. A static model of a critically polarized ceramic exposed to the longitudinal stress σ has been used to study the behavior of the piezocoefficient d_{33} of TsTS-83G ceramic and a quantitative analysis has been made of the 90° and 180° domain reorientations $A^{90°} = A^{90°}(\sigma)$ and $A^{180°} = A^{180°}(\sigma)$.

2. It has been shown that for TsTS-83G ceramic exposed to stresses $\sigma > 100 \times 10^6$ N/m^2 all possible 90° domain rotations occur and an antiparallel 180° domain structure forms.

[1] V. Z. Borodin, V. A. Doroshenko, O. P. Kramarov *et al.*, *Piezoelectric Materials and Transducers* [in Russian], Rostov State University Press, Rostov-on-Don (1971), pp. 49–58.
[2] H. H. A. Krueger, J. Acoust. Soc. Am. **12**, 636 (1967).
[3] H. F. Kay, Acta Crystallogr. **1**, 229 (1948).
[4] E. G. Fesenko, V. G. Gavrilyachenko, and A. F. Semenchev, *Domain Structure of Multiaxial Ferroelectric Crystals* [in Russian], Rostov State University Press, Rostov-on-Don (1990), 19 pp.
[5] S. M. Yufatova, A. F. Semenchev, V. G. Gavrilyachenko *et al.*, Pis'ma Zh. Tekh. Fiz **4**, 12 (1978) [Sov. Tech. Phys. Lett. **4**, 286 (1978)].
[6] V. A. Doroshenko, Dissertation for Candidate's Degree [in Russian], Rostov (1971).
[7] S. V. Bogdanov, B. M. Vul, and R. L. Razbash, Kristallografiya **6**, 72 (1961) [Sov. J. Crystallogr. **6**, 58 (1961)].
[8] E. V. Sinyakov and S. A. Flerova, Fiz. Tverd. Tela (Leningrad) **12**, 2728 (1970) [Sov. Phys. Solid State **12**, 2196 (1970)].
[9] V. A. Doroshenko, O. P. Kramarov, and G. N. Vitalinskaya, *Piezoelectric Transducers* [in Russian], Rostov State University Press, Rostov-on-Don (1972), pp. 57–65.
[10] A. F. Devonshire, Philos. Mag. **40**, 1040 (1949).
[11] M. Marutake, J. Phys. Soc. Jpn. **11**, 807 (1956).
[12] M. G. Minchina and V. P. Dudkevich, Zh. Tekh. Fiz **68**(7), 75 (1998) [Tech. Phys. **43**, 814 (1998)].
[13] *Polarization of Piezoceramics*, edited by E. G. Fesenko [in Russian], Rostov State University Press, Rostov-on-Don (1962), 18 pp.
[14] V. V. Kuleshov, Dissertation for Candidate's Degree [in Russian], Rostov (1980).
[15] E. I. Bondarenko, V. Yu. Topolov, and A. V. Turik, *Piezoactive Materials: Physics, Technology, and Application in Devices* [in Russian], Rostov State University Press, Rostov-on-Don (1991), pp. 33–37.

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Influence of sample dimensions on the propagation velocity of magnetoelastic waves in amorphous metal alloys

A. A. Gavrilyuk, A. V. Gavrilyuk, and N. P. Kovaleva

Irkutsk State Pedagogical University, 664011 Irkutsk, Russia

(Submitted November 11, 1997; resubmitted May 15, 1998)
Zh. Tekh. Fiz. **69**, 50–54 (June 1999)

An investigation is made of the propagation velocity of magnetoelastic vibrations in amorphous metal alloys as a function of the sample size in the direction of propagation of the magnetoelastic vibrations and the external magnetic field. The experimentally observed results are at variance with the conclusions of the model of uniform rotation of the magnetization usually used to describe the magnetoelastic properties of amorphous metal alloys. It is shown that changes in the structure of the domain walls (Bloch–Néel transition) in ribbons of amorphous metal alloys as a result of the action of the external magnetic field must be taken into account to obtain an adequate description of the observed results. © *1999 American Institute of Physics.* [S1063-7842(99)00906-X]

Studies of the magnetic properties of iron-based amorphous metal alloys are currently attracting widespread attention. These alloys have low coercive forces and high magnetostriction constants so that they can be used as sensors in various types of acoustic and ultrasonic transducers, delay lines, and so on. The possibility of using iron-based amorphous metal alloys for this purpose is attributable to the magnetoelastic vibrations induced in them under the action of an varying magnetic field. The propagation characteristics of magnetoelastic vibrations in amorphous metal alloys are mainly determined by the characteristics of rearrangement of their domain structure. Unfortunately, at present the relationship between the domain structure rearrangement processes in amorphous metal alloys and the propagation of acoustic vibrations in these alloys has not been studied sufficiently comprehensively, and this has delayed the practical application of these materials.

The aim of the present study is to investigate the propagation velocity of magnetoelastic waves as a function of the demagnetization factor of the sample in the direction of propagation of the magnetoelastic waves.

We shall consider ferromagnets with a positive magnetostriction constant ($\lambda_s > 0$) in the form of a narrow strip with the easy magnetization axis perpendicular to its length. The domain structure consists of domains of opposite magnetization separated by 180° walls (Fig. 1). This type of domain structure occurs in amorphous metal ribbons when they are annealed in a static magnetic field, which induces uniaxial anisotropy. A static magnetic field H and varying elastic stresses σ which excite magnetoelastic vibrations act along the length of the strip in the direction of the difficult magnetization axis. The action of these factors changes the orientation of the magnetization in the domains. It follows from the model of uniform rotation of the magnetization[1,2] that an expression for the angle of rotation α of the magnetization in the domains may be obtained from the condition for minimum domain structure energy W, which contains the uniaxial anisotropy energy, the energies of interaction of the magnetization with the external magnetic field H and the elastic stresses σ, and the magnetostatic energy caused by the appearance of magnetic scattering fields from the magnetization component perpendicular to the easy axis

$$W = K\sin^2\alpha - M_s H \mu_0 \sin\alpha - \frac{3}{2}\lambda_s \sigma \sin^2\alpha + \frac{1}{2} N M_s^2 \mu_0 \sin^2\alpha, \tag{1}$$

where K is the uniaxial anisotropy constant, μ_0 is the magnetic permeability of vacuum, M_s is the saturation magnetization, and N is the demagnetization factor in the direction of the sample length.

From the condition $dW/d/\alpha = 0$ the expression for the angle of rotation of the magnetization can be written as

$$\alpha = \arcsin\left(M_s \mu_0 H / (2K - 3\lambda_s \sigma + N M_s^2 \mu_0)\right). \tag{2}$$

The magnetoelastic deformation $\varepsilon_{\mathrm{m.e}} = (3/2)\lambda_s \sin^2\alpha$ can be given as

$$\varepsilon_{\mathrm{m.e}} = \frac{3}{2}\lambda_s [M_s^2 H^2 \mu_0^2 / (2K - 3\lambda_s \sigma + N M_s^2 \mu_0)^2]. \tag{3}$$

The elastic modulus in the magnetic field E_H can be determined from

$$(1/E_H) = (1/E_0) + d\varepsilon_{\mathrm{m.e}}/d\sigma, \tag{4}$$

where E_0 is the elastic modulus in the demagnetized state.

The expression for E_H can then be written as

$$E_H = E_0\{1 - [9\lambda_s^2 M_s^2 H^2 \mu_0^2 E_0 / ((2K - 3\lambda_s \sigma + N M_s^2 \mu_0)^3 + 9\lambda_s^2 M_s^2 H^2 \mu_0^2 E_0)]\}. \tag{5}$$

Using expressions which give the magnetoelastic resonance frequency f_r and the propagation velocity of the magnetoelastic vibrations $V_{\mathrm{m.e}}$ as a function of the elastic modulus

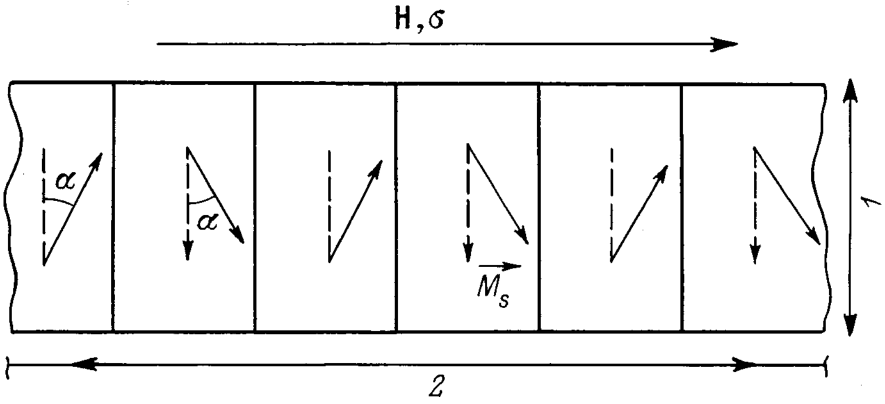


FIG. 1. Domain structure configuration being studied: *1* — easy magnetization axis, *2* — difficult magnetization axis.

$$f_r=(1/2L)\cdot(E_H/\rho)^{1/2}, \tag{6}$$

$$V_{\mathrm{m.e}}=2L\cdot f_r, \tag{7}$$

where L is the sample length in the direction of propagation of the magnetoelastic vibrations and ρ is the sample density, we obtain the following expression for $V_{\mathrm{m.e}}$:

$$V_{\mathrm{m.e}}=\{(E_0/\rho)[1-(9\lambda_s^2M_s^2H^2\mu_0^2E_0/((2K-3\lambda_s\sigma+NM_s^2\mu_0)^3+9\lambda_s^2M_s^2H^2\mu_0^2E_0))]\}^{1/2}. \tag{8}$$

According to Kim,[3] an approximate expression for the demagnetization factor N in the direction of the sample length may be written in the form

$$N=(1/L)/[(1/L)+(1/a)+(1/b)], \tag{9}$$

where a is the width and b is the sample thickness.

Figure 2 gives the calculated dependence of the propagation velocity of the magnetoelastic vibrations on the sample length L. For the calculations we used the following values of the sample parameters typical of iron-based amorphous metal alloys, and for the external magnetic field, the and elastic stresses: $M_s=5\times10^5$ A/m, $K=50$ J/m^3, $\lambda_s=3\times10^{-5}$, $\sigma=10^6$ Pa, $H=100$ A/m, $\rho=5000$ kg/mm^3, $E_0=1.4\times10^{11}$ Pa, $a=0.002$ m, and $b=2.5\times10^{-5}$ m.

It can be seen from this curve that as the sample length decreases, the propagation velocity of the magnetoelastic vibrations should increase monotonically. The calculations also show that as the magnetic field directed along the difficult magnetization axis increases, the dependence $V_{\mathrm{m.e}}(H)$ should decrease monotonically.

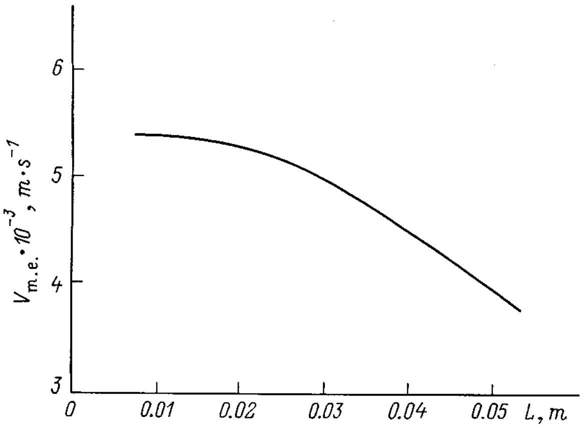


FIG. 2. Calculated dependence of the propagation velocity of the magnetoelastic waves $V_{\mathrm{m.e}}$ on the sample length L in the direction of propagation of these magnetoelastic vibrations.

In order to check these calculations experimentally, we investigated the propagation velocity of magnetoelastic vibrations as a function of the sample length in the direction of propagation of these vibrations in iron-based amorphous metal alloys.

The sample used was an $Fe_{81.5}B_{13.5}Si_{3.5}C_2$ amorphous metal alloy obtained by rapid quenching from a melt and annealed at 430° in a static magnetic field of 40 000 A/m for 20 min. During annealing the static magnetic field was directed perpendicular to the sample length. The sample was 2.5×10^{-5} m thick and 0.002 m wide. The length was varied between 0.045 and 0.015 m in 0.005 m steps. The propagation velocity of the magnetoacoustic vibrations was measured by a resonance–antiresonance method[4] as follows. A section of the ribbon sample is inserted in an induction coil which records the change in the magnetic flux caused by the action of magnetic fields on the sample. This coil together with the sample is then inserted inside a coil which generates a varying magnetic field which induces magnetoelastic vibrations in the sample. The varying magnetic field does not exceed 10 A/m. The system of coils is placed between Helmholtz coils which generate a static magnetic field. The static and varying magnetic fields are directed along the difficult magnetization axis. We measure the complex impedance of the recording coil whose maximum corresponds to the magnetoelastic resonance frequency f_r (Ref. 5). Using the experimentally determined magnetoelastic resonance frequency, we calculate the velocity of the magnetoelastic vibrations from the expression $V_{\mathrm{m.e}}=2Lf_z$

Figure 3 gives the propagation velocity of the magnetoelastic vibrations $V_{\mathrm{m.e}}$ as a function of the sample length L for various values of the external magnetic field directed along the difficult magnetization axis. An analysis of these curves indicates that the external magnetic field strongly influences their behavior. In weak magnetic fields (H=80–240 A/m) the curve $V_{\mathrm{m.e}}(L)$ has a minimum which shifts toward smaller L as the static magnetic field increases. In fairly strong magnetic fields (H = 320 A/m) for all L we observed a decrease in the propagation velocity of the magnetoelastic vibrations. Consequently, the calculated curves

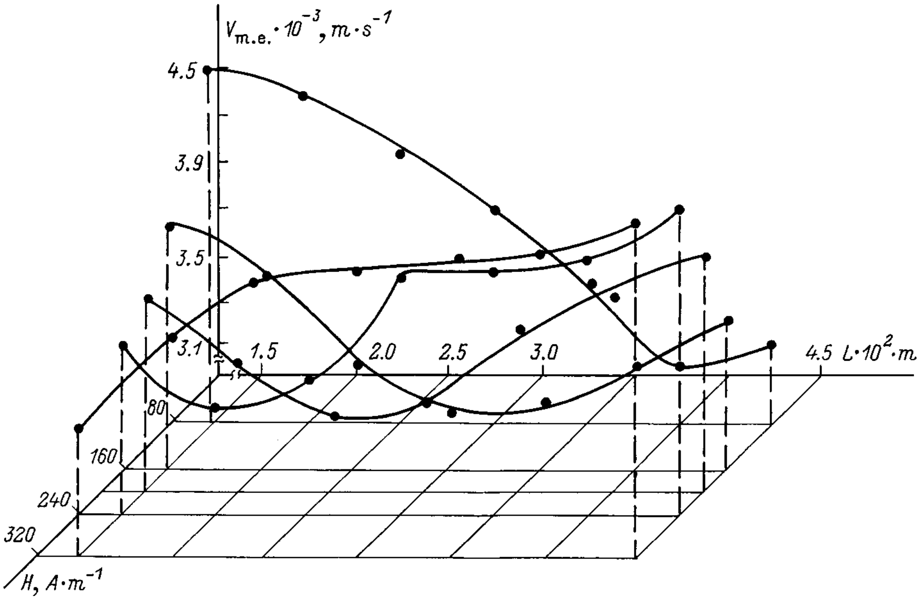


FIG. 3. Experimental dependences of the propagation velocity of the magnetoelastic vibrations $V_{\rm m.e}$ on the sample length L in the direction of propagation of the magnetoelastic waves for various values of the static magnetic field directed along the difficult magnetization axis.

(Fig. 2) only show satisfactory agreement with the experimental results for weak magnetic fields and small L.

Figure 4 gives the propagation velocity of the magnetoelastic vibrations $V_{\rm m.e}$ as a function of the static magnetic field H for $Fe_{81.5}B_{13.5}Si_{3.5}C_2$ samples of various lengths L. It can be seen that the curve $V_{\rm m.e}(H)$ has a minimum for all L. The position of this minimum is determined by the sample length L. The lower the value of L, the stronger the magnetic fields in which the minimum of $V_{\rm m.e}(H)$ is observed.

The model of uniform rotation of the magnetization[1,2] provides no explanation for the appearance of these minima of $V_{\rm m.e}(L)$ and $V_{\rm m.e}(H)$. The following explanation may be put forward for the experimental results. Annealing these samples in a magnetic field induces uniaxial anisotropy in the plane where the easy axis is perpendicular to the length of the ribbons. This leads to the formation of a striped domain structure with oppositely magnetized domains separated by 180° Bloch domain walls. When a varying magnetic field is applied in the direction of the difficult magnetization axis of narrow ribbons of amorphous metal alloys a few tens of micron thick, the domain walls oscillate around their equilibrium position. Brouha and van der Borst[5] showed that this effect may be explained using concepts of a Bloch–Néel transition of the domain-wall structure and redistribution of the magnetic poles at the edges of the sample under the action of a magnetic field. As a result of a Bloch–Néel transition of the domain wall structure, the equilibrium period of the striped domain structure corresponding to its energy

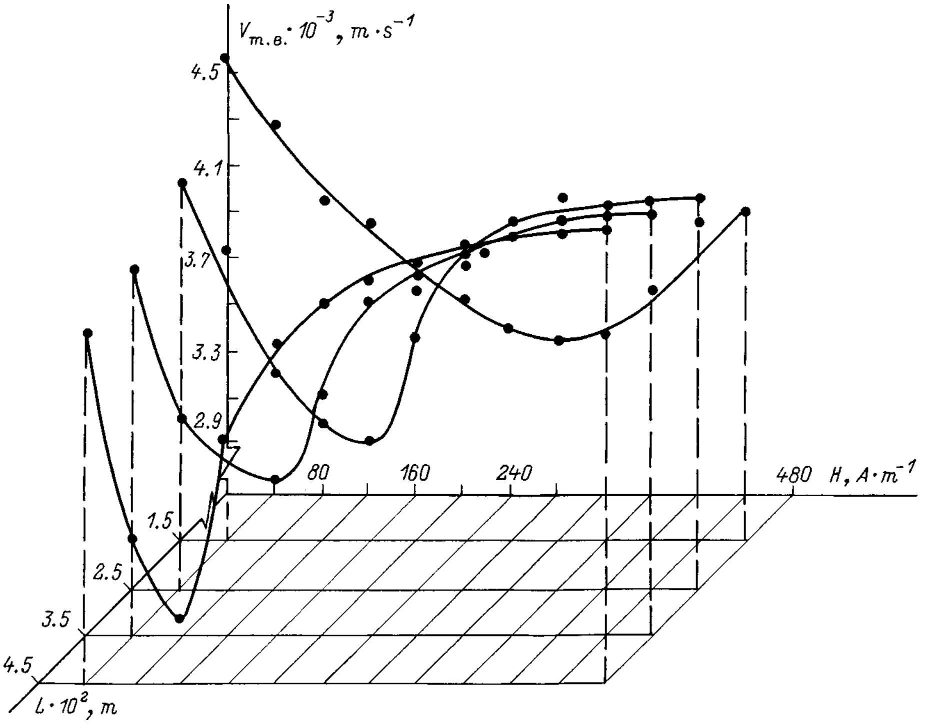


FIG. 4. Experimental dependences of the propagation velocity of the magnetoelastic vibrations $V_{\rm m.e}$ on the static magnetic field H directed along the difficult magnetization axis for various lengths L.

minimum decreases and the density of Néel domain walls $n(H)$ increases.[6] Taking into account the contribution of the energy of the Néel domain walls to the total energy of the domain structure, the expression for the propagation velocity of the magnetoelastic waves may be expressed in the form

$$V_{\mathrm{m.e}}=\{(E_0/\rho)[1-(9\lambda_s^2(M_sH\mu_0+2n\gamma)^2E_0/((2K-3\lambda_s\sigma+NM_s^2\mu_0+2n\gamma)^3+(9\lambda_sM_sH\mu_0+2n\gamma)^2E_0))]\}^{1/2}, \tag{10}$$

where γ is the energy density of the Néel domain walls.

In magnetic fields exceeding the fields for which the domain-wall structure undergoes a Bloch–Néel transition, the propagation velocity of the magnetoelastic vibrations increases as a result of an increase in the density of Néel domain walls *n*. Hence, the minima on the curves of the propagation velocity of the magnetoelastic vibrations should be observed near the field for which the domain-wall structure undergoes a Bloch–Néel transition. Middellhoek[7] showed that the energy density of the Néel domain walls in magnetically soft materials a few tens of micron thick is 8–10 J/m^3 and the energy density of the Bloch domain walls is 1–3 J/m^3. At these domain-wall energy densities the field for a Bloch–Néel transition does not exceed half the effective anisotropy field of the sample, which shows good agreement with the experimental results.

The effective anisotropy field and consequently the Bloch–Néel transition field are strongly influenced by the demagnetization factor of the sample in the direction of application of the external magnetic field. The smaller the sample length *L* in the direction of application of the field, the greater the demagnetization factor *N*. An increase in the demagnetization factor leads to a decrease in the angle of rotation of the magnetization and increases the field for a Bloch–Néel transition. In this case, the minimum on the curve of the propagation velocity of the magnetoelastic vibrations also shifts toward stronger fields.

In fairly strong fields exceeding the field for a Bloch–Néel transition the minimum on the curve of $V_{\mathrm{m.e}}(L)$ should be shifted toward smaller *L*. In the range of *L* being studied it is therefore probable that the curve of $V_{\mathrm{m.e}}(L)$ had no minimum in a magnetic field $H=320$ A/m.

From these investigations we can draw the following conclusions.

1. The dependences of the propagation velocity of magnetoelastic vibrations in iron-based amorphous metal alloys having a periodic domain structure are not monotonic functions of the sample length and external magnetic field. The results cannot be explained merely in terms of the model of uniform rotation of the magnetization usually used to describe the magnetoelastic properties of amorphous metal alloys.

2. The calculated and experimental results show satisfactory agreement when the domain wall structure of the sample is taken into account. The existence of a minimum on the curves of the propagation velocity of the magnetoelastic vibrations as a function of the external magnetic field and the sample length can be explained using concepts of a Bloch–Néel transition of the domain-wall structure.

3. The external magnetic field *H* strongly influences the behavior of $V_{\mathrm{m.e}}(L)$. As *H* increases, the minimum of $V_{\mathrm{m.e}}(L)$ shifts toward smaller *L*. This result may be explained by the fact that as the sample length decreases, the effective anisotropy field increases and therefore the field for a Bloch–Néel transition, which corresponds to the minimum propagation velocity of the magnetoelastic vibrations, also increases.

[1] J. D. Livingston, Phys. Status Solidi A **70**, 591 (1982).
[2] M. L. Spano, K. B. Hathaway, and H. T. Savage, J. Appl. Phys. **53**, 2667 (1982).
[3] P. D. Kim, Doctoral Dissertation [in Russian], Krasnoyarsk, (1988) 24 pp.
[4] T. O'Dell, Phys. Status Solidi A **74**, 565 (1982).
[5] M. Brouha and J. van der Borst, J. Appl. Phys. **50**, 7594 (1979).
[6] A. L. Petrov, A. A. Gavrilyuk, and S. M. Zubritskiĭ, Fiz. Met. Metalloved. **80**(6), 47 (1995).
[7] S. Middellhoek, J. Appl. Phys. **34**, 1054 (1963).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Microwave excitations of a domain wall in a cubic magnet with induced anisotropy

A. M. Alekseev, H. Dötsch, N. E. Kulagin, A. F. Popkov, and V. T. Synogach

Moscow Institute of Electronic Technology (Technical University), 103498 Moscow, Russia

(Submitted December 29, 1997; resubmitted October 23, 1998)
Zh. Tekh. Fiz. **69**, 55–62 (June 1999)

Theoretical and experimental studies are made of the spectra of spin wave excitations localized at a Bloch domain wall in a ferromagnet with combined cubic and biaxial anisotropy in [100]- or [110]-oriented bismuth-containing iron garnet films. This analysis of the spectra is used as the basis to calculate diagrams of stable states of a homogeneous Bloch domain wall as a function of the magnetic parameters in these materials. Results of measurements of the oscillation spectrum of an isolated domain wall in the mega- and gigahertz frequency ranges are presented. A description is given of an effect where the spin-wave excitations of the Goldstone translational branch and the high-frequency Gilinskii unidirectional mode are hybridized.
[S1063-7842(99)01006-5]

INTRODUCTION

The resonant response of domain walls to microwave electromagnetic radiation has been studied by many authors.[1–5] Lührmann *et al.*[5] observed antiphase resonant oscillations of the domain walls in films of easy-axis iron garnets in the gigahertz range. In recent studies of microwave excitations of the domain walls in easy-plane iron garnet films, high-frequency branches[6] associated with the Gilinskii mode[7] were observed in addition to the low-frequency Goldstone modes. As well as being of scientific interest, these investigations of the microwave excitations of the domain walls are also of technical interest for light modulation,[8] optical mode conversion, and other applications in magnetooptics.[9] It should be noted that potentially useful magnetic-film materials based on Bi-substituted iron garnets exhibit induced uniaxial and basis anisotropy as well as cubic anisotropy.[10] However, we know that the spectrum of the spin waves localized at a domain wall depends strongly on the anisotropy of the magnetic substance.[11] It is therefore interesting to analyze the spectrum of elementary excitations of a domain wall as a function of the ratio of the constants of magnetocrystalline and induced anisotropy. Some studies along these lines were reported in Ref. 12 for iron garnet films grown in the [110] plane. Here we report a detailed analysis of the spectrum of spin-wave excitations and construct diagrams showing the stability of Bloch domain walls in orthorhombic and tetragonal ferromagnets as a function of the magnetic parameters.

INITIAL EQUATIONS

The spin-wave excitations of the domain wall of a ferromagnet are described by the Landau–Lifshitz equations of magnetodynamics and magnetostatics

$$\frac{\partial \mathbf{M}}{\partial t} = -\gamma[\mathbf{M}\times\mathbf{H}_{\mathrm{eff}}], \tag{1}$$

$$\operatorname{div}(\mathbf{h}_d + 4\pi\mathbf{M}) = 0, \tag{2}$$

where $\mathbf{M}$ is the magnetization vector, γ is the gyromagnetic ratio, $\mathbf{H}_{\mathrm{eff}} = -(\delta F/\delta\mathbf{M})$ is the effective magnetic field, F is the thermodynamic potential of the ferromagnet, and $\mathbf{h}_d$ are the degaussing fields.

We shall first consider a cubic ferromagnet with orthorhombic induced anisotropy along the [110] crystal axis. We shall take the coordinate axes in the crystallographic directions $\mathbf{e}_x = [001]$, $\mathbf{e}_y = [110]$, and $\mathbf{e}_z = [\bar{1}10]$.

The thermodynamic potential can then be expressed in the form:

$$F = A(\nabla\mathbf{m})^2 - K_u(m_y)^2 - K_i(m_z)^2 + K_1\left[\frac{1}{4}(m_y^2 - m_z^2)^2 + m_x^2(1 - m_x^2)\right] - \mathbf{M}\mathbf{h}_d - \frac{\mathbf{h}_d^2}{8\pi}, \tag{3}$$

where A is the exchange interaction constant, $\mathbf{m}$ is the unit vector directed along the magnetization vector such that $\mathbf{M} = M\mathbf{m}$, and K_u, K_i, and K_1 are the constants of uniaxial, orthorhombic, and cubic anisotropy, respectively (all negative).

We introduce the polar coordinates for the direction of magnetization $\mathbf{m} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$ and define the magnetic potential $\Psi : \mathbf{h}_d = \nabla\Psi$. For convenience we normalize the x, y, and z coordinates and the wave vector $\mathbf{k}$ to $\sqrt{|K_u|/A}$, and the time t and the frequency ω to $2\gamma|K_u|/M$. The domain wall which we shall consider is described by a static solution of the Landau–Lifshitz equation (1)

$$\vartheta_0 = \frac{\pi}{2}, \quad \sin^2\varphi_0 = 1/(\cosh^2(z\Delta) + \varepsilon\sinh^2(z\Delta)),$$

$$\Psi_0 = 0, \tag{4}$$

where $\Delta = \sqrt{1-\beta}$, $\varepsilon = 3\beta/4(1-\beta)$, and $\beta = K_1/K_u$.

We shall seek small harmonic deviations of the spins from the equilibrium distribution (4) in the form of a small correction to the main solution

$$\vartheta = \frac{\pi}{2} + \theta(z)\sin(\omega t - k_x x - k_y y),$$

1063-7842/99/44(6)/7/$15.00

$$\varphi=\varphi_0+\phi(z)\cos(\omega t-k_x x-k_y y),$$

$$\Psi=\psi(z)\sin(\omega t-k_x x-k_y y), \qquad (5)$$

where $\psi(z)$ is normalized to $4\pi M\sqrt{A/|K_u|}$.

In the first approximation these satisfy the linearized system of equations (1) and (2), which may be expressed in the form

$$\frac{\partial^2\theta}{\partial z^2}=\left[q_i-\beta+k^2-\left(2-\frac{9}{2}\beta\right)\sin^2\varphi_0-\frac{9}{4}\sin^4\varphi_0\right]\theta$$

$$+\omega\phi+\frac{1}{Q}\frac{\partial\psi}{\partial z},$$

$$\frac{\partial^2\phi}{\partial z^2}=\omega\theta+\left[1-\beta+k^2-\left(2-\frac{13}{2}\beta\right)\sin^2\varphi_0\right.$$

$$\left.-6\beta\sin^4\varphi_0\right]\phi-\frac{1}{Q}(k_x\sin\varphi_0-k_y\cos\varphi_0)\psi,$$

$$\frac{\partial^2\psi}{\partial z^2}=\frac{\partial\theta}{\partial z}+(k_x\sin\varphi_0-k_y\cos\varphi_0)\phi+k^2\psi, \qquad (6)$$

where $k=\sqrt{k_x^2+k_y^2}$, $Q=|K_u|/2\pi M^2$, and $q_i=K_i/K_u$.

The boundary conditions have the form

$$\lim_{z\to\pm\infty}\{\theta,\phi,\psi\}=0. \qquad (7)$$

The equations (6) with the boundary conditions (7) were solved numerically by two different methods. One of these, described in Ref. 11, involves determining the spectrum from the condition that the Wronskian of the system (6) vanishes at the joining point of the right and left triplets of solutions obtained by Runge–Kutta integration (‘‘regula falsi’’ method). The initial condition for each solution is obtained from the asymptotic form of the corresponding independent solutions of the linearized problem which satisfy the boundary condition (7). For example, the asymptotic behavior of the solutions on the right half of the domain walls for $z\to\infty$ is determined by the triplet of vectors with the indices $p_{1,2,3}$, satisfying the characteristic equation

$$p^6-p^4(1+q_i-2\beta+Q^{-1}+3k^2)+p^2[(q_i-\beta+Q^{-1}$$

$$+k^2)(1-\beta+k^2)+k^2(1+q_i-2\beta+Q^{-1}(k_y/k)^2$$

$$+2k^2)-\omega^2]+k^2[\omega^2-(q_i-\beta+k^2)(1-\beta$$

$$+Q^{-1}(k_y/k)^2)]=0. \qquad (8)$$

By using this solution as the initial condition, after integrating as far as the joining point, we can obtain a triplet of vectors in the phase space of the system which span the vector of the required solution. Similarly, using the asymptotic form at infinity to the left of the wall, i.e., for $z\to-\infty$, after integrating in the opposite direction, we can obtain another triplet of vectors which should also be span the required solution. The condition for joining of the right and left solutions is equivalent to vanishing of the determinant (Wronskian) from the six independent basis vectors obtained. The second method is a modification of that described in Ref. 12. It reduces the problem of searching for the spectral dependence to the condition that the eigenvalue in the Sturm–Liouville problem vanishes for the operator equation

$$\hat{L}\mathbf{a}=\lambda\mathbf{a}, \qquad (9)$$

where λ is an eigenvalue and the operator $\hat{L}$ has the form

$$\hat{L}=-\hat{E}\frac{\partial^2}{\partial z^2}+\hat{A}\frac{\partial}{\partial z}+\hat{B},$$

$\hat{E}$ is the unit matrix,

$$\hat{A}=\begin{pmatrix}0&0&Q^{-1}\\0&0&0\\1&0&0\end{pmatrix},$$

$$\hat{B}=\begin{pmatrix}\begin{matrix}q_i-\beta+k^2-\\-\left(2-\frac{9}{2}\beta\right)\sin^2\varphi_0-\\-\frac{9}{4}\beta\sin^4\varphi_0\end{matrix} & \omega & 0\\ \omega & \begin{matrix}1-\beta+k^2-\\-\left(2-\frac{13}{2}\beta\right)\sin^2\varphi_0-\\-6\beta\sin^4\varphi_0\end{matrix} & \begin{matrix}-\frac{1}{Q}(k_x\sin\varphi_0-\\-k_y\cos\varphi_0)\end{matrix}\\ 0 & k_x\sin\varphi_0-k_y\cos\varphi_0 & k^2\end{pmatrix}.$$

In general the boundary condition can be written in matrix form

$$\left.\frac{\partial\mathbf{a}}{\partial z}\right|_{z=\pm\infty}=\hat{D}\mathbf{a}, \qquad (10)$$

where $\hat{D}=\hat{R}\hat{P}\hat{R}^{-1}$, $\hat{P}$ is the diagonal matrix of the characteristic exponential functions, and $\hat{R}$ is a matrix whose columns are constructed of the eigenvectors of the solutions of the linearized system (9) with $\lambda=0$ at infinity.

This last condition follows from the asymptotic form of the solutions at the edge of the domain wall. In fact, we write the general solution at infinity in the form

$$\mathbf{a}=\begin{pmatrix}\theta\\ \psi\\ \psi\end{pmatrix}=\hat{R}\mathbf{e},$$

$$\hat{R}=\begin{pmatrix}1&1&1\\c_1&c_2&c_3\\d_1&d_2&d_3\end{pmatrix},\quad \mathbf{e}=\begin{pmatrix}A\exp(p_1 z)\\B\exp(p_2 z)\\C\exp(p_3 z)\end{pmatrix},$$

where *A*, *B*, and *C* are arbitrary constants, c_i and d_i are the amplitudes of the eigenvectors of the linearized system (9) at infinity.

The vector of its spatial derivative $\partial\mathbf{a}/\partial z$ can then be related to the vector of the general solution $\mathbf{a}$ using the corresponding matrix, and specifically

$$\frac{\partial\mathbf{a}}{\partial z}=\hat{R}\hat{P}\hat{R}^{-1}\mathbf{a}=\hat{D}\mathbf{a},$$

where

$$\hat{P}=\begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{pmatrix},$$

$$\hat{D}=\hat{R}\hat{P}\hat{R}^{-1}.$$

It can be shown that the boundary condition (10) can be used not only for real but also for complex values of the characteristic exponent. Both methods are equally effective for wave numbers $k>0.005$ but cease to work for very small wave numbers $k<0.005$. In this range a special asymptotic analysis of the solutions is required as in Ref. 11.

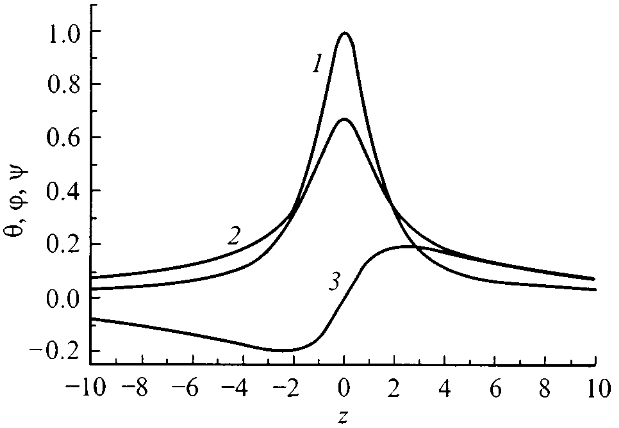


FIG. 2. Spectrum of the Goldstone mode (curve *1* in Fig. 1 for $k=0.23$; *1* — azimuthal angle φ, *2* — polar angle θ, and *3* — magnetic potential ψ.

SPECTRUM OF SPIN-WAVE EXCITATIONS OF BLOCH DOMAIN WALLS IN AN ORTHORHOMBIC FERROMAGNET

The methods described were used to make numerical calculations of the spectra of cubic ferromagnets with induced anisotropy. Not only the quantitative but also the qualitative behavior of the spectral dispersion curves varied as a function of the ratio of the magnetic parameters. In an orthorhombic ferromagnet described by the thermodynamic potential (3) the axisymmetric case considered by Gilinskii is obtained for $K_1=0$, $K_i=K_u$. In this case, in accordance with the theory[7] the spectrum of spin waves propagating along the domain wall perpendicular to the easy magnetization axis is asymmetric and initially contains three branches. Of these, two are the Goldstone translational modes and the third is a unidirectional high-frequency Gilinskii mode. At high wave numbers $k\gg 1$ we may observe additional high-frequency modes[11] localized at domain walls whose spectral branches become detached from the bottom of the continuous spectrum band. The spectrum of waves propagating along the anisotropy axis parallel to the magnetizations in the domains is symmetric and contains only two branches, which correspond to the translational modes at the beginning of the spectrum. In the presence of cubic or basis anisotropy, the asymmetry of the spectrum of spin waves propagating perpendicular to the magnetizations in the domains is intensified. The spectral branches of the high-frequency and adjacent translational branches repel each other so that for a certain value of the magnetic parameters the frequency of the translational mode may vanish at a finite wave number, as shown in Fig. 1 (see also Ref. 12). At this point the Bloch domain wall becomes unstable relative to nonuniform perturbations of its spins. Figure 2 shows the structure of the soft mode determining the instability of the Bloch domain wall for this case. The spectrum of the spin waves propagating parallel to the direction of magnetization in the domains has the usual form (Fig. 3. Figure 4 shows the calculated boundaries separating the region of stable Bloch domain wall in an orthorhombic ferromagnet with combined anisotropy from the region of parameters where this domain wall is unstable. Toward the inside of the region of instability, the low-frequency oscillation branch touching the abscissa begins to go over to negative frequencies so that a gap appears on the lower branch at positive frequencies. Because of the inversion symmetry of the spectrum $(\omega(k)=-\omega(-k))$, an additional branch appears in the positive half-plane in an identical region of wave numbers but of opposite sign. For a certain value of the anisotropy parameters this new oscillation branch may merge with the ‘‘old’’ Goldstone branch,

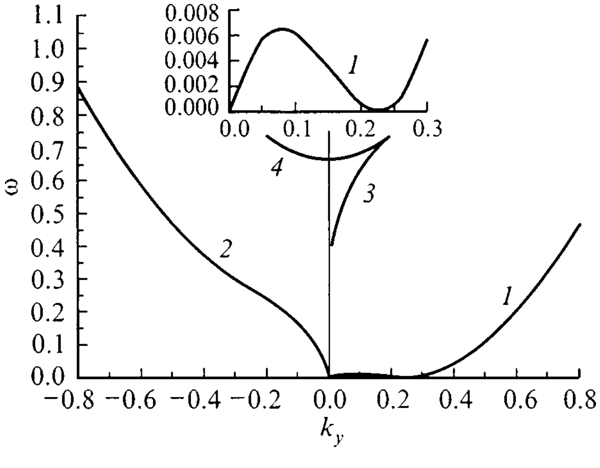


FIG. 1. Excitation spectrum of 180°-domain wall in an orthorhombic ferromagnet for $k_x=0$; $q_i=0.5$, $\beta=0.3$, and $Q=0.6624$; *1* and *2* — Goldstone mode, *3* — Gilinskii mode, and *4* — boundary of continuous spectrum.

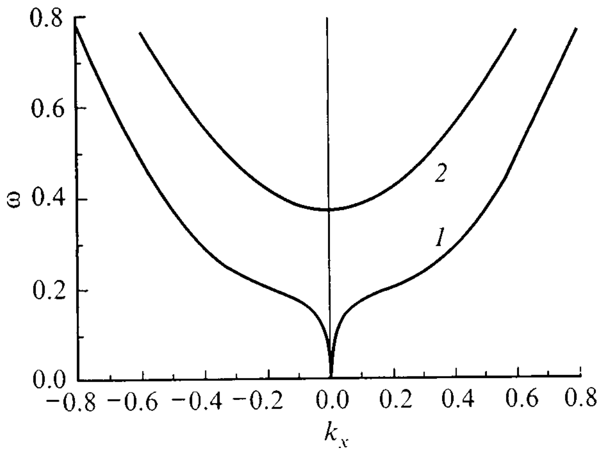


FIG. 3. Excitation spectrum of a 180° domain wall in an orthorhombic ferromagnet with $k_y=0$; $q_i=0.5$, $\beta=0.3$, and $Q=0.6624$; *1* — translational mode, *2* — boundary of continuous spectrum.

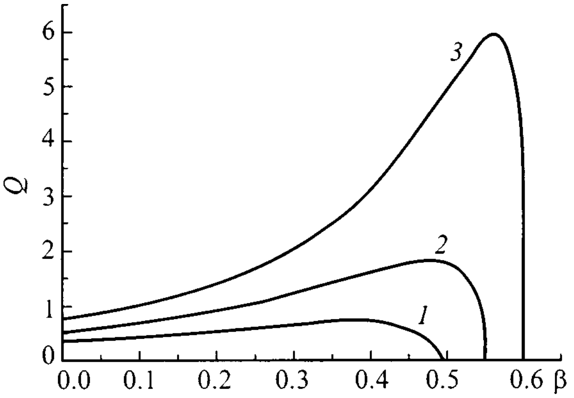


FIG. 4. Boundaries separating the regions of domain-wall stability and instability in an orthorhombic ferromagnet: $q_i = 0.5$ (*1*), 0.55 (*2*), 0.6 (*3*). (The region of instability lies above the curves).

disrupting it and forming a closed loop at the beginning of the spectrum, as shown in Fig. 5. Similar oscillation spectra appear in a stratified liquid flux having different densities and velocities under conditions of Kelvin–Helmholtz instability and turbulence.[13,14] The symmetric branches of the spin wave spectrum in cases where these propagate parallel to the magnetizations at the edge of the domain wall gradually become lower toward the inside of the unstable region until they come in contact and subsequently form a rough singularity at this point (Fig. 6). These spectral properties of the spin waves show that in the unstable region a homogeneous Bloch domain wall is a highly nonequilibrium formation. By analogy with the flow of a stratified liquid,[13,14] its evolution may be characterized by an exponential increase in the amplitudes of the negative-energy waves near the line of stability loss and strong Kelvin–Helmholtz instability far from this line. The final domain-wall structure may contain soliton-like and vortex formations.

SPECTRUM OF LOCALIZED SPIN WAVES AT A BLOCH DOMAIN WALL IN A TETRAGONAL FERROMAGNET

We shall analyze easy-plane films of cubic ferromagnets grown in the [100] plane. In the coordinate system $\mathbf{e}_x = [\overline{1}10]$, $\mathbf{e}_y = [001]$, $\mathbf{e}_z = [110]$ the thermodynamic potential for this case is written as follows:

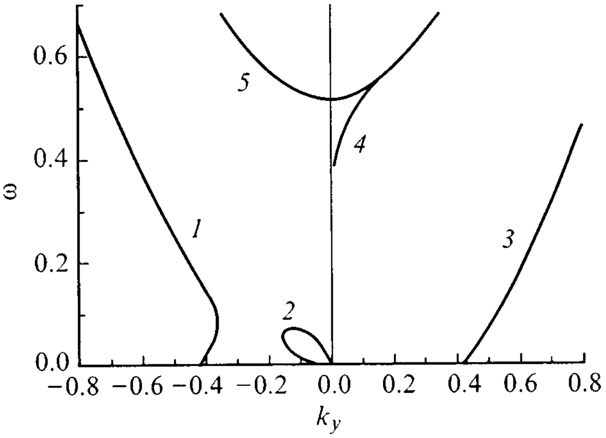


FIG. 5. Excitation spectrum of a 180° domain wall in an orthorhombic ferromagnet with $k_x = 0$, $q_i = 0.5$, $\beta = 0.3$, $Q = 1.6$; *1–3* — Goldstone mode, *4* — Gilinskii mode, and *5* — boundary of continuous spectrum.

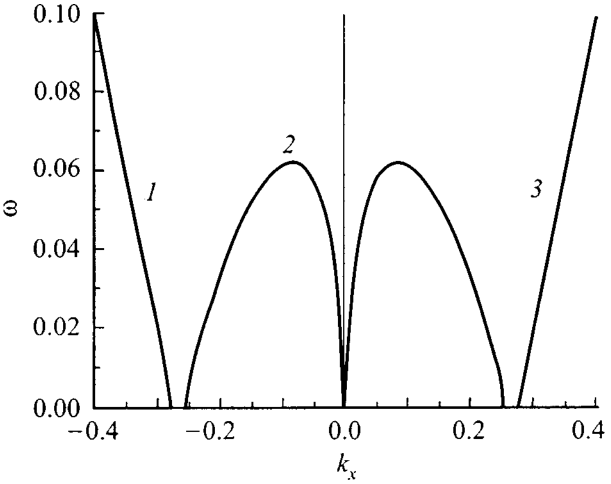


FIG. 6. Excitation spectrum of a 180° domain wall in an orthorhombic ferromagnet with $k_x = 0$, $q_i = 0.5$, $\beta = 0.3$, $Q = 1.6$; *1–3* — translational mode.

$$F = A(\nabla \mathbf{m})^2 - K_u (m_y)^2 + K_1 \left[\frac{1}{4} (m_x^2 - m_z^2)^2 + m_y^2 (1 - m_y^2) \right] - \mathbf{M}\mathbf{h}_d - \frac{\mathbf{h}_d^2}{8\pi}. \tag{11}$$

The structure of this Bloch domain wall obtained by solving the Landau–Lifshitz static equation is described by

$$\vartheta_0 = \frac{\pi}{2}, \quad \sin^2 \varphi_0 = 1/(\cosh^2(z\Delta) + \varepsilon \sinh^2(z\Delta)),$$

$$\Psi_0 = 0, \tag{12}$$

where

$$\Delta = \sqrt{1 - \beta/2}, \quad \varepsilon = \frac{3\beta}{4(1 - \beta/2)}, \quad \beta = K_1 / K_u.$$

In this case, the linearized system of equations under study has the form

$$\frac{\partial^2 \theta}{\partial z^2} = \left[\beta + k^2 - 2 \sin^2 \varphi_0 - \frac{9}{4} \beta \sin^4 \varphi_0 \right] \theta + \omega \phi + \frac{1}{Q} \frac{\partial \psi}{\partial z},$$

$$\frac{\partial^2 \phi}{\partial z^2} = \omega \theta + \left[1 + k^2 - 2 \sin^2 \varphi_0 - \frac{\beta}{2} (1 - 11 \sin^2 \varphi_0 + 12 \sin^4 \varphi_0) \right] \phi - \frac{1}{Q} (k_x \sin \varphi_0 - k_y \cos \varphi_0) \psi,$$

$$\frac{\partial^2 \psi}{\partial z^2} = \frac{\partial \theta}{\partial z} + (k_x \sin \varphi_0 - k_y \cos \varphi_0) \phi + k^2 \psi. \tag{13}$$

An analysis of the localized solutions of this system was made as in the previous case. A numerical experiment showed that the spin wave spectra typical of a stable homogeneous Bloch domain wall are not observed over the entire range of magnetic parameters for this particular case. A diagram of the Bloch domain wall states for this case (Fig. 7) shows the calculated boundaries of the region of strong

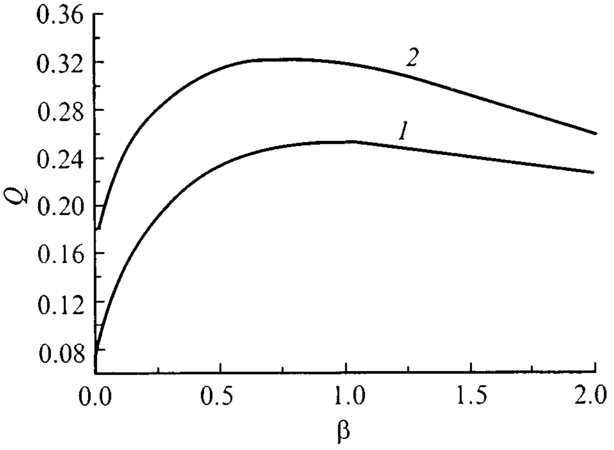


FIG. 7. Boundaries separating regions having different types of domain-wall instability in a tetragonal ferromagnet. Below curve *1* — region of weak instability, above curve *1* — region of existence of Goldstone mode loop, above curve *2* — region in which the spectrum for $k_y=0$ is similar to that shown in Fig. 6.

Kelvin–Helmholtz instability for a tetragonal-symmetry ferromagnet with combined anisotropy. In the region of weak instability the asymmetric spectrum of localized spin waves has no loops on the low-frequency branch at the beginning of the spectrum (Fig. 8). In the region of parameters above the critical curve a loop of the Goldstone mode is observed (Fig. 9). The spectrum of the symmetric propagation of spin waves along the magnetizations in the domains behaves as in the previous case of an unstable Bloch domain wall.

EXPERIMENTAL STUDY OF THE SPECTRA OF SPIN WAVES LOCALIZED AT DOMAIN WALLS, IN EASY-PLANE [100] AND [110] IRON GARNET FILMS

We investigated films of Bi-containing iron garnets grown epitaxially on [110]- and [100]-oriented gadolinium gallium garnet substrates, which correspond to the cases of orthorhombic and tetragonal ferromagnets considered above. All the films exhibited in-plane magnetization as a result of a suitable choice of cubic, uniaxial, and orthorhombic anisotropy parameters. In the [110]-oriented films the magnetization was directed along the $[00\bar{1}]$ axis (sample No. 1) or at an angle $\varphi\sim 18°$ to this axis (sample No. 2). The [100]-oriented films contained 180° domain walls parallel to the [010] and [011] directions, and also 90° domain walls parallel to the [010] and [001] axes (sample No. 3). The 180° domain walls studied exhibited a strong Bloch magnetization component perpendicular to the plane of the film. This was confirmed by the strong Faraday contrast at the walls observed in a polarizing microscope. Sample No. 3 also exhibited Bloch lines separating dark and light subdomains on the image of the domain wall. The numbering of the samples and the parameters of the materials were the same as in Ref. 6.

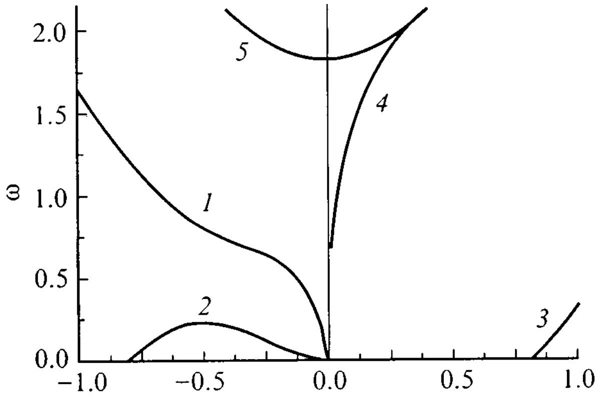


FIG. 8. Excitation spectrum of a 180° domain wall in a tetragonal ferromagnet with $k_x=0$, $\beta=0.5$, $Q=0.17$; *1–3* — Goldstone mode, *4* — Gilinskii mode, and *5* — boundary of continuous spectrum.

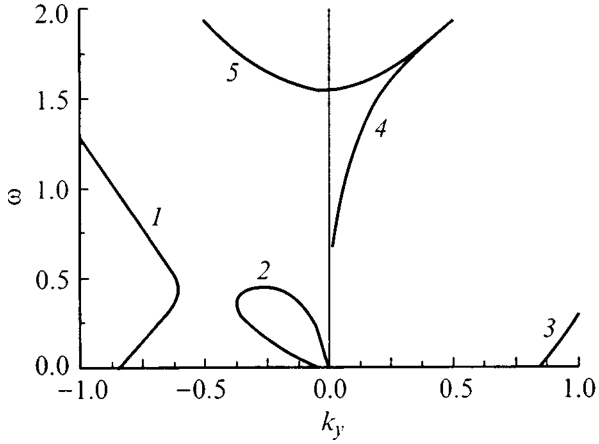


FIG. 9. Excitation spectrum of a 180° domain wall in a tetragonal ferromagnet with $k_x=0$, $\beta=0.5$, and $Q=0.25$; *1–3* — Goldstone mode, *4* — Gilinskii mode, and *5* — boundary of continuous spectrum.

In order to measure the response of the domain wall to an rf magnetic field, we used a planar structure comprising a slot line and a coplanar waveguide.[5,6] The dimensions of the structure did not exceed those of the domains ($\lesssim 100\,\mu$m) which allowed us to study the excitation of an isolated domain wall. The output signal from a microwave spin oscillator was fed to the structure input, i.e., the slot line. The signal produced by the domain-wall oscillations was recorded from the output, i.e., the coplanar waveguide. After being amplified and detected, these frequency dependences of the linear amplitude of the domain-wall oscillations were recorded using a digital storage oscilloscope and also recorded and processed on a PC.

Figures 10 and 11 show the measured frequency dependences of the domain-wall response in [110] films for samples Nos. 1 and 2, respectively. A set of narrow resonance lines with nonmonotonically increasing frequencies in the gigahertz range is clearly visible. We also observed a second set of almost equidistant resonances in a lower frequency range. Figure 12 shows the corresponding dispersion dependences of the resonance frequencies as a function of the resonance number.

Figure 13 shows a frequency scan of the domain-wall response measured in the [100] film (sample No. 3). The lower part of the figure shows the FMR signal measured from the same part of the film but in the absence of a domain wall (removed using an additional weak static magnetic field which was then switched off). The first resonance has a

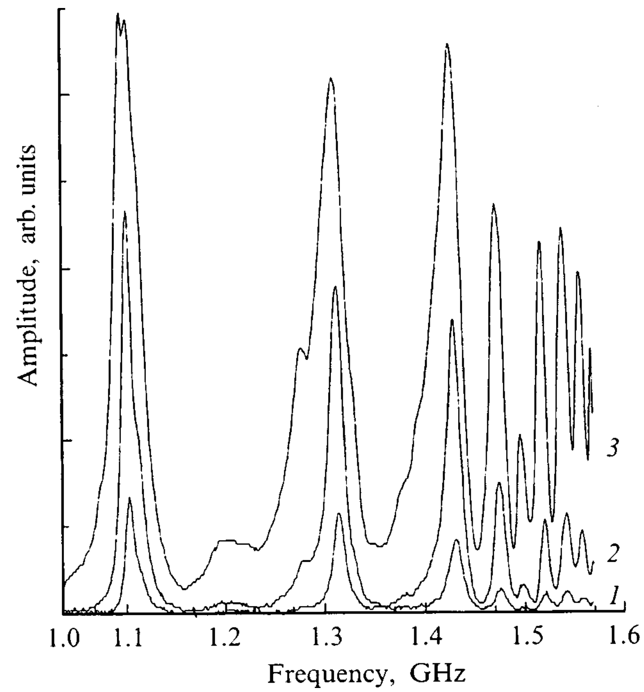


FIG. 10. Frequency response of an isolated 180° domain wall in a [110]-oriented bismuth garnet film (sample No. 1): *1–3* — oscillator power levels -35, -25, and -15 dBW, respectively.

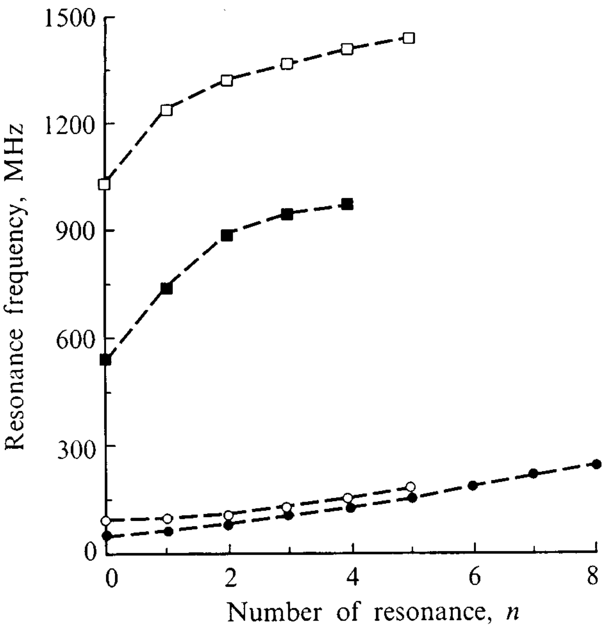


FIG. 12. Dispersion dependences of resonances frequencies as a function of the resonance number for sample No. 1 (■, ●) and No. 2 (□, ○); ●, ○ — resonances assigned to Goldstone mode, ■, □ — resonances assigned to high-frequency Gilinskii mode.

lower frequency than that for the [110] films although it has an equally narrow line width (4.5 MHz).

The frequencies of all the resonances increased with decreasing sample thickness and were almost independent of the other dimensions. From this it follows that both sets of resonances are caused by the excitation of standing domain-wall waves over the film thickness and belong to different branches of the spectrum. The low-frequency peaks with linear dispersion are associated with the translational Goldstone branch, i.e., standing flexural domain-wall waves. The high-frequency resonances with nonmonotonic dispersion belong to the unidirectional Gilinskii branch. In this range where the frequencies of the higher resonances of the Goldstone mode approached those of the lower resonances of the Gilinskii mode, we observed some overlap of the two systems of resonances (Fig. 11). This indicates that the two different branches in the spectrum of domain-wall excitations are hybridized. It should be noted that the type of standing wave formation observed in this particular case differs qualitatively from the classical case of excitation with a symmetric dispersion law. The forward-traveling wave belongs to the Gilinskii mode while the backward-traveling wave with the same ω belongs to the Goldstone mode. The superposition of these modes leads to the formation of a hybrid standing wave whose phase also depends on the spatial coordinates, as in the case of a pure Goldstone mode with asymmetric dispersion.[15] The condition for resonant excitation of a standing wave has the form $(k_+(\omega)-k_-(\omega))=2\pi n/d$, $n=0,\pm 1,\pm 2,\ldots$. Since the return wave has no gap unlike the forward wave, the hybrid standing wave will have a different gap value. Consequently, the phase velocity and gap calculated from the measured dispersion dependence of the high-frequency resonances $\omega(n)$ may differ appreciably from the values obtained in the Gilinskii theory (the experimental slit value for sample No 1 is ≈ 550 MHz and the theoretical value ≈ 140 MHz). A comparison between the theoretical and experimental results shows that the experimentally observed resonance frequencies are higher than the frequencies predicted theoretically using the model of an unbounded medium. This difference can also be attributed to the appearance of magnetic charges at the boundaries of the

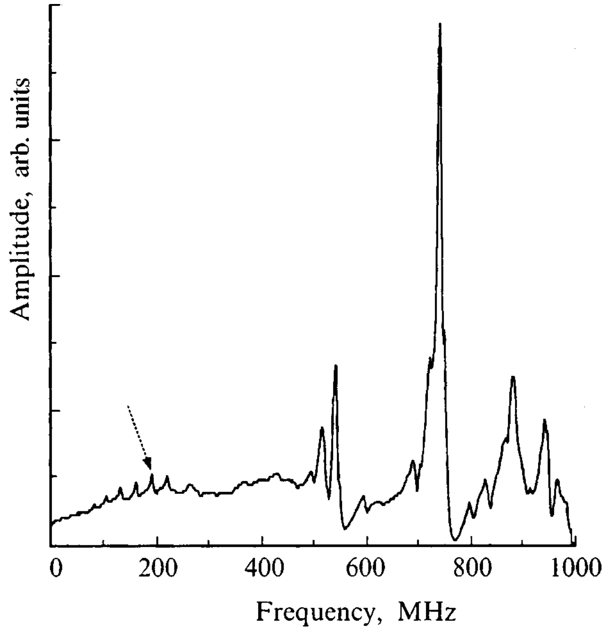


FIG. 11. Frequency response of an isolated 180° domain wall in a [110]-oriented film (sample No. 2). In addition to the high-frequency resonances, the arrow indicates low-frequency resonances assigned to the Goldstone mode.

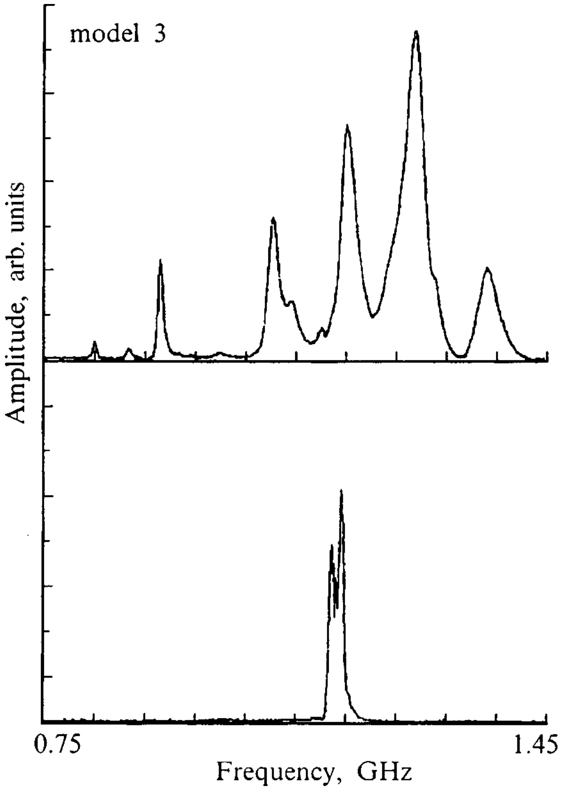


FIG. 13. Frequency dependence of the response of an isolated 180° domain wall in a [100]-oriented film (sample No. 3). The lower part of the figure shows the FMR signal measured for the same sample without a domain wall.

film which were neglected in the calculations.

CONCLUSIONS

To sum up, we have made theoretical calculations and experimental investigations of the spectra of localized spin waves at a domain wall in cubic ferromagnets with induced anisotropy. The calculations have shown that in [110]-oriented films there are regions of magnetic parameters in which the Bloch domain wall is unstable relative to nonuniform spin oscillations. We determined the soft mode for transition to the inhomogeneous state and the boundaries of the region of stability as a function of the constants of cubic and uniaxial anisotropy. We showed that the domain walls in [100]-oriented easy-plane films are unstable relative to nonuniform spin oscillations over the entire range of magnetic parameters. We calculated the boundaries of the regions of magnetic parameters in which strong Kelvin–Helmholtz instability may occur by analogy with the flow of a stratified hydrodynamic flux. Above this boundary soliton-like and vortex spin distributions may be predicted in an equilibrium domain-wall structure. Dispersion dependences of the resonant modes of the domain-wall oscillations were measured experimentally in the microwave frequency range up to the gigahertz range. The observed high-frequency modes of the spin-wave excitations were associated with the Gilinskii mode. A standing-wave resonance was observed in the high-frequency modes, which can be attributed to the hybridization of the modes in the translational and high-frequency spectral branches. The observed high-frequency resonances of the domain-wall oscillations may be used for magnetooptic light modulation in the gigahertz range.

This work was supported by a grant from the Russian Fund for Fundamental Research No. 97-02-16183.

[1] D. Polder and J. Smit, Rev. Mod. Phys. **25**, 89 (1979).
[2] J. O. Artman, Phys. Rev. **105**, 62 (1957).
[3] J. O. Artman and S. H. Charap, J. Appl. Phys. **50**, 2024 (1979).
[4] S. Y. Bi, D. J. Seagle, E. C. Myers *et al.*, IEEE Trans. Magn. **MAG-18**, 1337 (1982).
[5] B. Lührmann, H. Dötsch, and S. Sure, Appl. Phys. A: Solids Surf. **57**, 553 (1993).
[6] V. T. Synogach and H. Dötsch, Phys. Rev. B **54**, 15266 (1996).
[7] I. A. Gilinskiĭ, Zh. Éksp. Teor. Fiz. **68**, 1032 (1975) [Sov. Phys. JETP **41**, 511 (1975)].
[8] H. P. Winkler, H. Dotsch, B. Luhrmann, and S. Sure, J. Appl. Phys. **76**, 3272 (1994).
[9] H. Dotsch, A. Erdman, M. Fehndrich *et al.*, in *Nonlinear Microwave Processing: Towards a New Range of Devices*, edited by R. Marcelli and S. Nikitov (Kluwer, Drodrecht 1996).
[10] P. Hansen and K. Witter, J. Appl. Phys. **58**, 454 (1985).
[11] A. Mikhailov and I. Shimochin, Phys. Rev. B **48**, 9569 (1993).
[12] N. E. Kulagin, A. F. Popkov, V. T. Synogach, and H. Dötsch, J. Appl. Phys. **81**, 2336 (1997).
[13] R. A. Cairns, J. Fluid Mech. **92**, 1 (1979).
[14] S. A. Maslowe and L. G. Redekopp, J. Fluid Mech. **101**, 321 (1980).
[15] V. S. Gornakov, V. I. Nikitenko, I. A. Prudnikov, and V. T. Synogach, Phys. Rev. B **46**, 10 829 (1992).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Phase transitions in a titanium–silicon system under laser treatment in alkanes

A. M. Chaplanov and A. N. Shibko

Institute of Electronics, Academy of Sciences of Belarus, 220841 Minsk, Belarus

(Submitted January 14, 1998)
Zh. Tekh. Fiz. **69**, 63–66 (June 1999)

An investigation is made of the phase transitions taking place in a titanium–silicon composite exposed to $\lambda = 1.06\,\mu$m laser treatment in pentane and hexane. It is shown that the formation of titanium carbide, oxides, and silicides depends on the treatment parameters and conditions. The phase changes were investigated over the thickness of the film and their influence on the electrophysical parameters of the titanium–silicon contact was studied. © *1999 American Institute of Physics.* [S1063-7842(99)01106-X]

One of the most promising methods of obtaining Ohmic and rectifying contacts is laser treatment of materials, primarily because of its speed and the local nature of treating integrated circuits.[1,2] In the present paper we investigate the phase transitions in a titanium–silicon composite exposed to laser treatment in alkanes, i.e., pentane and hexane.

A 60 nm thick titanium film was deposited on a (111)-oriented *p*-type silicon substrate by electron beam deposition at a substrate temperature of 373 K. Before the film was deposited, the silicon wafer was treated chemically using a method described in Ref. 3. A $\lambda = 1.06\,\mu$m LTN-103 laser was used to treat the Ti–Si system as shown schematically in Fig. 1. A 5×5 mm sample positioned in a cell containing alkanes, either pentane (C_5H_{12}) or hexane (C_6H_{14}), was irradiated for 3 s using a scanning laser beam. The surface of the sample was treated uniformly for 5 s using the scanner. The laser radiation power was 1.5–7.0 W and was monitored with an IMO-2 power meter during the treatment process. The phase transitions taking place in the composite after laser treatment were investigated by electron diffraction analysis and by electron spectral and chemical analysis.

The investigations showed that laser treatment of a Ti–Si system leads to changes in the phase composition of the surface layer. Grain growth is observed accompanied by migration of grain boundaries and redistribution of defects in the crystal lattice. Titanium carbide, oxide, and silicide nuclei form and grow.

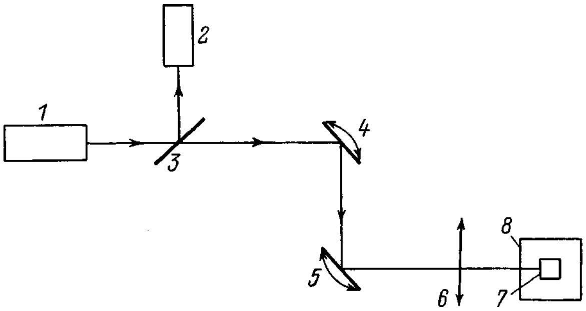


FIG. 1. Schematic of experimental apparatus: *1* — LTN-103 laser, *2* — IMO-2 laser radiation power meter, *3* — semitransparent mirror, *4* — mirror rotatable about *X* axis, *5* — mirror rotatable about *Y* axis, *6* — focusing lens, *7* — cell containing samples, and *8* — VUP-4 device.

After the Ti–Si system had been treated in pentane at 1.5 W laser radiation power, the electron diffraction patterns revealed diffraction rings assigned to the lowest titanium oxides Ti_2O_3 and Ti_3O_5 and titanium carbide TiC (Table I). When the laser radiation power was increased to 4.0 W, the electron diffraction patterns revealed not only titanium carbide and oxide rings but also diffraction rings ascribed to metal-enriched titanium silicide Ti_5Si_3. The oxide phases form at the surface of the composite as a result of interaction between the film and oxygen adsorbed during deposition and also from the ambient medium. Laser treatment of a titanium–silicon system in pentane is accompanied by the homolytic breaking of carbon–carbon bonds (cracking) in the carbon-containing liquid. Laser irradiation imparts sufficient energy to crack the pentane molecules, and the carbon atoms interact with the titanium film to form titanium carbide. A more detailed description of the formation of metal carbides on laser-irradiated thin metal films in carbon-containing liquids can be found in Ref. 4. As a result of the diffusion of carbon atoms into the film from the surrounding medium, numerous pores form in the surface layers both inside grains and along grain boundaries. As a result of laser treatment of Ti–Si in pentane at 7.0 W, the electron diffraction patterns reveal that in addition to the phases noted above, the highest-intensity peaks are from titanium silicide, TiSi. The titanium silicides are formed by the diffusion of silicon into the titanium film. The high concentration of defects and the grain boundaries in the metal film facilitate the diffusion of silicon to a considerable extent.[1] Laser annealing stimulates interdiffusion and interaction between the titanium and silicon. The sequence of the phase transitions reflects the interdiffusion kinetics in the Ti–Si system as the laser radiation power increases, resulting in the redistribution of titanium, carbon, oxygen, and silicon atoms and changes in the titanium–silicon interface. The change in the phase composition of the Ti–Si surface layer as a function of the treatment conditions may be represented by the following scheme:

$$\mathrm{Ti{-}Si} \xrightarrow{1.5\,\mathrm{W}} \mathrm{Ti_3O_5, Ti_2O_3, TiC} \xrightarrow{4.0\,\mathrm{W}} \mathrm{Ti_3O_5, TiC, Ti_5Si_3}$$

$$\xrightarrow{7.0\,\mathrm{W}} \mathrm{Ti_3O_5, TiC, TiSi, Ti_5Si_3}.$$

When a Ti–Si system is treated in hexane at a laser radiation power of 1.5 W, the electron diffraction patterns reveal TiC, Ti_3O_5, and Ti_2O_3 peaks (Table II and Fig. 2a). In addition to these phases, as the laser radiation power increases, nuclei of the Ti_5Si_3 phase also form and grow. When the system is treated in hexane at 7.0 W the electron diffraction pattern reveals peaks of the Ti_3O_5, TiC, Ti_5Si_3, and TiSi phases (Table II). An analysis of the results presented in Tables I and II shows that there are no significant differences in the phase compositions for treatment in pentane or hexane. However, for treatment in hexane the electron diffraction pattern reveals many diffraction rings ascribed to titanium carbide which indicates an increase in the titanium carbide content. Thus, we shall subsequently discuss the results obtained by treating the system in hexane.

In order to investigate the phase transitions taking place in a Ti–Si system after laser treatment in alkanes over the entire thickness of the titanium film, we etched a 20 nm thick film. As a result of 1.5 W laser irradiation of Ti–Si in hexane, at a depth of 20 nm from the surface the electron diffraction pattern reveals titanium silicide TiSi peaks in addition to the phases formed at the surface of the system. These TiSi peaks have the highest intensities. This indicates that as a result of 1.5 W treatment, titanium silicide forms at the titanium–silicon interface. When the irradiation power is increased to 4.0 W, growth of the TiSi titanium silicide phase is observed accompanied by nucleation of $TiSi_2$ enriched in the silicon modification S-54 (Table III). Titanium disilicide with an S-54 lattice was described by Yatsenko *et al.*[5] At a laser radiation power of 7.0 W the electron diffraction patterns reveal peaks of the TiSi and $TiSi_2$ phases in the same proportion (Fig. 2b). In addition to the silicides, titanium oxides and carbide are also present at a depth of 20 nm. This can be attributed to the diffusion of carbon atoms into the titanium film and the presence of oxygen adsorbed by the film during deposition which diffuses into the film as a result of degradation of the SiO_2 layer on the silicon surface.

An analysis of the results presented in Table IV shows that the phase composition of the system after laser treatment in alkanes at powers of 1-5–7.0 W at a depth of 40 nm consists mainly of titanium silicide and disilicide. The electron diffraction pattern of the treated samples also reveals isolated Si and SiO_2 (tridymite) peaks. A comparison of the results

TABLE I. Change in phase composition of Ti–Si system under laser treatment in pentane.

		E, W	
d, Å	1.5	4.0	7.0
4.28	Ti_3O_5	Ti_3O_5	Ti_3O_5
3.54	Ti_3O_5	Ti_3O_5	Ti_3O_5
3.14	Ti_3O_5	Ti_3O_5	Ti_3O_5
2.71	Ti_2O_3	···	···
2.57	Ti_2O_3	···	···
2.51	TiC	TiC	TiC
2.44	···	···	TiSi
2.20	···	Ti_5Si_3	Ti_5Si_3
2.19	···	···	TiSi
2.18	TiC	TiC	TiC
2.11	···	Ti_5Si_3	Ti_5Si_3
2.10	Ti_3O_5	···	···
1.96	···	···	TiSi
1.70	Ti_2O_3	···	···
1.54	TiC	TiC	TiC
1.51	···	Ti_5Si_3	···
1.31	TiC	TiC	TiC

Note: Here and subsequently, *d* is the interplanar distance and *E* is the laser radiation power.

TABLE II. Change in the phase composition of a Ti–Si system under laser treatment in hexane.

		E, W	
d, Å	1.5	4.0	7.0
4.28	Ti_3O_5	Ti_3O_5	Ti_3O_5
3.54	Ti_3O_5	Ti_3O_5	Ti_3O_5
3.14	Ti_3O_5	Ti_3O_5	Ti_3O_5
2.71	Ti_2O_3	···	···
2.57	Ti_2O_3	···	···
2.51	TiC	TiC	TiC
2.44	···	···	TiSi
2.20	···	Ti_5Si_3	Ti_5Si_3
2.19	···	···	TiSi
2.18	TiC	TiC	TiC
2.11	···	Ti_5Si_3	Ti_5Si_3
2.10	Ti_3O_5	···	···
1.96	···	···	TiSi
1.54	TiC	TiC	TiC
1.51	···	Ti_5Si_3	···
1.31	TiC	TiC	TiC
0.97	TiC	TiC	TiC

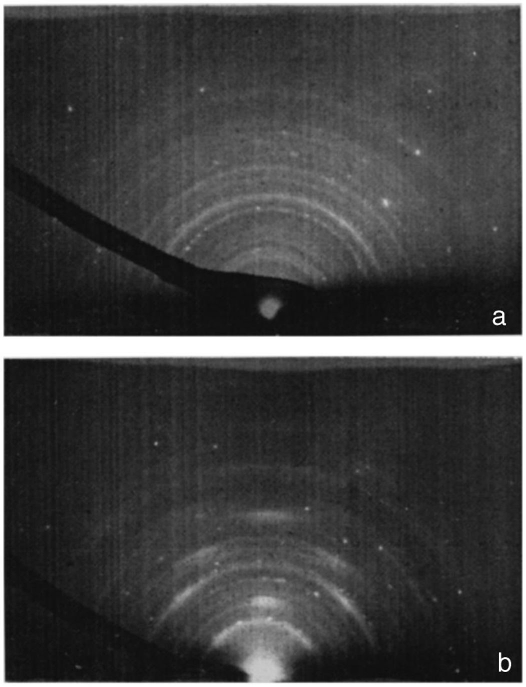


FIG. 2. Electron diffraction patterns of titanium–silicon system after laser treatment in hexane at 1.5 (a) and 7.0 W (b).

TABLE III. Change in the phase composition of a Ti–Si system in hexane after etching the film to a thickness of 20 nm.

	E, W		
d, Å	1.5	4.0	7.0
4.28	Ti_3O_5	Ti_3O_5	···
3.54	Ti_3O_5	Ti_3O_5	···
3.14	Ti_3O_5	Ti_3O_5	···
2.68	TiSi	TiSi	TiSi
2.57	···	···	Ti_2O_3
2.51	TiC	TiC	TiC
2.44	TiSi	TiSi	TiSi
2.34	···	TiSi	···
2.29	···	$TiSi_2$	$TiSi_2$
2.20	Ti_5Si_3	···	···
2.19	···	TiSi	TiSi
2.18	TiC	TiC	TiC
2.11	Ti_5Si_3	···	···
2.08	···	$TiSi_2$	$TiSi_2$
1.82	···	···	$TiSi_2$
1.70	···	···	Ti_2O_3
1.54	TiC	TiC	TiC
1.48	···	···	Ti_2O_3
1.44	TiSi	TiSi	TiSi
1.40	Ti_5Si_3	···	···
1.39	···	$TiSi_2$	$TiSi_2$
1.31	TiC	TiC	TiC

presented in Tables III and IV shows that no titanium oxides or titanium carbide exist at a depth of 40 nm. This is limited by the depth of diffusion of carbon atoms from the surrounding medium. The amount of oxygen diffusing from the SiO_2 layer into the titanium film is negligible and is insufficient to form titanium oxides. The presence of silicides indicates that silicon atoms diffuse through the SiO_2 into the titanium.

In order to study the interface of the system after laser treatment in alkanes, 50 and 60 nm surface layers were removed from the samples after treatment. The results are presented in Table V. In addition to peaks of titanium silicide with an S-54 lattice, the electron diffraction patterns also reveal peaks from SiO_2 (tridymite) and silicon. Thus, layer-by-layer etching of the surface layer revealed changes in the phase composition of a Ti–Si system under laser annealing. The phase composition changes as follows: $Ti/SiO_2/Si$ to Ti_xO_y–TiC–Ti_nSi_m/$TiSi_2$–SiO_2/Si, depending on the distance from the surface. Attention is drawn to the fact that SiO_2 crystallizes at the metal–semiconductor interface and becomes displaced inside the silicon.

TABLE IV. Change in the phase composition of a Ti–Si system after etching to 40 nm.

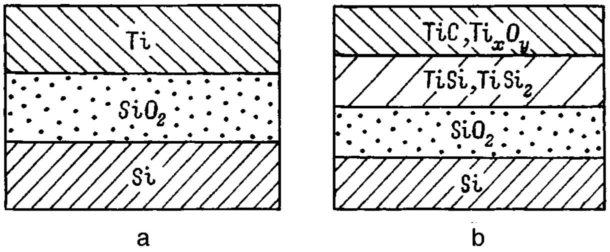
	E, W		
d, Å	1.5	4.0	7.0
4.26	SiO_2	SiO_2	SiO_2
4.08	SiO_2	SiO_2	SiO_2
3.12	Si	Si	Si
2.68	TiSi	TiSi	···
2.44	TiSi	TiSi	···
2.34	TiSi	TiSi	···
2.29	$TiSi_2$	$TiSi_2$	$TiSi_2$
2.19	$TiSi_2$	$TiSi_2$	$TiSi_2$
2.13	···	···	$TiSi_2$
2.08	$TiSi_2$	$TiSi_2$	$TiSi_2$
1.96	TiSi	TiSi	···
1.90	Si	Si	Si
1.82	$TiSi_2$	$TiSi_2$	$TiSi_2$
1.63	Si	Si	Si
1.49	$TiSi_2$	$TiSi_2$	$TiSi_2$
1.44	TiSi	···	···
1.39	···	···	$TiSi_2$
1.35	···	···	Si
1.31	$TiSi_2$	$TiSi_2$	$TiSi_2$
1.24	$TiSi_2$	$TiSi_2$	$TiSi_2$

TABLE V. Change in the phase composition of a Ti–Si system at $E=7.0$ W after etching at 50 (*A*) and 60 nm (*B*).

d, Å	*A*	*B*
4.26	SiO_2	SiO_2
4.08	SiO_2	SiO_2
3.80	SiO_2	SiO_2
3.61	···	SiO_2
3.23	···	SiO_2
3.12	Si	Si
2.96	SiO_2	SiO_2
2.48	SiO_2	SiO_2
2.29	$TiSi_2$	$TiSi_2$
2.19	$TiSi_2$	$TiSi_2$
2.13	$TiSi_2$	···
2.08	$TiSi_2$	···
2.07	···	SiO_2
1.93	···	SiO_2
1.90	Si	Si
1.63	Si	Si
1.24	Si	Si
1.10	···	Si
1.04	···	Si
0.92	···	Si

We therefore conclude that when a titanium–silicon system undergoes laser irradiation in alkanes, the surface layer of titanium interacts with the carbon, leading to the formation and growth of titanium carbide at the surface of the film, which is also accompanied by oxidation of the titanium. As a result of the silicon diffusing across the SiO_2 film and also as a result of its degradation, silicides form at the titanium–silicon dioxide interface. In consequence, the near-surface layer has a complex phase composition. The phase distribution over the depth of this layer can be seen from Fig. 3.

The results obtained by electron diffraction analysis cor-



a

b

FIG. 3. Diagram showing distribution of phase composition of titanium–silicon system over depth: a — initial sample (without treatment), b — after laser treatment at 7.0 W in alkanes.

relate with those obtained by electron spectral and chemical analysis.

Changes in the phase composition of the Ti–Si system and its interface lead to changes in the electrophysical parameters of the contact. The current–voltage characteristics were investigated using a method described in Ref. 6 which yielded the Schottky barrier height, the breakdown voltage, and the ideality factor. These investigations showed that within experimental error, the electrophysical parameters of a Ti–Si contact treated in pentane and hexane are identical. Under treatment the Schottky barrier height increases from 0.55 eV for the initial sample to 0.56 eV as a result of annealing in alkanes, and the breakdown voltage increases from 0.5 to 0.8 V. The increase in the Schottky barrier height is caused by the reactive diffusion of silicon under laser treatment and the subsequent formation of titanium silicide and disilicide. A negligible change in the Schottky barrier height can be attributed to the presence of $TiSi_2$ and was explained in detail in Ref. 1. The barrier height is also influenced by changes in the surface state density at the semiconductor under laser treatment. Redistribution of oxygen and diffusion of carbon in the surface layer and inside the system lead to an increase in the breakdown voltage. The influence of impurities in the contacts, which affect the characteristics of semiconductor devices, was described by Stricha *et al.*[7]

To sum up, by treating a Ti–Si system with $\lambda = 1.06\,\mu$m laser radiation in alkanes, we can obtain a contact with a specific phase composition. Treatment in alkanes stimulates the formation of titanium oxides, carbide, and silicides depending on the laser radiation power and the thickness of the titanium film. By exposing the system to laser treatment in pentane and hexane, we can produce a rectifying contact with specific electrophysical properties.

[1] S. P. Murarka, *Silicides for VLSI Applications* [Academic Press, New York (1983); Mir, Moscow (1986), 176 pp.].

[2] N. N. Rykalin, A. A. Uglov, and A. N. Kokora, *Laser Treatment of Materials* [in Russian], Mashinostroenie, Moscow (1975) 296 pp.

[3] *Handbook of Thin-Film Technology*, edited by L. I. Maissel and M. Glang [McGraw-Hill, New York (1970); Sovetskoe Radio, Moscow (1977), 664 pp.].

[4] A. M. Chaplanov and A. N. Shibko, Phys. Status Solidi A **120**, K37 (1990).

[5] O. B. Yatsenko, A. Ya. Tverdokhlebova, and É. A. Sadychev, Neorg. Mater. **28**, 536 (1992).

[6] E. H. Roderick, *Metal–Semiconductor Contacts* [Clarendon Press, Oxford (1978); Radio i Svyaz', Moscow (1982), 208 pp.].

[7] V. M. Strikha, E. V. Buzaneva, and I. A. Radzievskiĭ, *Schottky Barrier Semiconductor Devices* [in Russian], Sovetskoe Radio, Moscow (1974), 248 pp.

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

X-ray diffraction scattering and determination of the structural parameters of a film with a variable strain gradient

A. A. Dyshekov and Yu. P. Khapachev

Kabardino-Balkar State University, 360004 Nalchik, Russia

(Submitted February 24, 1998)
Zh. Tekh. Fiz. **69**, 67–70 (June 1999)

Characteristics of the x-ray diffraction field in structures with a variable strain gradient are analyzed using a model with an exponential profile. It is shown that the problem of reconstructing the structural parameters of the strained layer from the angular positions of the principal maximum and the oscillations is generally multivalued even when the strain varies monotonically over depth. Conditions are determined for which this problem can be solved. An analogy is identified with the results of an approach to determine the parameters of the strained layer based on using the integral characteristics of the diffraction reflection curve. © *1999 American Institute of Physics.* [S1063-7842(99)01206-4]

An analytic study of the strained surface region of a crystal by x-ray diffraction methods is essentially based on information on the wave field structure over the crystal depth in various angular intervals near the Bragg maximum.[1] In this case, the strain profile is described by a certain model which contains information on the structural parameters of the strained region: the layer thickness, the nature of the decrease over depth, and so on. A comprehensive analysis under various diffraction conditions for various possible ratios of the structural parameters can only be made for those profiles for which an accurate analytic solution of the relevant dynamic diffraction problem is known. In this context, particular importance is attached to the diffraction problem for an exponential profile

$$\varepsilon(z)=\varepsilon_0 e^{-Mz}, \tag{1}$$

where ε_0 is the strain amplitude, M is a positive value proportional to the strain gradient and determining some characteristic thickness over which the strain varies, and z is the coordinate along the normal directed into the crystal.

In general, the accurate solution of the dynamic diffraction problem for a crystal with an exponential strain profile (1) has the form[2,3]

$$E_{\mathbf{H}}(\tau)=C_1 e^{-i(\kappa+\kappa_0)\tau}F\left(1+\frac{i}{\mu}(\kappa+\kappa_0),1+\frac{i2\kappa_0}{\mu};\frac{i2\xi}{\mu}e^{-\mu\tau}\right)+C_2 e^{-i(\kappa-\kappa_0)\tau}F\times\left(1+\frac{i}{\mu}(\kappa-\kappa_0),1-\frac{i2\kappa_0}{\mu};\frac{i2\xi}{\mu}e^{-\mu\tau}\right), \tag{2}$$

where $F(a,c;x)$ is a degenerate hypergeometric function, C_1 and C_2 are arbitrary constants determined by the boundary conditions of the problem.

The following notation is introduced in Eq. (2):

$$\kappa=\left(-2\Delta\theta\sin 2\theta-\chi_0\left(1-\frac{\gamma_{\mathbf{H}}}{\gamma_0}\right)\right)\frac{\pi L}{2\gamma_{\mathbf{H}}\lambda},$$

$$\xi=-\frac{\pi L}{2\gamma_{\mathbf{H}}\lambda}4\sin^2\theta\left(\cos^2\varphi\pm\frac{1}{2}\cot\theta\sin 2\varphi\right)\varepsilon_0,$$

$$\delta_0=\left(\frac{L}{L_{\text{ext}}}\right)^2,\quad L_{\text{eta}}=\frac{\lambda\sqrt{|\gamma_{\mathbf{H}}|\gamma_0}}{\pi\eta|\chi_{\mathbf{H}}|},$$

$$\kappa_0^2=\kappa^2-\delta_0,\quad \mu=ML,$$

$\Delta\theta$ is the deviation from the exact Bragg angle θ, $\gamma_{\mathbf{H}}=-\sin(\theta\pm\varphi)$ and $\gamma_0=-\sin(\varphi\pm\theta)$ are the direction cosines of the wave vectors of the diffracted and refracted waves, φ is the angle of inclination of the diffraction plane to the crystal surface, L_{ext} is the extinction length, λ is the wavelength of the incident radiation, φ_0 is the strain amplitude, η is the polarization factor, $\tau=z/L$ is the coordinate normalized to the crystal thickness L, the "+" ("−") signs correspond to the diffraction geometry when the incident beam is at the angle $\theta-\varphi$ ($\theta+\varphi$) to the crystal surface.

We shall confine our analysis to the case of a steep strain gradient which corresponds to a strained surface region of small thickness. We take as the parameter the "effective thickness" $1/\mu$ of the strained region relative to the total crystal thickness

$$\frac{1}{\mu}=\frac{1}{ML}.$$

At this point, we must specify which quantity is assumed to be small relative to μ in the exact solution. Since we intend to use these formulas for arbitrary values of the strain (and specifically for the parameter ξ typical of the theory), this constraint will be determined only by the angular interval in which the diffraction reflection curve can be constructed. Hence, the angular region for which this approximation holds is limited by the constraints $|\kappa/\mu|\ll 1$ and $|\kappa_0/\mu|\ll 1$.

It follows from Eq. (2) that the angular values κ_0 and κ only appear in the function $F(a,c;x)$ in terms of the parameters a and c and the dependence on the strain is only relative to the argument x. This characteristic of the solution means

1063-7842/99/44(6)/3/$15.00

that for $F(a,c;x)$ we can find a so-called uniformly suitable asymptotic expansion[4] for arbitrary values of the strain and the angular range specified above.

An analysis of the solution (1) for a steep gradient shows[2,3] that the structure of the diffraction reflection curve broadly retains the form characteristic of an ideal crystal. At the same time, the presence of a strained surface region is manifest as additional modulation of the standing x-ray waves and changes in the phase ratios between them. As a result of this redistribution of the unitary wave field in the strained structure, the diffraction peak and the oscillation maxima are shifted from the angular positions for an ideal crystal.

The angular position of the main diffraction peak to within terms of the order $1/\mu^2$ is given by the following expression:[3]

$$\kappa(0)=-\rho(\delta_0)\frac{\mathrm{Si}(2\xi/\mu)}{\mu},$$

$$\rho(\delta_0)=\frac{15(2\delta_0^2+\delta_0+3)}{12\delta_0^2+7\delta_0+15}, \tag{3}$$

where $\mathrm{Si}(x)$ is the sine integral and the coefficient $\rho(\delta_0)$ determines the nature of the scattering.

It can be seen from Eq. (3) that the dependence of the angular shift on the strain amplitude is essentially nonlinear. Moreover, as a result of the oscillating behavior of the sine integral between ξ and $\Delta\kappa(0)$, generally only a multivalued correspondence exists. From Eq. (3) we obtain two characteristic limits: the kinematic limit ($\delta_0\rightarrow 0$) and the case of a thick crystal corresponding to the formal limit $\delta_0\rightarrow\infty$. For the kinematic limit, Eq. (3) gives

$$\kappa(0)_{\mathrm{kin}}=-3\frac{\mathrm{Si}(2\xi/\mu)}{\mu}, \tag{4}$$

and for a thick crystal the angular shift of the principal maximum is given by

$$\kappa(0)_{\mathrm{dyn}}=-\frac{5}{2}\frac{\mathrm{Si}(2\xi/\mu)}{\mu}. \tag{5}$$

Thus, these two cases only differ by the numerical coefficient. It follows from Eq. (3) that the angular shift of the principal maximum depends monotonically on δ_0 and the numerical difference between the two limiting cases does not exceed 20%. Therefore it is sufficient to use an average value for $\rho(\delta_0)=2.75$ in the following approximate expressions.

Similar calculations for the angular widths of the oscillation maxima yield the following result:[3]

$$\Delta\kappa(n)=\pi\left(1+\frac{1}{\mu}\mathrm{Cin}\left(\frac{2\xi}{\mu}\right)\right), \tag{6}$$

where the function

$$\mathrm{Cin}(x)=\int_0^x\frac{1-\cos y}{y}dy=-\mathrm{Ci}(x)+\ln(x)+\gamma$$

is expressed in terms of the cosine integral $\mathrm{Ci}(x)$ and $\gamma=0.577$ is the Euler constant.[5]

In formula (6) the number n of the oscillation maximum should not be too high for the expansion to hold. In deriving formula (6) we assumed that dynamic effects mainly appear only within the total diffraction reflection region so that in practice the kinematic approximation can be used for $n\geqslant 2$.

We shall analyze two variants for which the thickness of the strained surface region and the strain amplitude can be estimated for given angular shifts of the main diffraction peak and oscillations.

If $2\xi/\mu\gg 1$ (and in fact, even if $2\xi/\mu>2$), we have $\mathrm{Si}(x)\sim\pi/2$ and $\mathrm{Cin}(x)\sim\gamma+\ln(x)$ (Ref. 5). At this point, a distinguishing feature of x-ray diffraction appears in structures with a variable strain gradient: the angular position of the principal diffraction maximum is independent of the strain amplitude and is determined only by the thickness of the strained layer. A similar situation may be interpreted as the result of multiple inversion of the phase of the diffracted wave over the thickness of the strained layer with components of the initial ''perturbation'' being ''forgotten'' (strain amplitude). Consequently, this case of scattering has some analogy with a Markov process.

Thus, the criterion for this case being implemented experimentally will be that $\kappa(0)$ is constant within the error limits specified above, which is determined by the nature of the scattering (measuring $\rho(\delta_0)$) as a function of the reflection order.

Equations (3) and (6) yield the following approximate expressions for the thickness of the strained layer and the strain amplitude in the experimental units $\kappa(0)$ and $\Delta\kappa(n)$:

$$\frac{1}{\mu}=-0.23\kappa(0),\quad \xi=0.28\mu\exp\left(\left(\frac{\Delta\kappa(n)}{\pi}\right)\mu\right). \tag{7}$$

At this point, it should be noted that this expression for ξ is to a considerable extent illustrative. This is because, as given by Eq. (6), the angular width of the oscillations $\Delta\kappa(n)$ is primarily determined by the total crystal thickness. The influence of $\Delta\kappa(n)$ on the parameters of the strained surface region ξ and μ is considerably weaker here since it has the form of a small additive correction. Finally, we find that the exponential function in the expression for ξ in Eq. (7) is always very close to unity for the same μ.

The second variant corresponding to the condition $2\xi/\mu\ll 1$ is more frequently encountered in crystals with a strained layer. In cases of small values of the argument, for the functions $\mathrm{Si}(x)$ and $\mathrm{Cin}(x)$ we obtain the estimate $\mathrm{Si}(x)\sim x$ and $\mathrm{Cin}(x)\sim x^2/4$. Consequently, $\Delta\kappa(0)$ depends linearly on ξ, which should be observed experimentally. This result is fully compatible with a clear interpretation of the shift of the angular position of the principal diffraction maximum as the manifestation of some nontrivial ''diffraction averaging'' of the strain field over the thickness of the strained layer by the x-ray wave. Relations similar to (7) have the form

$$\frac{1}{\mu}=0.033\frac{\kappa^2(0)}{(\Delta\kappa(n)/\pi-1)},\quad \xi=0.18\mu^2\kappa(0). \tag{8}$$

Thus, in these limiting cases, information on the angular shifts of the principal maximum and the oscillations is quite sufficient primarily to estimate the thickness of the strained region and only as a secondary function to estimate strain amplitude. This problem can be solved uniquely (for the thickness of the strained region) in two alternative cases: under conditions where the ratio of the strain at the crystal surface to the thickness of the strained layer is small and conversely, when this ratio is large. Otherwise, additional information is required for an accurate solution of the problem.

We shall give the results of calculations of ξ and μ using formulas (7) and (8) for two cases. In order to determine the angular shifts, we constructed the theoretical diffraction reflection curve for given ξ and μ near the principal diffraction maximum using asymptotic representations of the exact solution in this particular case of a steep gradient.[2,3]

For the first case $2\xi/\mu \gg 1$ for the given values $\xi = 20$ and $\xi = 30$ for $\mu = 10$ we obtain: $\mu = 8.9$, $\xi = 5.03$, and $\mu = 9.7$, $\xi = 8.5$, respectively. Thus, the good approximation obtained for μ and the complete disparity for ξ (which differs almost fourfold) illustrates the reasoning put forward above on the analogy with Markov processes—loss of information on ξ.

For the second case $2\xi/\mu \ll 1$ the calculations yield $\xi = 1.2$, $\mu = 8.5$ for the true values $\xi = 1$ and $\mu = 8$, and $\xi = 2.02$, $\mu = 6.7$ for $\xi = 2$ and $\mu = 8$. Here no problems arise with loss of information and this accuracy for the approximate expressions can be considered to be quite satisfactory.

These results for the specific problem of diffraction in a crystal with an exponential strain profile can lead to the following generalization. The relations put forward above essentially establish qualitative criteria for a particular degree of solubility for the problem of determining the strain amplitude and the thickness of the strained region from x-ray diffraction data. This statement applies equally to analytic and numerical methods since the latter are based substantially on the angular positions and widths of the diffraction maxima.[6]

It is interesting to note that similar conclusions can also be drawn when the diffraction problem is considered in the so-called semikinematic approximation when the strained surface layer of the crystal is assumed to scatter kinematically and the substrate scatters like an ideal dynamic crystal. When solving this problem, Afanasev *et al.*[7] obtained a relationship linking the Fourier transform of the diffraction reflection curve far from the fundamental maximum with some effective thickness L_{eff}. In the notation used here this formula is given by

$$L_{\text{eff}} = \int_0^{1/\mu} \left(1 - \cos\left(-\frac{\xi}{\mu} e^{-\mu\iota}\right)\right) d\iota$$
$$= \frac{1}{\mu}\left(1 - \mathrm{Ci}\left(\frac{\xi}{\mu}\right) + \mathrm{Ci}\left(\frac{\xi}{\mu} e^{-1}\right)\right). \quad (9)$$

For the first of the characteristic limits specified above we obtain from formula (9)

$$L_{\text{eff}} = \frac{\xi^2}{4\mu^3}, \quad (10)$$

i.e., in cases of incomplete phase inversion over the thickness of the strained layer L_{eff} depends on the parameters of this layer and these can therefore be reconstructed from the x-ray data.

However, the estimate (9) for $2\xi/\mu \gg 1$ (multiple phase inversion) gives

$$L_{\text{eff}} = \frac{1}{\mu}\left(1 - \frac{\sin(\xi/\mu)}{\xi/\mu}\right). \quad (11)$$

Since the second term in Eq. (11) for large values of the argument oscillates near zero and is much less than unity, the effective thickness is merely reduced to the depth of the strained layer and does not depend on the strain amplitude. In this case we are therefore dealing with a loss of information on the parameters of the strained layer.

It is extremely important that this result can also be generalized to other strain profiles which decrease over depth, by estimating the integral (9). However, the important constraint remains that the strain must vary monotonically over depth. Mathematically this constraint reduces to the absence of steady-state points in the phase of the cosine integral (9) over the thickness of the strained layer. This in fact implies that a whole range of constant-sign strain profiles which decrease over depth at least no more weakly than exponentially can be included here.

To sum up, these results obtained for a crystal with an exponential strain profile suggest that these x-ray diffraction laws apply to all structures with a variable gradient where the strain varies monotonically over depth. An analogy is also obtained when integral characteristics such as the Fourier transforms of the diffraction reflection curves and the angular positions of the diffraction maxima are used to identify the structure of the strained layer.

This work was supported by the Russian Fund for Fundamental Research, Project No. 98-02-16151.

[1] A. M. Afanas'ev and V. G. Kon, Zh. Éksp. Teor. Fiz. **74**, 300 (1978) [Sov. Phys. JETP **47**, 154 (1978)].
[2] A. A. Dyshekov and Yu. P. Khapachev, Poverkhnost' No. 3, (1997).
[3] A. A. Dyshekov and Yu. P. Khapachev, Poverkhnost' No. 3, (1997).
[4] A.-H. Nayfeh, *Perturbation Methods* [Wiley, New York (1973); Mir, Moscow (1976), 455 pp.].
[5] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun [Dover, New York (1965); Nauka, Moscow (1979), 832 pp.].
[6] V. A. Bushuev, Yu. P. Khapachev, and V. V. Lider, Pis'ma Zh. Tekh. Fiz. **19**(23), 74 (1993) [Tech. Phys. Lett. **19**, 766 (1993)].
[7] A. M. Afanasev, M. V. Kovalchuk, E. K. Kovev, and V. G. Kohn, Phys. Status Solidi A **42**, 415 (1977).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Nonlinear dynamics of the space charge in a photorefractive crystal under pulsed photoinjection

Yu. I. Kuz'min

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

(Submitted March 11, 1999)
Zh. Tekh. Fiz. **69**, 71–79 (June 1999)

An analysis is made of the nonlinear relaxation of the charge formed in a photoconducting material as a result of the absorption of a light pulse. Calculations are made of the distribution profiles of the excess carriers and the density of the photoinduced charge in various transport regimes. The influence of contact phenomena on the dynamics of the space charge is examined and solutions are obtained for boundary conditions corresponding to a blocking contact with an illuminated electrode and optical breakdown of the contact. An investigation is made of the evolution of the integral pulse response under transverse electrooptic modulation. The results are suitable for analyzing two geometric recording systems for entering information into a space–time light modulator with arbitrary absorption and for recording a holographic grating in the limit of low absorption coefficients. © *1999 American Institute of Physics.* [S1063-7842(99)01306-9]

INTRODUCTION

Studies of nonlinear space-charge relaxation processes in photorefractive materials are topical for two main reasons. First, photorefractive materials are opening up new prospects for developing active elements in optical information processing and storage systems (including holographic memory).[1–5] Second, photoinjection processes and transport of nonequilibrium photoinduced charges in these materials lead to the formation of photorefractive waves, which are of separate scientific interest.[6–12]

Here, a two-dimensional nonsteady-state problem is solved for the nonlinear relaxation of the space charge formed in a photorefractive crystal under pulsed illumination. A detailed analysis is made of a variant corresponding to the storage of information in a space–time light modulator.[1,12,13] However, the results are suitable for analyzing the pulsed recording of holographic gratings[1,5,14] in the limit of low absorption coefficients. The difference between these variants is determined by the form of the spatial modulation function of the photogeneration in accordance with the system for storing information in the optical memory. The case of an arbitrary absorption coefficient is considered for a space–time light modulator. Allowance is also made for the depletion of impurity centers during photogeneration and the influence of the injection current on the dynamics of the space charge is considered. Two sets of boundary conditions are analyzed, corresponding to a blocking contact with an illuminated electrode and optical breakdown of the contact.

FORMULATION OF THE PROBLEM AND MATHEMATICAL MODEL

We analyze an infinite single-crystal wafer of thickness L to which we apply a static electric field created by a voltage source U (Fig. 1). The x axis is directed along the crystal thickness and the y axis is directed along the surface. The dark conductivity of the material is assumed to be zero. At time $t=0$ the crystal undergoes pulsed illumination. The optical intensity distribution has the form $I(\mathbf{r},t)=W\delta(t)R(\mathbf{r})$, where $\mathbf{r}\equiv x\mathbf{e}_x+y\mathbf{e}_y$, W is the exposure, $\delta(t)$ is the Dirac delta function, and $R(\mathbf{r})=X(x)Y(y)$ is the spatial modulation function.

Figure 1 shows the direction of light for the two most commonly used experimental systems. In the first case, when information is coupled into a space–time modulator, light is incident on the crystal along the x axis (Fig. 1a) and the modulation function $X(x)$ is determined by the Bouguer–Lambert law $X(x)=\exp(-\alpha x)$, where α is the absorption coefficient. In the second case, when a holographic grating is recorded (Fig. 1b), light propagates along the y axis and for $\alpha L \ll 1$ the intensity is modulated along the x axis by the

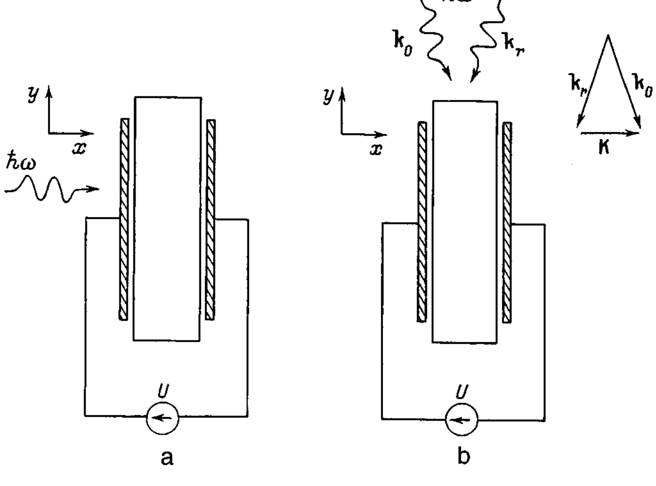


FIG. 1. Schematic of information recording process: a — input of information to space–time light modulator, b — hologram recording, $\mathbf{k}_0$ and $\mathbf{k}_r$ are the wave vectors of the object and reference waves, respectively.

function $X(x)=1+\xi\cos(Kx)$, where K is the wave vector of the recorded grating and ξ is the modulation index ($0<\xi<1$).

The crystal contains two types of traps: shallow traps in thermal equilibrium with the conduction band and deep impurity centers whose photoionization leads to the appearance of free carriers migrating in the external field. The carrier lifetime in the deep traps significantly exceeds the transit time. The drift of the nonequilibrium carriers is limited by the space charge defined by the carriers themselves. Consequently, the transport process is nonlinear: Ohm's law is not obeyed since the charge migrates in a self-consistent field.

The influence of shallow traps on the transport process is taken into account in the fast recapture approximation.[15] We assume that $\tau_t \ll \tau_f$, where τ_t is the trapping time (the average carrier lifetime in an extended state), τ_f is the free time (the average lifetime in a trapping level). In this case, the equation describing the trapping and release of carriers from shallow traps $\partial n_t/\partial t = n_f/\tau_t - n_t/\tau_f$, reduces to $\tau_f/\tau_t = n_t/n_f$, where n_f and n_t are the concentrations of free and trapped carriers, respectively. This means that instead of the transport of free carriers having the mobility μ_0, which determines the conductivity for the extended states, we can analyze the motion of quasifree carriers having the concentration $n=n_f+n_t$ and the effective drift mobility $\mu=\mu_0 n_f/(n_f+n_t)$.

The relaxation of the photoinduced charge is described by the following system of equations written in dimensionless variables:

$$\nabla\cdot(n(\mathbf{r},t)\nabla\varphi(\mathbf{r},t))-\frac{\partial}{\partial t}n(\mathbf{r},t)=0, \tag{1}$$

$$\nabla^2\varphi(\mathbf{r},t)=p(\mathbf{r},t)-n(\mathbf{r},t), \tag{2}$$

$$\frac{\partial}{\partial t}p(\mathbf{r},t)=\frac{\alpha}{N}I(\mathbf{r},t)(N-p(\mathbf{r},t)), \tag{3}$$

where (1) is the equation of continuity, (2) is the Poisson equation, and (3) describes the ionization of the impurity centers. Here φ is the potential and N and p are the concentrations of impurity centers and ionized impurities, respectively. On changing to dimensionless variables, we use the following normalization:

$$r\equiv\frac{r'}{L},\quad t\equiv\frac{t'}{t_T},\quad \varphi\equiv\frac{\varphi'}{U},\quad \alpha\equiv\alpha' L,$$

$$(n,N,p)\equiv(n',N',p')\frac{4\pi qL^2}{\varepsilon U},\quad W\equiv W'\frac{4\pi qLQ}{\hbar\omega\varepsilon U},$$

where the corresponding dimensional variables are denoted by the prime, $t_T\equiv L^2/(\mu U)$ is the carrier transit time in an unperturbed field, q is the carrier charge, Q is the photoionization quantum yield of an impurity center, ω is the light frequency, and ε is the permittivity of the crystal for polarization with relaxation times considerably shorter than the transit time.

The initial conditions correspond to the absence of any ionized centers before illumination, and local electroneutrality at the end of illumination

$$p(\mathbf{r},0-0)=0,\quad n(\mathbf{r},0+0)=p(\mathbf{r},0+0). \tag{4}$$

The boundary conditions are determined by the carrier concentration at the illuminated surface according to the type of contact, the applied electric field, and the type of radiation intensity distribution over the y axis

$$n(0,y,t)=n(0,y,0+0), \tag{5}$$

$$\varphi(0,y,t)=1,\quad \varphi(1,y,t)=0, \tag{6}$$

$$\frac{\partial}{\partial y}\varphi(x,0,t)=0, \tag{7}$$

$$\frac{\partial}{\partial y}\varphi(x,y_L,t)=0. \tag{8}$$

Condition (8) corresponds either to translational symmetry of the potential for the periodic function $Y(y)=Y(y+y_L)$, or screening of the field generated by the space charge, by the mirror image charges for the damped function $Y(y)$. In this last case, condition (8) holds for $y_L \gg L$. The integral of the photoionization equation (4) has the form

$$p(\mathbf{r},t)=N(1-\exp(-F(\mathbf{r})))H(t), \tag{9}$$

where $F(\mathbf{r})\equiv(\alpha W/N)R(\mathbf{r})$, $H(t)$ is the Heaviside function

$$H(t)\equiv\begin{cases}0 & t<0,\\ 1 & t>0.\end{cases}$$

The equation of continuity (1) may be transformed to a convenient form for the subsequent discretization

$$\frac{\partial}{\partial t}n(\mathbf{r},t)-\nabla\varphi(\mathbf{r},t)\cdot\nabla n(\mathbf{r},t)+n(\mathbf{r},t)(n(\mathbf{r},t)-p(\mathbf{r},t))=0. \tag{10}$$

Thus, Eqs. (2), (9), and (10) form a closed system to determine the carrier concentration $n(\mathbf{r},t)$ and the potential $\varphi(\mathbf{r},t)$.

BRIEF DESCRIPTION OF THE FINITE DIFFERENCE ALGORITHM

The equations are discretized using a three-dimensional mesh whose nodes are defined by the coordinates

$$x_i=i\Delta x,\quad y_j=j\Delta y,\quad t_k=k\Delta t, \tag{11}$$

where Δx, Δy, and Δt are the mesh steps along the appropriate coordinates, $i=0,1,\dots,i_{\max}$; $j=0,1,\dots,j_{\max}$; $k=0,1,\dots,k_{\max}$.

Applying an integrointerpolation method to discretize the nonlinear equation of continuity (10) in the bulk of the crystal yields the following iteration difference equations written for half-integer points of the initial mesh (11):

$$\hat{D}_t n_{\mathbf{m}}^{(l+1)}-\hat{D}_x\varphi_{\mathbf{m}}^{(l)}\cdot\hat{D}_x n_{\mathbf{m}}^{(l+1)}-\hat{D}_y\varphi_{\mathbf{m}}^{(l)}\cdot\hat{D}_y n_{\mathbf{m}}^{(l+1)} +n_{\mathbf{m}}^{(l+1)}(n_{\mathbf{m}}^{(l)}-p_{\mathbf{m}})=0, \tag{12}$$

where l is the number of the layer-by-layer iteration for fixed k and all values with the multiindex $\mathbf{m}=\|i+1/2,\ j+1/2,\ k+1/2\|$ corresponds to the center of a mesh cell (11) (Fig. 2a).

The difference derivatives have the form

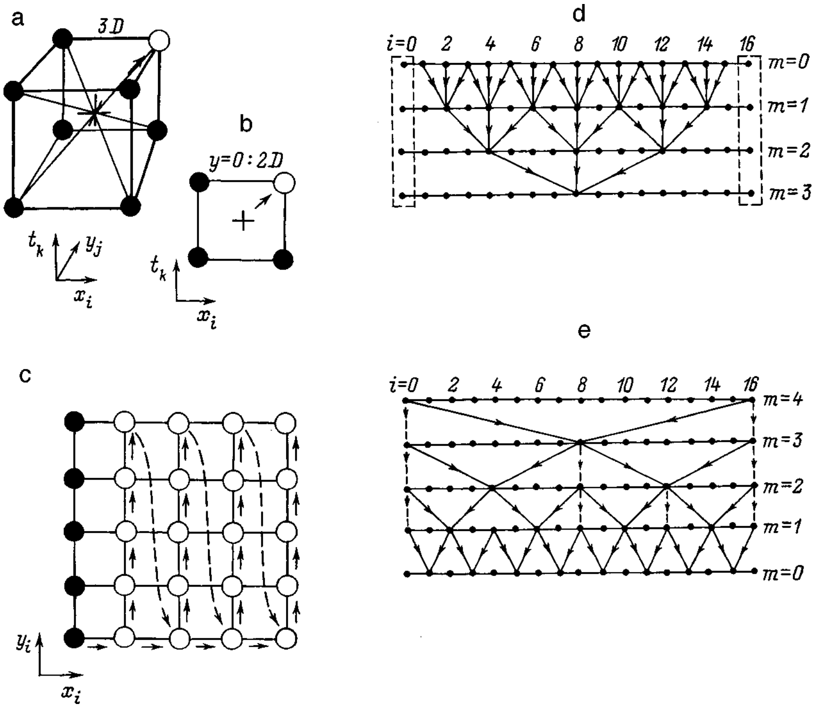


FIG. 2. Construction of a finite-difference algorithm: patterns of difference schemes (15) (a) and (16) (b); sequence of calculations of the carrier concentrations at the $(k+1)$ mesh layer at fixed time (c): direct run (d) and inverse formulation (e) of cyclic reduction for solving the Poisson equation.

$$\hat{D}_t n_{\mathbf{m}} \equiv \frac{1}{4\Delta t}(n^{k+1}_{i+1,j+1}+n^{k+1}_{i+1,j}+n^{k+1}_{i,j+1}+n^{k+1}_{i,j}-n^{k}_{i+1,j+1} -n^{k}_{i+1,j}-n^{k}_{i,j+1}-n^{k}_{ij}),$$

$$\hat{D}_x n_{\mathbf{m}} \equiv \frac{1}{4\Delta x}(n^{k+1}_{i+1,j+1}+n^{k}_{i+1,j+1}+n^{k+1}_{i+1,j}+n^{k}_{i+1,j} -n^{k+1}_{i,j+1}-n^{k}_{i,j+1}-n^{k+1}_{ij}-n^{k}_{ij}),$$

$$\hat{D}_y n_{\mathbf{m}} \equiv \frac{1}{4\Delta y}(n^{k+1}_{i+1,j+1}+n^{k}_{i+1,j+1}+n^{k+1}_{i,j+1}+n^{k}_{i,j+1} -n^{k+1}_{i+1,j}-n^{k}_{i+1,j}-n^{k+1}_{ij}-n^{k}_{ij}).$$

Difference differentiation of the potential is performed similarly. The values of n and p are averaged over all points in an elementary cell of the mesh

$$n_{\mathbf{m}}=\frac{1}{8}\sum_{s=i}^{i+1}\sum_{r=j}^{j+1}\sum_{q=k}^{k+1} n^{q}_{sr}, \quad p_{\mathbf{m}}=\frac{1}{8}\sum_{s=i}^{i+1}\sum_{r=j}^{j+1}\sum_{q=k}^{k+1} p^{q}_{sr}.$$

Discretization of the equation of continuity (10) in the plane $y=0$ where condition (7) holds is performed separately. The corresponding difference equation written for half-integer points of the two-dimensional mesh is obtained from the initial three-dimensional mesh (11) for $j=0$

$$\hat{\partial}_t n^{(l+1)}_{\boldsymbol{\mu}} - \hat{\partial}_x \varphi^{(l)}_{\boldsymbol{\mu}} \cdot \hat{\partial}_x n^{(l+1)}_{\boldsymbol{\mu}} + n^{(l+1)}_{\boldsymbol{\mu}}(n^{(l)}_{\boldsymbol{\mu}}-p_{\boldsymbol{\mu}})=0, \tag{13}$$

where l is the iteration number, $\boldsymbol{\mu}\equiv\|i+1/2,\,0,\,k+1/2\|$, $n_{\boldsymbol{\mu}}=\frac{1}{4}\sum_{s=i}^{i+1}\sum_{q=k}^{k+1}n^{q}_{s,0}$, $p_{\boldsymbol{\mu}}=\frac{1}{4}\sum_{s=i}^{i+1}\sum_{q=k}^{k+1}p^{q}_{s,0}$,

$$\hat{\partial}_t n_{\boldsymbol{\mu}} \equiv \frac{1}{2\Delta t}(n^{k+1}_{i+1,0}+n^{k+1}_{i,0}-n^{k}_{i+1,0}-n^{k}_{i,0}),$$

$$\hat{\partial}_x n_{\boldsymbol{\mu}} \equiv \frac{1}{2\Delta x}(n^{k+1}_{i+1,0}+n^{k}_{i+1,0}-n^{k+1}_{i,0}-n^{k}_{i,0}).$$

The boundary conditions are obtained directly from Eqs. (4) and (5)

$$n^{0}_{ij}=p^{0}_{ij}, \quad n^{k}_{0,j}=n^{0}_{0,j}. \tag{14}$$

Using the difference equation (12) we can derive the following explicit two-layer difference scheme to calculate n at the $(l+1)$ iteration:

$$n^{k+1}_{i+1,j+1}(1+u^{+})=n^{k+1}_{i,j+1}(-1+\nu^{-})+n^{k+1}_{i+1,j}(-1-\nu^{+}) +n^{k+1}_{ij}(-1+u^{-})+n^{k}_{i+1,j+1} \times(1-u^{+})+n^{k}_{i,j+1}(1+\nu^{-})+n^{k}_{i+1,j} \times(1-\nu^{+})+n^{k}_{ij}(1+u^{-}), \tag{15}$$

where

$$u^{\pm}\equiv\frac{\Delta t}{\Delta x}\left(-\hat{D}_x\varphi^{(l)}_{\mathbf{m}}-\frac{\Delta x}{\Delta y}\hat{D}_y\varphi^{(l)}_{\mathbf{m}}\pm\frac{\Delta x}{2}(n^{(l)}_{\mathbf{m}}-p_{\mathbf{m}})\right);$$

$$\nu^{\pm}\equiv\frac{\Delta t}{\Delta x}\left(-\hat{D}_x\varphi^{(l)}_{\mathbf{m}}+\frac{\Delta x}{\Delta y}\hat{D}_y\varphi^{(l)}_{\mathbf{m}}\pm\frac{\Delta x}{2}(n^{(l)}_{\mathbf{m}}-p_{\mathbf{m}})\right).$$

Figure 2a shows the pattern of the difference scheme (15). In order to trigger the scheme (15), we require values of the carrier concentration $n^{k}_{0,j}$ at the crystal surface at $x=0$ defined by the second boundary condition (14) and also values of the concentration $n^{k+1}_{i+1,0}$ in the plane $y=0$, calculated using a difference scheme obtained from Eq. (13):

$$n^{k+1}_{i+1,0}(1+a^{+})=n^{k+1}_{i,0}(-1+a^{-})+n^{k}_{i+1,0}(1-a^{+}) +n^{k}_{i,0}(1+a^{-}), \tag{16}$$

where

$$a^{\pm}\equiv\frac{\Delta t}{\Delta x}\left(-\hat{\partial}_x\varphi^{(l)}_{\boldsymbol{\mu}}\pm\frac{\Delta x}{2}(n^{(l)}_{\boldsymbol{\mu}}-p_{\boldsymbol{\mu}})\right).$$

The pattern of the difference scheme (16) is shown in Fig. 2b. The initial values n_{ij}^0 required to trigger the schemes (15) and (16) are calculated from the first boundary condition (14) and the integral of the photoionization equation (9). The calculations are performed as follows: the carrier concentration n is calculated at the $(k+1)$th layer of the mesh (11) in the unperturbed field initially in the plane $y=0$ using scheme (16) and then using scheme (15) for all points in the bulk of the crystal having the same i value on the $(k+1)$th layer. The concentration is then calculated for the next value of i, first for $j=0$ using the scheme (16) and then for $j\neq 0$ using the scheme (15), and so on. The sequence of calculations is illustrated in Fig. 2c.

After finding the carrier concentration at the $(k+1)$th layer we calculate the field created by the resultant space charge, by solving the Poisson equation (2) in a rectangle on whose opposite sides Dirichlet conditions (6) and Neumann conditions (7), (8) are set. Then the carrier concentration is calculated again, but in the perturbed field. If the values obtained differ from the previous ones, the iterations are repeated until convergence is achieved.

We use a cyclic reduction method to solve the Poisson equation on a mesh defined for a fixed time in the rectangle $\{0\leqslant x\leqslant 1, 0\leqslant y\leqslant y_L\}$.

For the potential $V\equiv\varphi-1+x$ of the field created by the space charge with the volume density $f\equiv n-p$, the boundary conditions (6) will be homogeneous. The difference problem for the Poisson equation $\nabla^2 V=-f$ may be written in the form

$$C\mathbf{V}=\mathbf{F},$$

where

$$C\equiv\left\|\begin{array}{cccccc} E & 0 & 0 & 0 & \ldots & 0 \\ -E & C & -E & 0 & \ldots & 0 \\ 0 & -E & C & -E & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & -E & C & -E \\ 0 & \ldots & 0 & 0 & 0 & E \end{array}\right\|,$$

$$\mathbf{V}\equiv\left\|\begin{array}{c} \mathbf{V}_0 \\ \mathbf{V}_1 \\ \ldots \\ \mathbf{V}_{i_{\max}} \end{array}\right\|, \quad \mathbf{F}\equiv\left\|\begin{array}{c} \mathbf{F}_0 \\ \mathbf{F}_1 \\ \ldots \\ \mathbf{F}_{j_{\max}} \end{array}\right\|,$$

$$\mathbf{V}_i\equiv\|V_{i,0}V_{i,1}\ldots V_{ij_{\max}}\|^T, \quad i=0,1,\ldots,i_{\max},$$

$$\mathbf{F}_i\equiv\begin{cases} (\Delta x)^2\|f_{i,0}f_{i,1}\ldots f_{ij_{\max}}\|^T, & i=1,2,\ldots,i_{\max}-1, \\ \|0\ \ 0\ \ \ldots\ \ 0\|^T, & i=0,i_{\max}, \end{cases}$$

E is a unit matrix, 0 is the null matrix of dimensions $(j_{\max}+1)^2$, and the superscript T denotes transposition, the matrix elements C have the form

$$C_{rs}=2(1+\beta)\delta_{rs}-\beta(\delta_{r,s+1}+\delta_{r+1,s}+\delta_{r1}\delta_{2s} + \delta_{r,j_{\max}+1}\delta_{j_{\max},s}),$$

where $\beta\equiv(\Delta x/\Delta y)^2$, δ_{rs} is the Kronecker delta, and $r,s=1,2,\ldots,j_{\max}+1$.

Thus, the solution of the Poisson equation reduces to inversion of the partitioned matrix C for which we use the following variant of the cyclic reduction method.

The direct run involves calculating the vectors $\mathbf{P}_i^{(m)}$ and $\mathbf{Q}_i^{(m)}$

$$\begin{aligned} \mathbf{P}_i^{(m)}&=\mathbf{P}_i^{(m-1)}+(C^{(m-1)})^{-1}(\mathbf{Q}_i^{(m-1)}+\mathbf{P}_{i-M}^{(m-1)}+\mathbf{P}_{i+M}^{(m-1)}), \\ \mathbf{Q}_i^{(m)}&=2\mathbf{P}_i^{(m-1)}+Q_{i-M}^{(m-1)}+\mathbf{Q}_{i+M}^{(m-1)}, \\ \mathbf{Q}_i^{(0)}&=\mathbf{F}_i, \quad \mathbf{P}_i^{(0)}=0, \end{aligned} \tag{17}$$

m is the reduction number, $m=1,2,\ldots,n$; $M\equiv 2^{m-1}$, $i=2^m, 2\times 2^m, 3\times 2^m, \ldots, i_{\max}-2^m$.

The inverse formulation involves reconstructing the unknowns

$$\mathbf{V}_i=\mathbf{P}_i^{(m-1)}+(C^{(m-1)})^{-1}(\mathbf{Q}_i^{(m-1)}+\mathbf{V}_{i-M}+\mathbf{V}_{i+M}), \tag{18}$$

$m=n+1, n, n-1,\ldots,1$; $M\equiv 2^{m-1}$, $i=M, 3M, 5M,\ldots, i_{\max}-M$.

The constraint on the number of mesh points in the direction of reduction (along the x axis) is $i_{\max}=2^{n+1}$, where n is the number of reduction steps.

Inversion of the matrices $C^{(m-1)}$ involves solving the vector equations (17) and (18) by successively inverting the matrix factors in the expansion

$$C^{(m-1)}=\prod_{l=1}^{M} C_{l,m-1}, \text{ where } C_{l,m-1}=C-2E\cos\frac{(2l-1)\pi}{2^m}.$$

The matrices $C_{l,m-1}$, like the initial matrix C, are tridiagonal with diagonal predominance and are inverted using a right tridiagonal inversion.

The calculation algorithm using a cyclic reduction for $i_{\max}=16$ and $n=3$ is shown schematically in Figs. 2d and 2e.

RESULTS AND DISCUSSION

The space charge dynamics depend fundamentally on the properties of the contact between the illuminated electrode and the photorefractive crystal. For a blocking contact the boundary condition (5) has the form $n(0,y,t)=0$ both before and after the action of light. In this case, nonequilibrium carriers form in the bulk of the crystal exclusively as a result of photoinjection and none are transported from the electrode. Another case is also possible when, under the action of light, the contact is converted from blocking to injecting and undergoes optical breakdown. In this case, the optical pulse causes an abrupt increase in the carrier concentration at the irradiated contact from zero to a finite value determined by the integral of the photoionization equation (9) and the boundary conditions (4) and (5) $n(0,y,t)=N[1-\exp(-F(0,y))]H(t)$.

Thus, after photoexcited carriers have formed, excess carriers will be continuously injected into the crystal from the broken-down contact.

Figure 3 shows a schematic band diagram of the contact under optical breakdown. For simplicity the surface states are not shown. It is assumed that before illumination (Fig.

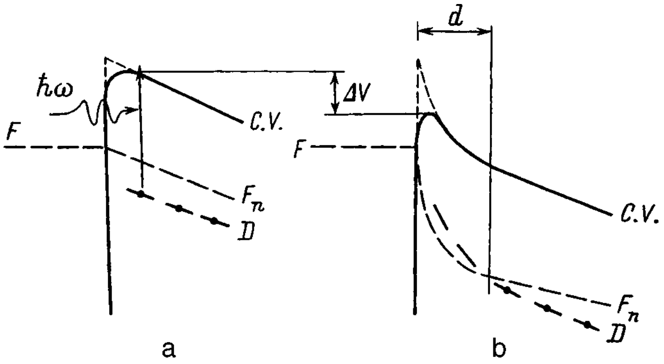


FIG. 3. Band diagram of contact before (a) and after (b) exposure to light pulse during optical breakdown: ΔV — reduction in the barrier height, d — depth of depletion zone, F — Fermi level, F_n — Fermi quasilevel, and D — impurity levels. The bottom of the conduction band (*C.V*) neglecting the potential of the mirror image forces is shown by the dashed curve.

3a) the contact is neutral, i.e., the flat band condition is satisfied. Prior to illumination an external bias is applied to the contact and all the transition processes associated with charge transfer at capacitances have time to be completed. The potential barrier impeding carrier transport (to be specific, electrons) from the cathode is assumed to be sufficiently large so that the contact can be considered to be blocking prior to illumination. Under the action of the light pulse, the impurity centers undergo photoionization. As a result of the external bias, the photoexcited carriers leave the contact region and residual ionized impurities create a space charge field. The formation of a depleted region has features in common with the similar process observed during the formation of a Schottky barrier but is distinguished by optical rather than thermal ionization of impurities and takes place at a constantly applied external bias. The combined effect of the space charge field in the depleted region, the potential of the mirror image forces, and the external electric field cause a reduction in the barrier height (similar to the Schottky effect). Consequently, after the pulsed illumination the barrier will no longer impede the injection of carriers from the contact into the bulk of the crystal (Fig. 3b).

We shall now analyze the distributions of the carrier concentration n and the space charge density f obtained for various boundary conditions. The calculations were made for spatial modulation functions given by

$$X(x)=\exp(-\alpha x),$$

$$Y(y)=\begin{cases}\dfrac{1}{2}\left(1+\cos\dfrac{2\pi y}{b}\right) & -\dfrac{b}{2}<y<\dfrac{b}{2},\\[2ex] 0 & y<-\dfrac{b}{2},\quad y>\dfrac{b}{2},\end{cases} \tag{19}$$

where b is the width of the irradiated region.

This modulation corresponds to exposure of the band positioned at the surface of the crystal perpendicular to the x and y axes when information is coupled into a space–time light modulator (Fig. 1a). The calculations were made for equal mesh steps along the spatial coordinates ($\Delta x=\Delta y$) with a relative accuracy of 10^{-2} for the carrier concentration and 0.1 for the transverse field. The time step $\Delta t=10^{-2}$ ensured that the Courant condition was satisfied. The results discussed subsequently were obtained for a mesh of dimensions $i_{\max}=64$, $j_{\max}=64$, and $k_{\max}=120$ for $\alpha=1$, $b=1$, and $N=10^6$.

We shall initially consider the case of optical breakdown of the contact. Figure 4 shows contours of the carrier concentration and the charge density at time $t=0.6$ for two values of the exposure: $W=0.1$ and 10. The calculated parameter is the product of the contour number and the quantization step. Figure 4a shows the linear motion of a packet of free carriers: at low exposures the resulting space charge field does not distort the external field and the dis-

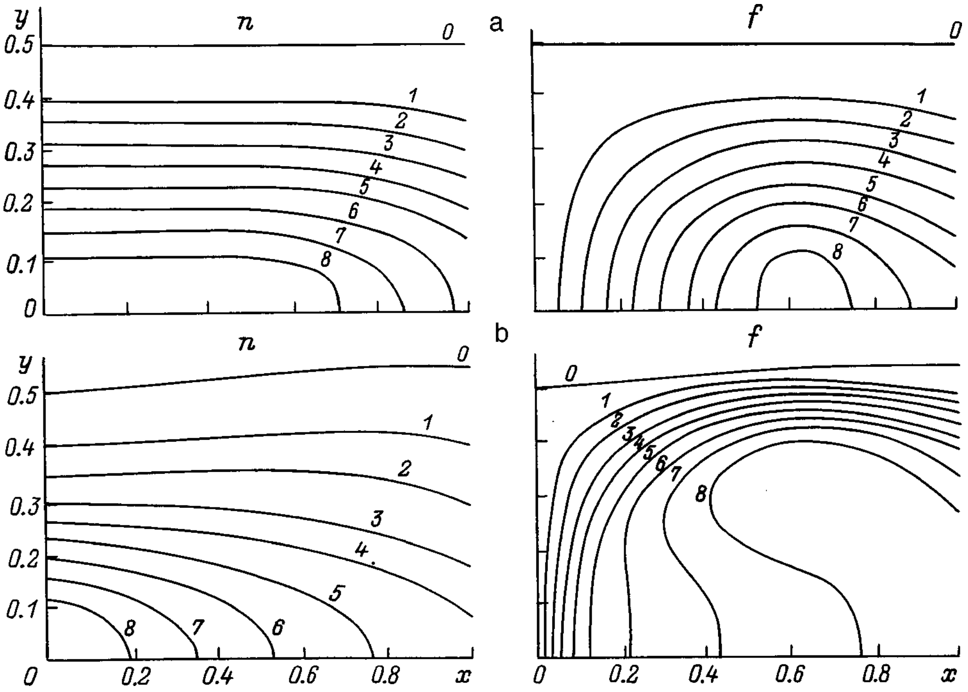


FIG. 4. Distribution profiles of excess carrier concentration (n) and space charge density (f) during optical breakdown of contact: a — linear regime (exposure $W=0.1$, quantization step $h_n=1.11\times10^{-2}$, $h_f=4.79\times10^{-3}$); b — nonlinear regime (exposure $W=10$, quantization steps $h_n=1.11$, $h_f=0.114$).

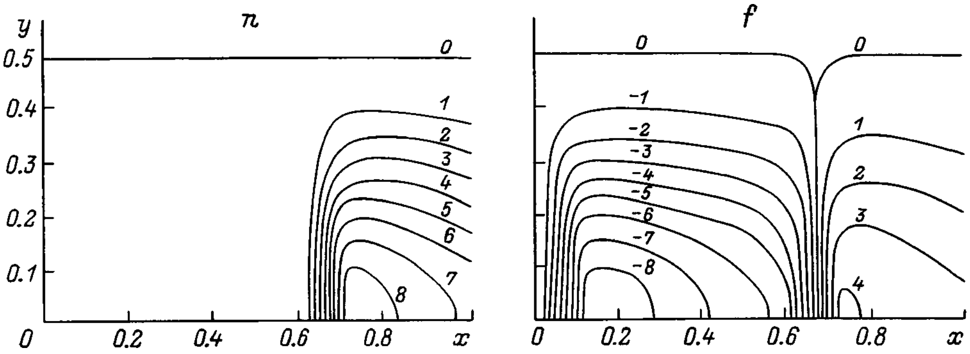


FIG. 5. Distribution profiles of space carrier concentration (n) and excess charge density (f) for blocking contact in linear transport regime (quantization steps $h_n = 9.73\times10^{-3}$, $h_f = 9.47\times10^{-3}$).

tance covered by the leading edge of the carriers increases linearly with time. The nonuniform spatial distribution of the ionized impurity centers leads to the formation of a maximum space charge density which propagates with the leading section of the carrier packet. Excess carriers are continuously injected from the contact broken down by the optical pulse and these completely compensate the charge of the ionized impurities. Thus, the space charge density has the same sign over the entire volume of the crystal. After passage of the photoexcited carriers, a current of excess carriers injected from the contact will flow through the crystal. The density of this current is modulated over the crystal surface by the function $Y(y)$ according to the intensity distribution under illumination. Therefore the recorded information is stored as long as the external bias is switched on. Since the optical breakdown of the contact is a reversible effect (provided that electrical breakdown of the contact region has not occurred as a result of summation of the external field and the space charge field), the devices returns to the initial state when the external bias is switched off.

As the exposure is increased, the contact goes over to a nonlinear transport regime as a result of confinement by the space charge. The resulting space charge of the excess carriers slows the motion of the packet and pushes it sideways. This transverse carrier drift leads to the formation of lateral ''tongues'' in the charge density distribution (Fig. 4b). Unlike the linear regime (Fig. 4a), under higher exposure the initial carrier concentration profile, defined by the spatial modulation of the radiation, becomes distorted by the space charge. At the same time, the sign of the space charge density remains constant over the entire volume of the crystal, as in the linear regime.

We shall now consider the more interesting case from the practical point of view when the illuminated contact remains blocking after the propagation of the light pulse. Results of calculating the carrier concentration and space charge density for this case are plotted in Fig. 5 which gives contours of the distributions formed at time $t = 0.6$ for the exposure $W = 0.1$, which corresponds to a linear transport regime. At $x \approx 0.6$ we can clearly identify a front which bounds the moving packet of excess carriers. This is followed by space charge of opposite sign generated by the ionized impurity centers. Since no carriers are injected from the contact, the charge of the ionized impurities remains uncompensated and the space charge density changes its sign at the bounding carrier front. After the packet of photoexcited carriers has left the bulk of the crystal under the action of the external field (at $t > 1$), the sign of the space charge density remains constant. The space charge of the ionized impurity centers creates an electric field which forms a recorded image in the crystal as a result of the electrooptic effect. It should be noted that this field and thus the written information are conserved during the dielectric relaxation time and after the external bias has been switched off.

An important parameter characterizing the possibility of using a photorefractive material for optical storage of information is the diffraction efficiency during readout. For the selected geometric system (Fig. 1a) the diffraction efficiency is determined by the integral pulse response under transverse electrooptic modulation

$$\Phi(y,t) \equiv -\int_0^1 dx \frac{\partial}{\partial x}\varphi(x,y,t).$$

For a given spatial modulation function $Y(y)$ of the type (19) the pulse response Φ characterizes the information writing efficiency when one line is input to a space–time modulator. Figure 6 shows the evolution of the pulse response calculated for $W = 0.2$, $\alpha = 0.5$, $b = 1$, and $N = 400$. The selected combination of dimensionless parameters is typical of a PRIZ space–time light modulator using bismuth silicate $Bi_{12}SiO_{20}$ (Ref. 13). The writing wavelength is 0.51 μm with

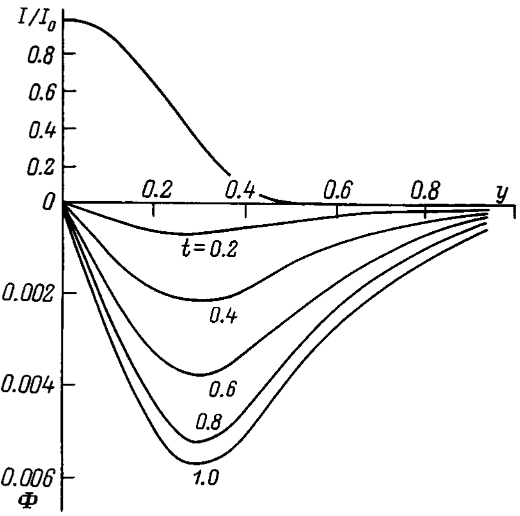


FIG. 6. Distribution of optical intensity I/I_0 along crystal surface and evolution of pulse response Φ under transverse electrooptic modulation.

the exposure $W'=0.1\ \mu\mathrm{J/cm^2}$ which corresponds to an optical pulse of duration 1 μs with an intensity $I_0=100\ \mathrm{mW/cm^2}$ at the line center. For the given exposure a linear transport regime is established and the distributions of the carrier concentration and the space charge density are similar to those shown in Fig. 5. The width of the exposed band is equal to the crystal thickness $b'=L=0.5\ \mathrm{mm}$. For $U=2\ \mathrm{kV}$ and $\mu=0.03\ \mathrm{cm^2/V\cdot s}$ the transit time is $\tau_T=42\ \mu\mathrm{s}$. The concentration of impurity centers with a 100% photoionization quantum yield is $N'=10^{16}\ \mathrm{cm^{-3}}$, the permittivity of the crystal is $\varepsilon=56$, and the absorption coefficient $\alpha'=10\ \mathrm{cm^{-1}}$. Figure 6 also shows the distribution of the optical intensity along the crystal surface defined by the spatial modulation function $I(y)/I_0=Y(y)$ as given by Eq. (19). For negative values of the coordinates the graphs are symmetric since the pulse response is described by the odd function $\Phi(-y,t)=-\Phi(y,t)$. The figure clearly shows a continuous rise in the pulse response with time as far as $t=1$. For times exceeding the transit time $(t>1)$ the coordinate dependence of the pulse response remains unchanged since after extraction of the photoinjected carriers the electric field is only created by the space charge of the ionized impurity centers. The maximum integral of the transverse field intensity is achieved at the point $y\approx 0.31$ where the optical intensity is 31% of the value at the center of the line. At the edge of the exposed band where the intensity falls to zero (at $y=0.5$) the pulse response is 67% of its maximum which characterizes the ''spreading'' of the recorded image.

CONCLUSIONS

An analysis has been made of the pulsed recording of information in a photorefractive crystal. As the exposure increases, the nonlinearity caused by the space charge limitation of the photoinjected carriers becomes significant. Calculations were made of the distribution profiles of the excess carriers and the space charge density in the linear and nonlinear transport regimes. The dynamics of the space charge is determined by contact phenomena at the interface between the illuminated electrode and the photorefractive crystal and differs appreciably for a blocking contact and for optical breakdown of the contact. The space–time distributions of the pulse response under transverse electrooptic modulation were calculated for the efficiency characteristic of the information recording process. The results are suitable for analyzing the process of entering information into a space–time light modulator with an arbitrary absorption coefficient and also for analyzing the pulsed recording of holographic gratings in the limit of low absorption coefficients.

To conclude, the author would like to thank all delegates at the M. P. Petrov Laboratory Seminar for fruitful discussions of the problem.

[1] M. P. Petrov, S. I. Stepanov, and A. V. Khomenko, *Photorefractive Crystals in Coherent Optical Systems* (Springer-Verlag, Berlin, 1991), 257 pp.

[2] P. Günter and J. -P. Huignard, *Photorefractive Materials and Their Applications,* Part 1 (Springer-Verlag, Berlin, 1989), 295 pp.

[3] P. Aing, C. Alves, G. Pauliat *et al.*, Opt. Mem. Neural Netw. **3**(2), 167 (1994).

[4] F. H. Mok, G. W. Burr, and D. Psaltis, Opt. Mem. Neural Netw. **3**(2), 119 (1994).

[5] P. J. Van Heerden, Appl. Opt. **2**, 393 (1963).

[6] G. C. Duree, L. Shultz, Jr., G. J. Salamo *et al.*, Phys. Rev. Lett. **71**, 533 (1993).

[7] B. Grosignani, M. Segev, D. Englin *et al.*, J. Opt. Soc. Am. B **10**, 446 (1993).

[8] M. Taya, M. Bashaw, M. M. Fejer *et al.*, Phys. Rev. A **52**, 3095 (1995).

[9] G. S. Garcia Quirino, J. J. Sanchez Mondragon, and S. Stepanov, Phys. Rev. A **51**, 1571 (1995).

[10] V. V. Bryksin, L. K. Korovin, and Yu. I. Kuz'min, Fiz. Tverd. Tela (Leningrad) **28**, 2728 (1986) [Sov. Phys. Solid State **28**, 1528 (1986)].

[11] V. V. Bryksin, L. K. Korovin, and Yu. I. Kuz'min, Fiz. Tverd. Tela (Leningrad) **29**, 1323 (1987) [Sov. Phys. Solid State **29**, 757 (1987)].

[12] V. V. Bryksin, L. K. Korovin, and Yu. I. Kuz'min, Fiz. Tverd. Tela (Leningrad) **28**, 148 (1986) [Sov. Phys. Solid State **28**, 79 (1986)].

[13] A. M. Bliznetsov, Yu. I. Kuz'min, and A. V. Khomenko, Zh. Tekh. Fiz **58**, 618 (1988) [Sov. Phys. Tech. Phys. **33**, 374 (1988)].

[14] Yu. I. Kuz'min, Pis'ma Zh. Tekh. Fiz **23**(18), 37(1997) [Tech. Phys. Lett. **23**, 711 (1997)].

[15] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* [Clarendon Press, Oxford (1971); Vol. 1, Mir, Moscow (1982), 368 pp.].

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Panoramic measurements of electron beam densities by Thomson scattering of laser radiation

A. P. Solov'ev, O. V. Zyuryukina, and K. I. Svinolupov

Research Institute of Mechanics and Physics, Saratov State University, 410071 Saratov, Russia

(Submitted September 14, 1997; resubmitted March 13, 1998)
Zh. Tekh. Fiz. **69**, 80–83 (June 1999)

The possibility of using a panoramic detector with a television signal-recording system in an apparatus to observe Thomson scattering of laser radiation by a nonrelativistic electron beam is discussed. Panoramas of Thomson and Rayleigh scattering and of the electron beam luminescence are presented. Estimates are given of the sensitivity and spatial resolution of the apparatus. Results of panoramic and single-point methods of investigation are compared. Possibilities for extending the range of the Thomson scattering method to measure the density distribution in nonrelativstic electron beams are discussed. © *1999 American Institute of Physics.* [S1063-7842(99)01406-3]

INTRODUCTION

An experimental method of investigating the parameters of electron and plasma fluxes based on the Thomson scattering of laser radiation has extensive possibilities, providing local measurements and introducing no distortions into the flux. Panoramic studies of the structure of electron beams and plasmas, where the spatial distribution of the signals is recorded simultaneously and not one point after another, are of considerable interest, with the number of points which can be monitored in space depending on the number of observation channels. However, investigations are difficult because of the small scattering signal and the need to isolate this against an appreciable noise background. Particular difficulties are encountered in studies of nonrelativistic electron beams having densities of the order of $10^{10}\,\mathrm{cm}^{-3}$ or lower. Panoramic measurements have been made for plasmas having electron densities of 10^{13}–$10^{16}\,\mathrm{cm}^{-3}$ (Ref. 1). In the proposed study we assess the possibility of panoramic investigations of the parameters of electron beams using Thomson scattering of ruby laser radiation in an experimental system described in detail in Ref. 2, but using a television signal-recording system.

NEW ELEMENTS AND THEIR PARAMETERS

The new system differs from that described in Ref. 2 in that it incorporates facilities for panoramic detection of the measured signals. The detection system includes an objective for coupling out the scattered radiation, a brightness amplifier consisting of two series-connected image converters, an LI-702 superkremnikon video camera, and a recording system comprising a television receiver and an S8-9A oscilloscope. The region inside the diagnostic chamber bounded by an aperture diaphragm was scanned on the television screen and the brightness of a single line of the television frame was analyzed using the oscilloscope. The line and its specific region were selected by a synchronization system. As in Ref. 2, a diaphragm D_2 was installed in the objective to align the system, and this was removed during the measurements. An image of the diaphragm D_2 was projected onto the fiber unit of the image converter and was then used as a reference point to determine the center of the probing region, i.e., the point of intersection of the axes of the laser and electron beams for a suitably selected line of the television frame and its region.

The resolution R of the detection system in the observation space was determined by a second image converter, taking into account the linear magnification of the image by the objective $M=0.6$ and the first image converter $M_1=0.63$, and also allowing for a decrease in the resolution by a factor of $\sqrt{2}$ accompanying the passage through the pair of fiber-optic disks of the two image converters. An experimental check using a reference standard gave $R=5\,\mathrm{mm}^{-1}$ and agreed with the theoretical estimate. The spatially resolved volume V, which can be represented to a certain approximation as a cylinder of length and cross section determined by the laser beam diameter d in the observation zone and by the linear resolution of the detection system, was $V\cong 4\times 10^{-2}\,\mathrm{mm}^3$.

The least sensitive element of the detection system is the superkremnikon. The working illumination of the photocathode for LI-702 tubes is $\cong 5\times 10^{-3}$ lx (Ref. 3) and each element of the image on the photocathode, corresponding to a single pixel of the 30 μm diameter superkremnikon detector should receive a radiation energy $E_s\cong 10^{-16}$ J at 550 nm i.e., approximately 10^3 photons over the frame scanning period of 40 ms. Then, having equated E_s to the corresponding Thomson scattering energy, we obtain an expression for the electron density n_1 for which the Thomson scattering signal of the laser radiation creates the working illumination on the superkremnikon photocathode

$$n_1=E_sS(E_0K_1K_2K\sigma\Omega V_1)^{-1},$$

where σ is the cross section for Thomson scattering, E_0 is the laser pulse energy, S is the cross-sectional area of the laser beam in the probed region, $K_{1,2}=10^2$ are the brightness gains of the first and second image converters, $K=0.3$ is the total transmission coefficient of all the optical elements of the detection system at the Thomson scattering wavelength $\lambda = 630\,\mathrm{nm}$, $\Omega = 2.6\times 10^{-2}$ sr is the solid angle of observation, V_1 is the volume from which the Thomson scattering enters a single pixel and which, by analogy with the resolved volume V, can be represented to some approximation as a cylinder of length equal to the laser beam diameter d in the observation zone and whose cross section is determined by the size of the superkremnikon pixel L_c taking into account the coefficients of magnification of the optical image of the two image converters $M_1 = M_2$ and the objective for coupling out the scattered radiation.

For the parameters given above and a laser energy $E_0 = 30$ J the value of n_1 is approximately $10^9\,\mathrm{cm}^{-3}$ and at the entrance to the first image converter the energy scattered from the spatially resolved volume corresponds to approximately a single photon. Since the quantum yield of the image converter photocathode is several times less than unity $(\cong 0.2)$, a single photoelectron from the photocathode of the first image converter produces an illumination several times greater than the working illumination per frame scanning period in the appropriate region of the superkremnikon photocathode. Thus, the detection system is sufficiently sensitive to record isolated photons.

Since the Thomson scattering signal is small and probabilistic, a series of measurements is required to obtain reliable data and the lower the electron density, the longer the series required. In a real experimental system, the presence of optical noise necessitates a further increase in the number of measurements. The optical noise is made up of luminescence from the gas in the lamp under the action of the electron beam N_e, luminescence from the gun, luminescence from the elements of the receiving channel under the influence of these, and the laser noise N_L (Ref. 2). This laser noise was ten times lower than the luminescence and four times less than the Thomson scattering signal at the point of maximum electron density, which was approximately $8\times 10^9\,\mathrm{cm}^{-3}$. The optical noise was investigated by analyzing the panorama on the television screen and oscilloscope traces of the luminescence in the probing zone, which are clear and convenient for television recording.

EXPERIMENTAL RESULTS AND DISCUSSION

Panoramic measurements were made of the Thomson and Rayleigh scattering of laser radiation by a nonrelativistic electron beam and air, respectively, and results of measurements obtained using panoramic and single-point[2] methods of signal recording were compared. The luminescence distribution over the beam cross section was recorded separately to obtain the Thomson scattering panorama. Figure 1 shows panoramas of the Thomson scattering signals H_T and the electron beam luminescence N_e averaged over a series of 70 shots for a laser beam energy of 30 J, the lines of the television frame oriented at right angles to the electron beam image, and a pressure of 2.6×10^{-5} Pa in the diagnostic chamber. In contrast to Ref. 2, the luminescence and Thomson scattering curves, normalized to their maxima, differ. This can be attributed to the energy distribution in the laser beam cross section since the laser beam was not displaced relative to the electron beam together with the observation point, as in Ref. 2, and the laser beam diameter is approximately equal to the electron beam diameter. Figure 2 shows the distribution of the Rayleigh signal H_R over the laser beam cross section measured for a single laser shot at a pressure of 266 GPa in the chamber. This distribution reflects the energy distribution of the probe radiation over the electron beam cross section when a Thomson scattering panorama is recorded. Thus, the distribution profile obtained for the Thomson signal is narrower than the luminescence profile. If the distribution obtained for the Thomson signal is normalized to the Rayleigh scattering distribution, the curve $H_{TR} = \{H_{TG}\}/\{H_{RG}\}$, as can be seen from Fig. 3, agrees fairly well with the luminescence distribution, as in Ref. 2. Here $\{H_{TG}\}$ and $\{H_{RG}\}$ are the experimentally determined Thomson $\{H_T\}$ and Rayleigh $\{H_R\}$ scattering distributions approximated by a Gauss function. Note that in this system

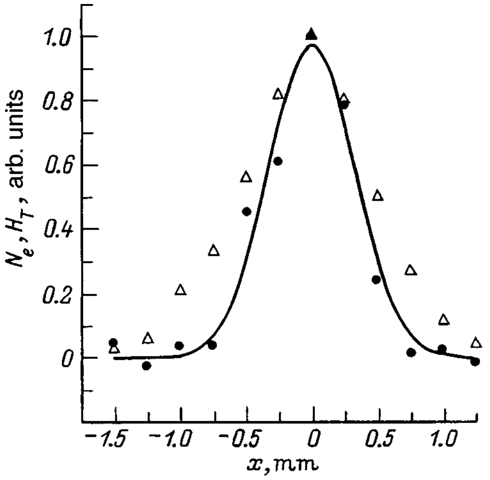


FIG. 1. Panoramas of beam luminescence signals N_e (△) and Thomson scattering H_T (●). Solid curve — approximation of H_T using Gauss function (H_{TG}).

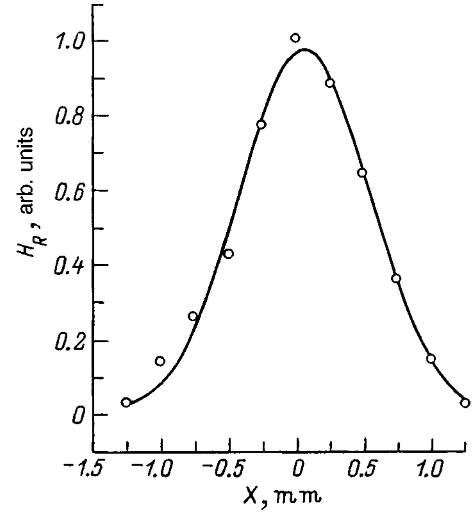


FIG. 2. Distribution of Rayleigh signal H_R over laser beam cross section. Circles — experimental data, solid curve — approximation of H_R using Gauss function (H_{RG}).

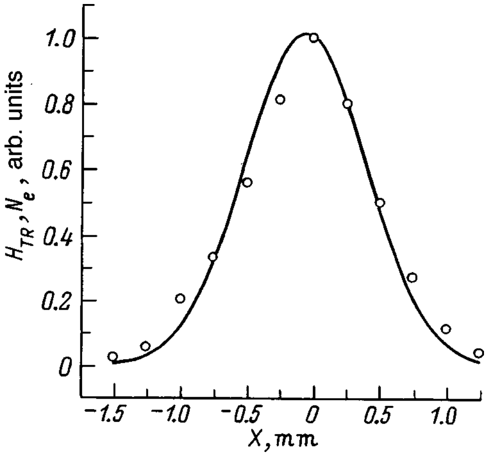


FIG. 3. Distribution of normalized Thomson signal H_{TR} over beam cross section: ○ — experimentally determined beam luminescence data N_e.

where the diameters of the electron and laser beams are approximately the same, as a result of a decrease in the laser beam energy, the Thomson scattering panorama at the beam edges was obtained by averaging over a large series of measurements (around 70) although 35 laser shots were sufficient for the central region.

Variation of the laser beam energy during the recording of a Thomson scattering panorama over the electron beam cross section can be eliminated by rotating the television frame line about the axis of the laser beam, which was easily achieved by rotating one end of the bunch of optical fibers which transfer the image from the image converter to the superkremnikon. Since in this design of diagnostic chamber the electron beam, laser beam, and direction of observation lie in the same plane, the television frame line was also oriented along the axis of the beam image. In this case, the beam luminescence should have a maximum and remain almost constant over the entire 8 mm region of observation. However, the oscilloscope trace of the luminescence shown in Fig. 4 reveals a clearly defined peak which decays rapidly toward the center of the probing region, i.e., toward the point of intersection of the axes of the probe laser and the electron beam. Observations on the television screen revealed that this peak can be attributed to cathode luminescence which illuminated the knife of the first screening diaphragm in the channel for coupling out the scattered radiation. As a result, in the probing region (around 3 mm) there was a slight increase in the overall noise background (on the right of Fig. 5). The luminescence distribution in Fig. 5 (curve *1*) was obtained by averaging over 50 traces. Also shown is the panorama of the Thomson scattering signal H_T (curve *2*) which has a well-defined peak at the center of the probing region which accounts for a third of the luminescence signal at a pressure of 2.9×10^{-5} Pa in the diagnostic chamber.

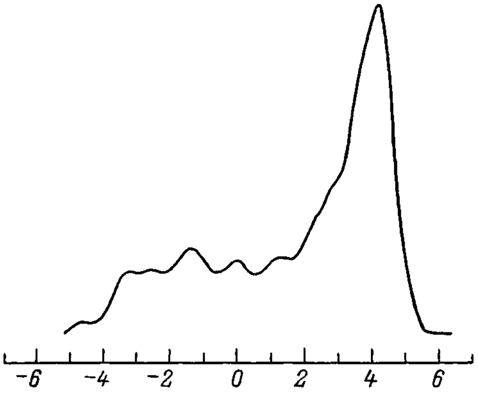


FIG. 4. Oscilloscope trace of luminescence on beam axis with television frame lines oriented along the axis of the beam image; 0 — center of probing region.

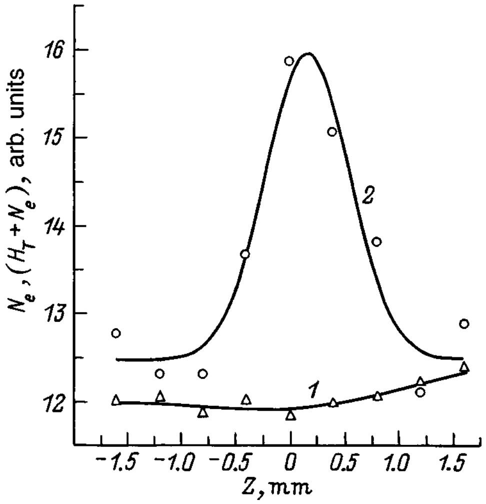


FIG. 5. Panoramas of luminescence N_e (*1*) and Thomson scattering signals H_T (*2*) when the frame lines are oriented along the axis of the beam image; 0 — center of probing region.

A comparison of the measurements obtained by the panoramic and single-channel[2] procedure showed that the ratio of the Thomson scattering signal to the luminescence is slightly greater for the single-channel procedure. This can be attributed to an increase in the exposure time of the luminescence signals on the superkremnikon detector. In the panoramic procedure this time was equal to the frame scanning period, i.e., 40 ms, whereas for the single-channel technique using a photomultiplier as the optical detector,[2] the exposure time was $\cong$10 ms. For the panoramic detector, the luminescence can be reduced substantially by real paths, for instance, by inserting an optical switch at the image converter input which switches on the power supply to one image converter during the laser pulse or switches it off in the absence of a pulse, or by modifying the electronic power supply circuit of the superkremnikon so that the superkremnikon detector was cleared directly before the laser shot near the line (frame) where the image of the part of the beam being studied is situated. By reducing the luminescence exposure time from 40 to 1–2 ms, i.e., down to the laser pulse duration, it is possible to increase the ratio of the Thomson scattering signal to the luminescence approximately 20 times and reduce the series of measurements required to obtain reliable data.

In our view, it is fairly difficult to use Thomson scattering at lower densities of the order of 10^9-10^8 cm^{-3} and in particular, this requires serious modifications to the apparatus used by the authors. As the density decreases, it is desirable to compensate for the decrease in the useful signal reaching the first image converter by increasing the laser radiation energy and the solid angle of observation or by using a substantially longer series of measurements to increase the signal to noise ratio. The length of the series of measurements can be determined experimentally from the constraint that the mean square deviation should not exceed a given value. For example, in studies[4] of the scattering spectrum of Nd laser radiation in a plasma having a density of $10^{12}-10^{13}\,\mathrm{cm}^{-3}$ using a twelve-channel detection system, because of the low laser pulse energy and the small solid angle, the useful signal in the channel with the lowest input energy corresponded to the signal in our apparatus at a density of the order of $10^8\,\mathrm{cm}^{-3}$. A series of 900 laser shots was required to measure this signal in Ref. 4.

When the sensitivity of the apparatus is enhanced, an increase in the useful signal H_T is accompanied by a proportional increase in the laser noise N_L, and under multipass probing this increase is even greater. Thus, the ratio of the signal to the laser noise at best remains constant. In the present apparatus, the magnitude of the laser noise corresponded to the Thomson signal at a density of $10^9\,\mathrm{cm}^{-3}$. At lower densities the ratio H_T/N_L becomes less than 1 and it is desirable to search for methods of reducing the laser noise. The beam luminescence decreases in proportion to the decrease in the density n and causes no deterioration in the ratio of the signal H_T to the total noise. Interference filters in the detection system reduced the laser noise approximately 10^5 times at the wavelength of the scattered light detuned by 60 nm from the probe laser wavelength with a 40% useful signal transmission. Casey *et al.*[5] reported that a polychromator consisting of three holographic gratings based on the principle of a vario-illuminator, suppressed the noise by a factor of 10^{10} at the wavelength of a ruby probe laser but the transmission coefficient of the polychromator was less than 25% in the range 650–690 nm. Using a similar polychromator in the present apparatus could reduce the laser noise by several orders of magnitude, i.e., could substantially increase the ratio of the signal to the laser noise and almost eliminate its influence.

Our estimates and the experimental single-point and panoramic observations of Thomson scattering by a nonrelativistic electron beam having a density of the order of $8\times10^9\,\mathrm{cm}^{-3}$ and also the success achieved in the development of high-power lasers, highly sensitive detectors, and laser noise suppression systems[5] suggest that the unique possibilities offered by Thomson scattering will be of interest not only for research purposes but also for measuring the parameters of real beams.

This work was partially supported by the Russian Committee for Higher Education (Grant No. GR-72-96 and the program ''Laser Physics and Laser Systems'' MLTs, Moscow State University).

[1] G. T. Razdobarin and D. A. Shcheglov, Diagnost. Plazmy No. 6, 88 (1989).
[2] B. G. Tsikin, L. E. Dolotov, O. V. Zyuryukina, and A. P. Solov'ev, Zh. Tekh. Fiz **61**(1), 149 (1991) [Sov. Phys. Tech. Phys. **36**, 89 (1991)].
[3] M. J. Eccles, M. E. Sim, and K. P. Tritton, *Low Light Level Detectors in Astronomy* [Cambridge University Press, Cambridge, 1983; Mir, Moscow, 1986].
[4] P. Jauernik, H. Kempkens, and J. Uhlenbusch, Plasma Phys. Controlled Fusion **29**, 1615 (1987).
[5] S. A. Casey and J. H. Irby, Rev. Sci. Instrum. **57**, 1804 (1986).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Asymptotic form of the nonsteady motion of a charged-particle gas

N. D. Naumov

Central Physicotechnical Institute, 141300 Sergiev Posad, Russia

(Submitted March 5, 1998)
Zh. Tekh. Fiz. **69**, 84–87 (June 1999)

A solution is obtained for the problem of expansion of a gas of an inhomogeneous cloud of rotating particles under the action of its space charge. It is shown that with time the process of expansion of the cloud becomes self-similar. © *1999 American Institute of Physics.* [S1063-7842(99)01506-8]

INTRODUCTION

The construction of analytic solutions of gasdynamics equations has attracted interest because these are one of the methods of studying the properties of nonlinear systems.[1–5] Analytic solutions of the self-consistent equations of motion for a charged-particle gas can generally be obtained under certain assumptions which simplify the real formulation of the problem. Nevertheless, these solutions are of practical interest and can be used for estimates and also for testing numerical simulation programs. Self-similar solutions play a special role for gasdynamic systems since under certain conditions they serve as intermediate asymptotic forms. For nonself-similar processes, details of the initial stage are ''forgotten'' and the process becomes self-similar.

In the present paper this behavior is demonstrated by solving the problem of expansion of a rotating cloud of charged particles under the influence of space charge. Two types of rotation of particles in a spherical cloud are possible, for which the characteristic of the rotational motion depends only on the radial coordinate and the time. In the first type, the particles only undergo ordered motion in the meridional direction $V_\theta \neq 0$, $V_\varphi = 0$. For this model it is comparatively easy to construct a self-similar solution of the gasdynamic equations corresponding to a homogeneous charged-particle cloud. It is found that in this case, a self-consistent nonself-similar solution can also be obtained for an inhomogeneous cluster. For the second type which corresponds to disordered particle rotation, i.e., $V_\theta = 0$, $V_\varphi = 0$, but the corresponding diagonal components of the pressure tensor are nonzero, a similar solution can be obtained using a kinetic description.

SOLUTION OF GASDYNAMIC EQUATIONS

In the absence of azimuthal motion, the Euler equation for a cold spherical cloud of charged particles has the form

$$\left(\frac{\partial}{\partial t} + V_r \frac{\partial}{\partial r}\right) V_r - \frac{1}{r} V_\theta^2 = \frac{e}{m} E, \tag{1}$$

$$\left(\frac{\partial}{\partial t} + V_r \frac{\partial}{\partial r}\right) V_\theta + \frac{1}{r} V_r V_\theta = 0. \tag{2}$$

Here we assume that the particle density and the radial velocity of the cloud depend only on the radial coordinate and time. From Eq. (2) for the function $F = rV_\theta$ we find

$$\left(\frac{\partial}{\partial t} + V_r \frac{\partial}{\partial r}\right) F = 0. \tag{3}$$

The simplest solution of Eq. (3) has the form $F = C$, where C is a constant. A nonsteady-state solution of this equation is obtained for nonsteady motion for which the velocity is proportional to the distance from the center of symmetry $V_r = r\dot{a}/a$, where a is the radius of the sphere, and the dot denotes differentiation with respect to time. In this case, Eq. (3) gives $F = \Omega r^2 a_0^2 / a^2$, where Ω is a certain constant. For this type of motion the particle density is typically uniformly distributed over the bulk of the cloud so that for the collective field we have $E = 4\pi e n r/3$. Finally, we find from Eq. (1) that the cloud radius satisfies the following equation:

$$\ddot{a} = \frac{a_0^3}{a^2}\left(\omega_0^2 + \Omega^2 \frac{a_0}{a}\right),$$

where $\omega_0^2 = Ne^2/ma_0^3$, and N is the total number of particles in the cloud.

The solution of this equation may be expressed in the parametric form[6]

$$a = \frac{a_0}{1+\varepsilon}(\varepsilon \cosh \psi + 1),$$

$$\omega_0 t = \frac{1}{(1+\varepsilon)^{3/2}}(\varepsilon \sinh \psi + \psi), \tag{4}$$

where for an initially stationary gas we have $\varepsilon = 1 + \Omega^2/\omega_0^2$.

Note that this solution of the self-consistent problem in the form of a spherical cloud with an abrupt boundary belongs to a class of self-similar solutions, as can easily be established by introducing the self-similar variable $\xi = r/a$,

$$n(r,t) = n_0 \Gamma \frac{a_0^3}{a^3}, \quad V_r(r,t) = \xi \dot{a} \Gamma,$$

$$V_\theta(r,t) = \xi \Omega \Gamma \frac{a_0^2}{a}, \quad \Gamma = H(1 - \xi^2).$$

Here $H(x)$ is the Heaviside step function. A more general solution of Eq. (3) can be constructed for an inhomogeneous particle cloud. For this we need to bear in mind that the expression for the radial gas velocity can be given in the form:

$$V_r(r,t)=\Lambda(t,\rho(r,t)). \tag{5}$$

Here the function Λ has the form

$$\Lambda(t,r_0)=\frac{\partial s(t,r_0)}{\partial t},$$

where $s=s(t,r_0)$ is the coordinate of the radial gas layer considered as a function of time and the value of this coordinate at zero time; the function $\rho(r,t)$ is a solution of the transcendental equation $s(t,\rho)=r$, i.e., $s(t,\rho(r,t))\equiv r$.

In terms of Lagrange variables the radial velocity of this element is given by

$$v_r(t,r_0)=\Lambda(t,r_0). \tag{6}$$

Therefore, this structure of the expression for the radial gas velocity is a consequence of changing from a Lagrangian to a Eulerian description of the motion. Then, it is easy to verify that $F(r,t)=W(\rho(r,t))$, where $W(r)$ is a certain function, is a solution of Eq. (3). Thus, in Euler variables the transverse velocity component is given by

$$V_\theta(r,t)=\frac{W(\rho(r,t))}{r}.$$

In Lagrange variables this velocity has the form

$$v_\theta(t,r_0)=\frac{W(r_0)}{s(t,r_0)}.$$

These results indicate that the function $W(r)$ is determined by the initial transverse velocity distribution of the cloud

$$W(r)=rV_{\theta 0}(r),\quad V_{\theta 0}(r)\equiv V_\theta(r,0).$$

Provided that the layers of particles move radially one behind the other without overtaking, the collective field acting on a particular layer of gas is determined by its initial position r_0 and by the defined initial particle density distribution $n(r,0)=n_0\nu(r)$

$$E=4\pi n_0\frac{e}{s^2}q(r_0),\quad q(r_0)=\int_0^{r_0}\nu(x)x^2dx.$$

The condition for conservation of the layer mass during motion of the gas has the form $4\pi n(t,r_0)s^2ds=4\pi n_0\nu(r_0)r_0^2dr_0$. The particle density is then given by

$$n(t,r_0)=n_0\nu(r_0)\frac{r_0^2}{s^2R},\quad R(t,r_0)=\frac{\partial s(t,r_0)}{\partial r_0}. \tag{7}$$

It follows from Eq. (1) that the radial velocity of the gas layer changes under the action of the collective field and also under the influence of the centrifugal force

$$\ddot{s}=\omega^2\frac{q(n_0)}{s^2}+\frac{W^2(r_0)}{s^3}, \tag{8}$$

where $\omega^2=4\pi n_0e^2/m$.

The initial conditions for Eq. (8) have the form $s(0,r_0)=r_0$, $\dot{s}(0,r_0)=u(r_0)$, where $u(r)$ is the given initial distribution of the radial velocity.

SOLUTION OF THE VLASOV EQUATION

A solution for a spherically symmetric cloud with disordered particle rotation can be obtained using a kinetic description. We shall analyze the following distribution function which describes the initial state of this cloud with a nonuniform particle density $n(r,0)=n_0\nu(r)$ and the initial radial velocity distribution function $V_r(r,0)=u(r)$,

$$f_0(X,I)=\frac{n_0}{\pi}r^2\nu(r)\,\delta(p_r-mu(r))\,\delta(I-m^2W^2(r)).$$

Here, for conciseness X denotes the set of variables r, p_r; $I=P_\theta^2+P_\varphi^2/\sin^2\theta$; $P_\theta=rp_\theta$, and $P_\varphi=r\sin\theta p_\varphi$ are the components of the generalized momentum. For this distribution function the transverse components of the gasdynamic velocity are zero

$$V_\theta=\frac{1}{mn}\int p_\theta f_0d^3p=0,\quad V_\varphi=\frac{1}{mn}\int p_\varphi f_0d^3p=0.$$

For the pressure tensor only the diagonal transverse components are nonzero

$$\Pi_{\theta\theta}=\frac{1}{m}\int p_\theta^2f_0d^3p=\frac{mn_0}{2mr^2}\nu(r)W^2(r),$$

$$\Pi_{\varphi\varphi}=\frac{1}{m}\int p_\varphi^2f_0d^3p=\frac{mn_0}{2mr^2}\nu(r)W^2(r).$$

Therefore, the function $W(r)$ characterizes the initial dependence of the degree of disordered particle rotation on the radial coordinate. Since for a central field I remains constant, in the Vlasov equation written in spherical coordinates we need to change from the variable θ to the new variable I. As a result, we obtain the following equation for the distribution function $f(X,I,t)$:

$$Lf\equiv\left(\frac{\partial}{\partial t}+\frac{p_r}{m}\frac{\partial}{\partial r}-\frac{\partial U}{\partial r}\frac{\partial}{\partial p_r}\right)f=0, \tag{9}$$

where $U=e\Phi+I/2mr^2$ and Φ is the potential of the collective field.

For the initial distribution function (8) a solution of Eq. (9) can be constructed using a singular solution method which plays the role of the Green function of the operator L (Refs. 6 and 7)

$$H(t)f(X,I,t)=\int G(X,X_0;t)f_0(X_0,I)dX_0,$$

$$LG(X,X_0;t)=\delta(t)\,\delta(X-X_0),$$

$$G(X,X_0;t)=H(t)\,\delta(r-r(t;X_0))\,\delta(p_r-p_r(t;X_0)).$$

Here $r(t;X_0),p_r(t;X_0)$ is the law of radial motion of a single particle in a field having the potential energy U satisfying the conditions $r(0;X_0)=r_0$, and $p_r(0;X_0)=p_{r0}$. This possibility arises because in this case, the nonsteady self-consistent problem reduces to calculating the one-dimensional motion of a cold charged-particle gas in a combination of collective and centrifugal fields.

Finally the distribution function of the charged-particle cloud is given by

$$f(X,I,t)=\frac{n_0}{\pi}\nu(\rho(r,t))\frac{\rho^2(r,t)}{Q(r,t)}\delta(p_r - m\Lambda(t,\rho(r,t)))\delta(I-m^2W^2(\rho(r,t))),$$

where the notation $Q(r,t)=R(t,\rho(r,t))$ is introduced in addition to the quantities used previously.

This distribution function yields the same expressions (7) and (5) for the particle density and the radial velocity. The following relations can be used to calculate the pressure tensor

$$\Pi_{\theta\theta}(s,t)=\Pi_{\varphi\varphi}(s,t)=\frac{mn_0}{2s^2}\nu(r_0)W^2(r_0). \tag{10}$$

By systematically varying r_0 in small steps, expressions (6), (7), and (10) can be used to determine the distribution of the gasdynamic characteristics of the cloud at any given time t.

DYNAMICS OF THE CLOUD

As for a homogeneous cloud, the solution of Eq. (8) is written in the parametric form

$$s=\frac{r_0}{1+\varepsilon}(\varepsilon\cosh\psi+1),$$

$$\omega t=\sqrt{\frac{r_0^3}{q(1+\varepsilon)^3}}(\varepsilon\sinh\psi+\psi). \tag{11}$$

For simplicity the expansion of the cloud under the influence of its space charge is subsequently analyzed for $u(r)=0$; in this case, we have $\varepsilon=1+W^2(r_0)/\omega^2 r_0 q(r_0)$. From the expressions (11) we obtain for the function R

$$R=\frac{1}{1+\varepsilon}\left[1+\varepsilon\cosh\psi+r_0\varepsilon\sinh\psi\frac{\partial\psi}{\partial r_0}+\frac{r_0}{1+\varepsilon}(\cosh\psi-1)\frac{\partial\varepsilon}{\partial r_0}\right],$$

$$\frac{\partial\psi}{\partial r_0}=-\frac{1}{2s(1+\varepsilon)}\left[(\varepsilon\sinh\psi+\psi)\times\left(3-r_0^3\frac{\nu}{q}-\frac{3r_0}{1+\varepsilon}\frac{\partial\varepsilon}{\partial r_0}\right)+2r_0\sinh\psi\frac{\partial\varepsilon}{\partial r_0}\right].$$

Quite clearly, a simpler expression for the function R is obtained when the eccentricity does not depend on the initial position of the layer $\varepsilon=1+\lambda$, i.e., if $W^2(r)=\lambda\omega^2 rq(r)$, where λ is a certain constant. For this choice of initial particle rotation we have

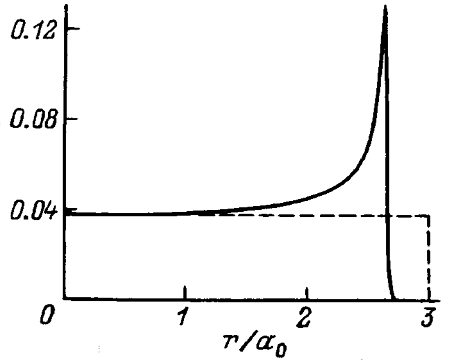


FIG. 1.

$$R=\frac{1}{1+\varepsilon}\left[1+\varepsilon\cosh\psi-\frac{\varepsilon}{2}\sinh\psi\left(3-r_0^3\frac{\nu}{q}\right)\frac{\varepsilon\sinh\psi+\psi}{\varepsilon\cosh\psi+1}\right]. \tag{12}$$

Let us assume that the initial particle density distribution is defined as a cloud with a diffuse boundary

$$\nu(r)=\exp\left(\frac{r^3}{a_0^3}\right). \tag{13}$$

The results of calculating the initial particle density distribution in the cloud using expressions (7), (11) and (12) for $\omega t=4$ in the case $\varepsilon=1.25$ are shown by the solid curve in Fig. 1. The dashed curve gives the self-similar solution (4) for this time and the same value of ε.

The formation of a particle density peak in this example illustrates the resulting limitation of the range of validity of the results, which may arise if the initial assumption that the particle layers move without overtaking is violated as the cloud expands. This is expressed as the condition $R(t_k,r_0)=0$ being satisfied at time t_k for various initial positions of the layers as a result of which the particle density tends to infinity (so-called gradient catastrophes).[8,9]

A comparatively simple expression for the function R is also obtained when

$$\frac{r_0}{1+\varepsilon}\frac{\partial\varepsilon}{\partial r_0}=1-r_0^3\frac{\nu}{3q}.$$

This condition is satisfied if

$$W^2(r_0)=r_0\omega^2 q(r_0)(\varepsilon-1),\quad \varepsilon=\mu\frac{r_0}{q^{1/3}}-1, \tag{14}$$

where μ is a certain constant.

In this case, we have

$$R=\frac{\nu r_0^3}{3q(1+\varepsilon)}(1+\varepsilon\cosh\psi)+\left(1-\frac{\nu r_0^3}{3q}\right)\frac{\varepsilon\cosh\psi+\varepsilon}{\varepsilon\cosh\psi+1}. \tag{15}$$

It can be seen from expression (15) that if the condition $3q(r)\geqslant\nu(r)r^3$ is satisfied, the cloud expands without any overtaking. In particular, this inequality is satisfied for an initial diffuse distribution (13). Thus, in this case the above constraint is not imposed on the time interval during which the analytic expressions can be used.

For large values of the parameter $\sinh\psi \simeq \cosh\psi$, Eqs. (4), (11), (14), and (15) therefore give for the quantities used here

$$a \simeq a_0 \omega t \sqrt{\mu} 3^{-1/3}, \quad s \simeq \omega t \sqrt{\mu} q^{1/3}, \quad R \simeq \frac{\nu}{3q} s r_0^2.$$

From this it follows that

$$n_0 \frac{\nu r_0^2}{s^2 R} \simeq \frac{3 n_0}{(\omega t \sqrt{\mu})^3} \simeq n_0 \frac{a_0^3}{a^3},$$

i.e., with time the details of the initial particle density distribution become insignificant and the cloud expansion process becomes self-similar. Similar behavior of the particle density should be predicted after overtaking in the example considered above.

[1] M. V. Kuzelev and A. A. Rukhadze, *Electrodynamics of Dense Electron Beams in a Plasma* [in Russian], Nauka, Moscow (1990).

[2] I. M. Aleshin and L. S. Kuz'menkov, Vestn. Mosk. Univ., Fiz., Astron. **35**(2), 89 (1994).

[3] Yu. M. Aliev and L. Stenflo, Phys. Scr. **50**, 701 (1994).

[4] P. A. Polyakov, Pis'ma Zh. Tekh. Fiz. **21**(19), 46 (1995) [Tech. Phys. Lett. **21**, 789 (1995)].

[5] N. D. Naumov, Pis'ma Zh. Tekh. Fiz. **22**(19), 89 (1996) [Tech. Phys. Lett. **22**, 817 (1996)].

[6] L. D. Landau and E. M. Lifshitz, *Mechanics*, 3rd ed. [Pergamon Press, Oxford, 1976; Nauka, Moscow, 1973].

[7] N. D. Naumov, Dokl. Akad. Nauk **357**, 759 (1997) [Phys. Dokl. **42**, 326 (1997)].

[8] V. P. Bykov, A. V. Gerasimov, and V. O. Turin, Usp. Fiz. Nauk No. 8, 955 (1995).

[9] A. A. Samarskiĭ and A. P. Mikhaĭlov, *Mathematical Modeling* [in Russian], Nauka, Moscow (1997).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Steady-state longitudinal conductivity of an electron boundary layer

A. V. Ivlev, K. B. Pavlov, and M. A. Yakovlev

N. E. Bauman Moscow State Technical University, 107005 Moscow, Russia

(Submitted January 8, 1998)
Zh. Tekh. Fiz. **69**, 88–93 (June 1999)

An analytic solution is obtained for the problem of electron motion in an electron boundary layer along the surface of a conductor in a static electric field. Calculations are made of the longitudinal conductivity near the surface in the limit of weak and strong Coulomb nonideality of the layer electrons. It is shown that under certain conditions the boundary conductivity may greatly exceed the conductivity in the bulk of the conductor. © *1999 American Institute of Physics.* [S1063-7842(99)01606-2]

INTRODUCTION

In Refs. 1 and 2 the present authors investigated the properties of an electron boundary layer which exists near the free surfaces of conductors and semiconductors. Ivlev *et al.*[3] also examined the influence of an electron boundary layer on the propagation of rf electromagnetic radiation and the ionization processes of a dense gas near the surface. They showed that the main factor determining the electrophysical properties of the electron boundary layer is the Coulomb nonideality parameter of the electrons $\gamma = e^2/4\pi\varepsilon_0\langle r\rangle\varepsilon_{\rm kin}$, where $\langle r\rangle$ is the average interelectron distance and $\varepsilon_{\rm kin}$ is the average kinetic energy of the electrons ($\sim kT$ for nondegenerate and $\sim\varepsilon_F$ for degenerate electrons, respectively).

For small γ ($\gamma \ll 1$) the boundary-layer electrons are similar to an ideal collisionless gas[1] and the collective processes in the electron boundary layer are described by a collisionless transport equation. In the opposite case, if the electrons form a strongly nonideal system ($\gamma \gg 1$), they are similar to a fluid and equations from the mechanics of continuous media must be used to describe the collective processes.[1,2]

The temperature and concentration of conduction electrons in a conductor or semiconductor (we shall subsequently use the term ''conducting condensed material''—CCM) can vary fairly widely if this CCM undergoes nonequilibrium heating by pulses of several picoseconds duration or shorter. Then, the lattice temperature remains almost unchanged and the electron temperature may reach several electronvolts.[4] The electron concentration may vary as a function of the band gap (for semiconductors) between arbitrarily low values and concentrations corresponding to those of a metal. Thus, when a CCM is exposed to pulsed heating, it is possible to have various combinations of the electron nonideality parameter γ, the Fermi energy ε_F, and the temperature T for which CCM and boundary-layer electrons form either a nonideal (degenerate or nondegenerate) or ideal (nondegenerate) Coulomb system.[5]

Note that under certain conditions the longitudinal conductivity near the CCM surface should substantially exceed that in the bulk of the material. This conclusion can be reached a priori, without making accurate calculations, on the basis of simple qualitative reasoning. Regardless of whether the electron component is an ideal gas or an electron liquid, the main ''delaying factor'' for the mass motion of electrons in a CCM are collisions with the crystal lattice. Motion of electrons outside the CCM (along the surface) is either limited by viscous friction forces ($\gamma \gg 1$, electron liquid) which are much greater than the collisional friction forces in a CCM, or it is not limited at all ($\gamma \ll 1$, ideal electron gas). Thus, if the effective time for free (without collisions with the lattice) electron motion above the CCM surface is much greater than the time for free electron drift in the CCM, the electron current at the boundary induced by the application of a longitudinal electric field, should greatly exceed the current in the bulk of the CCM. Consequently, the aim of the proposed study is to analyze the steady-state conductivity near the surface of a CCM for ideal and nonideal electron components.

CONDUCTIVITY IN THE CASE $\gamma \ll 1$

We shall consider the following problem: a CCM occupies the half-space $z<0$, the electron component of the CCM has the temperature *T*, and forms an ideal Coulomb system. A static electric field *E* is directed along the surface of the CCM (along the *x* axis). We need to determine the perturbation of the electron distribution function induced by the field *E* near the surface and the conductivity of the CCM as a function of *z*.

The distributions of the electron concentration *n* and the electric potential Φ of the electron boundary layer in the region $z>0$ have the form[6,7]

$$\Phi(\xi) = -\frac{kT}{e}[1+2\ln(1+\xi)],$$

$$n(\xi) = \frac{n_m}{e}(1+\xi)^{-2}, \tag{1}$$

where $\xi = z/L$, $L = 2\sqrt{e}\,D$, $D = \sqrt{\varepsilon_0 kT/2e^2 n_m}$ is the Debye length corresponding to the electron concentration in the CCM n_m, and e is the base of natural logarithms.

In the region $z<0$, the concentration and potential vary considerably more rapidly with increasing distance from the boundary $\Phi \propto -\exp[-z/D]$, $n \propto \exp[e\Phi/kT]$, and at the distance $\simeq (1.5-2)D$ the electron concentration differs negligibly from n_m. Consequently, in the following calculations we can assume $\Phi \simeq 0$ and $n \simeq n_m$ for $z<0$, and at the point $z=0$ the concentration and potential undergo a jump

$$\Phi(\xi=0_+)=-\frac{kT}{e}, \quad n(\xi=0_+)=\frac{n_m}{e}. \tag{2}$$

Therefore the electron-filled region $z>0$ will subsequently be called the electron boundary layer.

It is impossible to obtain an exact solution of the transport equation in the region $z<0$ which takes into account electron scattering by collisions with the lattice and impurities because of the complex form of the collision integral. Thus, we shall use a simpler linear representation of the collision integral, that is the isotropic τ-approximation.[8] This approximation sufficiently accurately describes the collisional kinetics of the electrons in a conductor if these are scattered by impurities and interelectron collisions can be neglected. In this case, the steady-state transport equation has the form

$$v_z \frac{\partial f}{\partial z} - \frac{e}{m} E(z) \frac{\partial f}{\partial v_x} + \frac{e}{m} \frac{d\Phi}{dz} \frac{\partial f}{\partial v_z} = -\nu_{\text{eff}}(f-f^{(0)}), \tag{3}$$

where ν_{eff} is the effective scattering frequency.

Having written the required distribution function in the conventional format

$$f=f^{(0)}+f^{(1)}, \quad |f^{(1)}| \ll |f^{(0)}|$$

and giving $f^{(1)}$ in the form

$$f=\begin{cases} f_+, & v_z>0, \\ f_-, & v_z<0, \end{cases}$$

we obtain the solution of Eq. (3) (Ref. 9)

$$f^{(1)}_{(\pm)}(z,\mathbf{v})=\exp[\mp\Lambda(\varepsilon,z)]\left[A_\pm(\varepsilon) \pm \frac{e}{m} \frac{\partial f^{(0)}}{\partial v_x} \int_z^z \frac{E(z')\exp[\pm\Lambda(\varepsilon,z')]}{\sqrt{\frac{2}{m}[\varepsilon+e\Phi(z')]}} dz' \right], \tag{4}$$

where

$$\Lambda(\varepsilon,z)=\int_z^z \frac{\nu_{\text{eff}}}{\sqrt{\frac{2}{m}[\varepsilon+e\Phi(z')]}} dz',$$

$$\varepsilon=\frac{1}{2}mv_z^2-e\Phi(z).$$

Here $A_\pm(\varepsilon)$ are the integration functions determined from the boundary conditions of the problem and integration over z' is performed along the electron trajectory.

Steady-state collisionless motion of boundary-layer electrons in the region $z>0$ is described by Eq. (3) without the right-hand side and in consequence, the solution of the transport equation for $z>0$ is given by formula (4) setting $\nu_{\text{eff}}=0$. However, it should be noted that an exact solution of the transport equation[10] can be obtained in the region $z>0$

$$f_\pm(z,\mathbf{v})=f_0\left(v_x+\frac{e}{m}\int_0^\tau \widetilde{E}(\tau',v_z)d\tau', v_y, \pm\sqrt{v_z^2-\frac{2e}{m}\Phi} \right),$$

$$\tau=\int_z^z \frac{dz'}{\sqrt{v_z^2+\frac{2e}{m}[\Phi(z')-\Phi(z)]}}=\Psi(v_z,z),$$

$$\widetilde{E}(\tau,v_z)=E(\Psi^{-1}(\tau,v_z)), \tag{5}$$

where f_0 is the equilibrium distribution function in the absence of the fields E and Φ.

We denote by m and l the functions assigned to the regions $z<0$ and $z>0$, respectively. The unperturbed functions $f_m^{(0)}$ and $f_l^{(0)}$ are Maxwellian distribution functions

$$f^{(0)}=n(z)\left(\frac{m}{2\pi kT}\right)^{3/2}\exp\left(-\frac{mv^2}{2kT}\right).$$

The thickness of the electron boundary layer is fairly small so that we can naturally assume that E does not depend on z. This means that the function $f^{(1)}_{m+}$, corresponding to a uniform flux from the depth of the CCM also does not depend on z. Consequently, using Eq. (4) we find that $A_{m+}=0$ $f^{(1)}_{m+}$ has the form

$$f^{(1)}_{m+}=\frac{eE}{m\nu_{\text{eff}}}\frac{\partial f^{(0)}}{\partial v_x}.$$

We write the boundary conditions determining the integration functions $A_{m-}(\varepsilon)$ and $A_{l\pm}(\varepsilon)$. Using the condition of continuity of the functions $f^{(1)}_\pm$ for given ε, we obtain

$$z=0: \ f^{(1)}_{l\pm}=f^{(1)}_{m\pm},$$

$$z=z^*: \ f^{(1)}_{l+}=f^{(1)}_{l-}, \tag{6}$$

where z^* is the classical turning point determined from the condition

$$e\Phi(z^*)=-\varepsilon. \tag{7}$$

Using expressions (2), (4), and (6) we derive an expression for A_{m-}

$$A_{m-}=2\Theta(|v_z|-\sqrt{2}\,v_e)\frac{eE}{m}\frac{\partial f^{(0)}}{\partial v_x}\tau_0^*(|v_z|), \tag{8}$$

where

$$\tau_0^*(v_z)=\int_0^{z^*}\frac{dz'}{\sqrt{v_z^2+\frac{2e}{m}\Phi(z')}} \tag{9}$$

is the time for electron motion in the field Φ from the surface $z=0_+$ to the turning point $z=z^*$, $\Theta(x)$ is the Heaviside step function, and $v_e=\sqrt{kT/m}$ is the electron thermal velocity.

Substituting the expressions obtained for $A_\pm$ into Eq. (4) we determine the final form of the distribution function in the CCM

$$f_{m+}^{(1)}=\frac{eE}{m\nu_{\mathrm{eff}}}\frac{\partial f^{(0)}}{\partial v_x},$$

$$f_{m-}^{(1)}=\frac{eE}{m\nu_{\mathrm{eff}}}\frac{\partial f^{(0)}}{\partial v_x}$$

$$\times\left[1+2\nu_{\mathrm{eff}}\Theta(|v_z|-\sqrt{2}v_e)\tau_0^*(|v_z|)\exp\left(\frac{\nu_{\mathrm{eff}}}{|v_z|}z\right)\right]. \tag{10}$$

The physical meaning of the expression obtained for $f_{m-}^{(1)}$ is quite obvious. Only those electrons whose velocity in the z direction exceeds the threshold value $\sqrt{-2e\Phi(0_+)/m}=\sqrt{2}v_e$ leave the CCM. An electron leaving the CCM migrates in the field Φ to the right as far as the turning point z^*, reaching it within the time $\tau_0^*(v_z)$. Hence, by the time the electron returns to the CCM, it has acquired the additional momentum $2eE\tau_0^*(v_z)$ in the direction of the z axis.

We now use formula (10) to calculate the conductivity of the CCM σ_m as a function of the depth $z<0$ measured from the surface. Using the relation between the current j and the conductivity σ

$$j_x(z)=\sigma_m(z)\,E=-e\int v_x f_m^{(1)}d\mathbf{v},$$

and also representing the function $f_m^{(1)}$ as the combination $f_{m\pm}^{(1)}$, we obtain

$$\sigma_m(z)=\sigma_m^0\left[1-\frac{\nu_{\mathrm{eff}}}{n_m}\int_{v_z>\sqrt{2}v_e}v_x\tau_0^*(v_z)\frac{\partial f^{(0)}}{\partial v_x}\exp\left(\frac{\nu_{\mathrm{eff}}}{v_z}z\right)d\mathbf{v}\right] \tag{11}$$

Here σ_m^0 is the conductivity in the bulk of the CCM. Integrating Eq. (11) over the velocities v_x and v_y and transforming the integral (9) using Eq. (1), we obtain the final expression for the conductivity

$$\frac{\sigma_m(z)}{\sigma_m^0}=1+\frac{1}{\sqrt{e}}\frac{\nu_{\mathrm{eff}}}{\omega_p}$$

$$\times\int_0^\infty \operatorname{erf} w\exp\left(-w^2+\frac{\nu_{\mathrm{eff}}z}{2v_e\sqrt{w^2+1/2}}\right)$$

$$\times\frac{w}{\sqrt{w^2+1/2}}dw, \tag{12}$$

where $\operatorname{erf} w$ is the error function

$$\operatorname{erf} w=\frac{2}{\sqrt{\pi}}\int_0^w e^{-\zeta}d\zeta,$$

$\omega_p=\sqrt{e^2n_m/m\varepsilon_0}$ is the plasma frequency of the CCM electrons.

It follows from formula (12) that the boundary conductivity is directly proportional to the effective electron scattering frequency in the CCM and inversely proportional to the plasma frequency corresponding to the concentration of conduction electrons. Thus, the increase in the boundary conductivity should be greater, the lower the electron mobility in the CCM and the lower their concentration. The frequency ν_{eff} determined by scattering at impurities may reach $\nu_{\mathrm{eff}}\sim10^{16}\,\mathrm{s}^{-1}$. If the CCM is a semiconductor with the band gap $\Delta E\simeq3$–$4\,\mathrm{eV}$, whose electron component undergoes pulsed heating to the temperature $T\simeq0.3$–$1\,\mathrm{eV}$, the concentration n_m may vary in the range $n_m\sim3\times10^{25}$–$3\times10^{26}\,\mathrm{m}^{-3}$, which corresponds to the plasma frequency $\omega_p\sim(10^{14}$–$10^{15})\,\mathrm{s}^{-1}$. Figure 1 gives the dimensionless conductivity $\sigma_m(z)/\sigma_m^0$ plotted as a function of the dimensionless coordinate $\nu_{\mathrm{eff}}z/v_e$, obtained using formula (12). The parameter $\nu_{\mathrm{eff}}/\omega_p$ is taken to be 100. The boundary value of the dimensionless conductivity $\sigma_m(0)/\sigma_m^0$ is

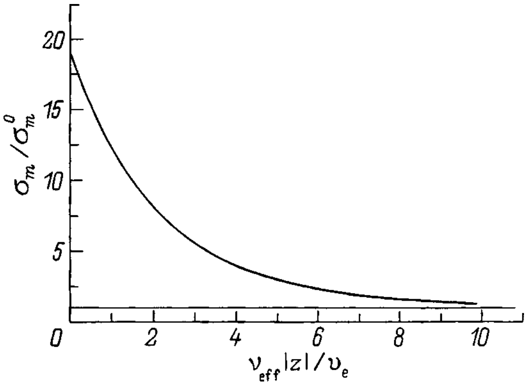


FIG. 1. Steady-state electrical conductivity of conducting condensed material as a function of distance from its surface.

$$\frac{\sigma_m(0)}{\sigma_m^0}=1+\frac{1}{2\sqrt{\pi e}}\frac{\nu_{\mathrm{eff}}}{\omega_p}\int_0^\infty\frac{e^{-r}dr}{\sqrt{r^2+1}(r^2+1/2)}$$

$$\simeq1+0.18\frac{\nu_{\mathrm{eff}}}{\omega_p}.$$

CONDUCTIVITY IN THE CASE $\gamma>1$

We shall now consider a variant where the electron component in the CCM forms a nonideal system i.e., is similar to a liquid. In the CCM (regardless of the degree of nonideality), the conductivity is determined by electron collisions with the lattice because the forces produced by these collisions are always much greater than the viscous friction forces.[11] Under these conditions no f_+ and f_- discontinuities occur in the electron liquid because of strong interparticle interaction at the boundary $z=0$. Consequently for $\gamma\geqslant1$ the conductivity of the CCM in the region $z\leqslant0$ differs negligibly from σ_m^0.

Unlike an electron gas, in the electron liquid of an electron boundary layer filling the region $z>0$ there is a mechanism for internal dissipation of energy caused by the viscous friction forces: the force $\mathbf{f}_{\mathrm{visc}}\simeq(mn)^{-1}\eta\nabla^2\mathbf{u}$ acts on an electron, causing the energy acquired in the field E to be dissipated. Consequently, the viscous friction forces determine the unique characteristic of a boundary layer of nonideal electrons: the conductivity therein is determined by the electron viscosity.[1] For this reason, the formulation of the problem in this particular case of a nonideal electron component will differ from the formulation in the limit $\gamma\ll1$ in that the conductivity will be determined in that part of the electron boundary layer near the surface where the electrons form a nonideal Coulomb system.

No analytic expressions are available to describe the distribution of the concentration of nonideal nondegenerate electrons. The distribution of nonideal degenerate electrons obtained by the density functional method[12] in the limit $T=0$ has the form of a power dependence $n \propto (z/L_{TF})^{-6}$ near the surface and an exponential asymptotic curve for large z (L_{TF} is the Thomas–Fermi radius). However, at fairly high temperatures the electrons rapidly become nondegenerate as the density decreases with increasing z. Thus, for the calculations we shall use the approximate power dependence[1]

$$n(z) \simeq n_0(n_m)\left(1+\frac{z}{L_h}\right)^{-h}, \tag{13}$$

where L_h and h are the effective spatial scale and the exponent which depend on n_m and vary in the ranges $L_{TF} \leqslant L_h \leqslant L$, and $6 \geqslant h \geqslant 2$.

Far from the surface of the CCM the dependence (13) yields the distribution of an ideal electron gas (with the exponent $h=2$).

We shall use the equation for the moments of the distribution function[13] to describe the electron motion in the electron boundary layer. The equation for the first moment (equation of continuity) is satisfied identically since the motion takes place along the x axis. The steady-state equation for the second moment has the form

$$en(z)E+\frac{d\pi_{xz}}{dz}=0, \tag{14}$$

where π_{xz} is the xz component of the stress tensor expressed in terms of the distribution function

$$\pi_{xz}=m\int (v_x-u_x)\, v_z f(\mathbf{v})d\mathbf{v}, \tag{15}$$

u_x is mass velocity of the electrons along the surface.

In the region of the electron boundary layer filled with the electron liquid π_{xz} has the form

$$\pi_{xz}=-\eta\frac{du_x}{dz}, \tag{16}$$

where η is the coefficient of viscosity of the electron liquid.

The dependence of η on n was obtained using a molecular dynamics method[1,14]

$$\eta(n)=\eta_g\left[1+\beta\left(\frac{n}{n_b}\right)^{\alpha}\right], \tag{17}$$

where η_g is the viscosity of the electron gas, which does not depend on n, n_b is the concentration determined from the condition $\gamma(n_b)\simeq 1$, and α and β are fitting parameters which vary between zero and unity.

Thus, if the CCM electrons form a strongly nonideal Coulomb system, the electron boundary layer near the surface is a layer of electron liquid with the concentration distribution (13). The motion of the liquid in the field E is described by Eq. (14) together with Eq. (16). On reaching an arbitrary boundary z_b (the point where $n \simeq n_b$), the liquid is converted to an ideal electron gas filling the region $z \geqslant z_b$ and distributed according to the law

$$n(z) \simeq n_b\left(1+\frac{z-z_b}{L_b}\right)^{-2}, \tag{18}$$

where $L_b=2D(n_b)$.

The electron motion for $z \geqslant z_b$ is described using Eq. (14) with the tensor component (15).

The coordinate z_b is selected so that for any $0<z\leqslant z_b$ the hydrodynamic approximation describes the behavior of the electron liquid sufficiently accurately. We shall integrate Eq. (14) over z between $z=z_b$ and $z=\infty$. Taking into account Eqs. (16) and (18), we obtain

$$eEL_b n_b+\pi_{xz}(\infty)+\eta(z_b)\left.\frac{du_x}{dz}\right|_{z_b}=0. \tag{19}$$

We show that $\pi_{xz}(\infty)=0$ for which we calculate $\pi_{xz}(z)$ for $z>z_b$. The electron gas in this region is ideal and collisionless so that we can use the solution of the collisionless transport equation (15). In accordance with Eq. (5) the electron distribution function f at point z has the form

$$f(z,\mathbf{v})=\begin{cases} f^{(0)}(z,v_x-u_x+\delta u_x,v_y,v_z), & v_z>0,\\ f^{(0)}(z,v_x-u_x-\delta u_x,v_y,v_z), & v_z<0,\end{cases} \tag{20}$$

where $f^{(0)}$ is a Maxwellian distribution function, and u_x and δu_x are given by

$$u_x=u_x(z_b)+\frac{eE}{m}\int_{z_b}^{z*}\frac{dz'}{\sqrt{v_z^2+\dfrac{2e}{m}[\Phi(z')-\Phi(z)]}},$$

$$\delta u_x=\frac{eE}{m}\int_{z}^{z*}\frac{dz'}{\sqrt{v_z^2+\dfrac{2e}{m}[\Phi(z')-\Phi(z)]}}, \tag{21}$$

where $u_x(z_b)$ is the mass velocity at the boundary z_b.

The turning point $z*$ is determined from the condition

$$e\Phi(z*)=e\Phi(z)-\frac{1}{2}mv_z^2.$$

Substituting Eqs. (20) and (21) into Eq. (15), we obtain the required expression for π_{xz} after various transformations

$$\begin{aligned}\pi_{xz}(z)&=4\sqrt{2}n(z)v_e\frac{eE}{\omega_p(z)}\int_0^{\infty} w\,\mathrm{erf}\,we^{-\omega}dw\\ &=eEL_b\sqrt{n_b n(z)},\end{aligned} \tag{22}$$

where $n(z)$ corresponds to Eq. (18).

Thus, we find $\pi_{xz}(z)\to 0$ for $z\to\infty$ and Eq. (19) has the form

$$\eta(z_b)\left.\frac{du_x}{dz}\right|_{z_b}=-eEL_b n_b. \tag{23}$$

Expression (23) must be taken to be a boundary condition imposed at the point z_b on the flow of electron liquid. The second quite obvious condition

$$[u_x]|_{z=0}=0 \tag{24}$$

is set at the boundary of the CCM and requires continuity of the longitudinal velocity u_x. The equation describing the flow in the region $0\leqslant z\leqslant z_b$, after substituting Eqs. (13), (16), and (17) into Eq. (14) has the form

$$\frac{d}{d\xi}\left[(1+\beta\tilde{n}^{\alpha}(1+\xi)^{-\alpha h})\frac{du}{d\xi}\right]=\frac{en_0L_h^2E}{\eta_g}(1+\xi)^{-h}, \tag{25}$$

where $\xi=z/L_h$, $\xi_b=z_b/L_h$, $\tilde{n}=n_0/n_b=\gamma^3(n_m)$ (the symbol x is omitted here and subsequently).

Integrating Eq. (25) allowing for the boundary conditions (23) and (24) and using the relation $\sigma E=-enu$, we obtain the following expression for the conductivity of the electron boundary layer:

$$\frac{\sigma(\xi)}{\sigma_m^0}=\frac{n_0}{n_m}(1+\xi)^{-h}\left[1+\frac{e^2n_0^2L_h^2}{\eta_g\sigma_m^0(h-1)}\right.$$
$$\left.\times\int_0^{\xi}\frac{(1+\xi')^{1-h}-\tilde{n}^{-1}(L_b/L_h)}{1+\beta\tilde{n}^{\alpha}(1+\xi')^{-\alpha h}}d\xi'\right]. \tag{26}$$

The integral in Eq. (26) may be expressed in terms of incomplete beta functions but it is somewhat inconvenient to use the expression obtained for the following analysis so that we shall use the following reasoning. For large $\gamma(n_m)$ the parameter $\tilde{n}$ is large and thus the one in the denominator of the integral (26) can be neglected as far as the limit ξ_b, at which $\tilde{n}^{\alpha}(1+\xi)^{-\alpha h}\simeq 1$. For the same reasons we can neglect the second term in the numerator of the integrand. As a result of integrating, we obtain the final expression for the dimensionless conductivity

$$\frac{\sigma(\xi)}{\sigma_m^0}=\frac{n_0}{n_m}(1+\xi)^{-h}\left[1+\frac{\beta^{-1}\tilde{n}^{-\alpha}\Theta}{(h-1)(2-(1-\alpha)h)}\right.$$
$$\left.\times[(1+\xi)^{2-(1-\alpha)h}-1]\right]. \tag{27}$$

where the following notation is introduced

$$\Theta=\frac{e^2n_0^2L_h^2}{\eta_g\sigma_m^0}.$$

When solving Eq. (25), we assumed that the temperature of the electron liquid does not vary along the ξ coordinate. Such an assumption is quite justified if the electron thermal velocity in the electron boundary layer far exceeds their mass velocity. In fact, if we use the solution obtained (27), it follows from the steady-state heat conduction equation

$$\kappa\frac{d^2T}{dz^2}+\sigma E^2=0$$

(where $\kappa\sim v_e\langle r\rangle$ is the thermal diffusivity) that the characteristic scale of temperature variation is $l_T\sim(v_e/|u|)\langle r\rangle\gg L_h$, i.e., in the weak perturbation approximation the temperature can be considered to be uniform.

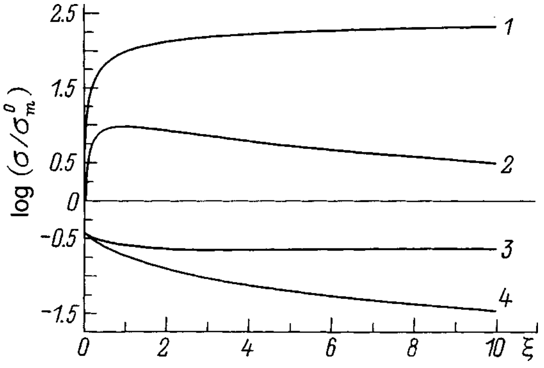


FIG. 2. Steady-state electrical conductivity of electron boundary layer as a function of distance from conductor surface $\beta^1\tilde{n}^{-\alpha}\Theta$: *1, 2* — 100, *3* — 0.1, and *4* — 1; $h=1.2$ (*1*), 2 (*2, 4*), and 1.2 (*3*).

We shall examine how the solution (27) varies as a function of the parameters of the problem. If $\Theta>\beta\tilde{n}^{\alpha}h(h-1)$ and $h<2/(2-\alpha)$, the conductivity of the electron boundary layer increases monotonically with increasing distance from the CCM boundary; if the opposite inequalities are satisfied, the conductivity decreases monotonically. If $\Theta>\beta\tilde{n}^{\alpha}h(h-1)$ and $h>2/(2-\alpha)$, the conductivity initially increases and then, after reaching a maximum, begins to decrease. If the opposite inequalities are satisfied, the conductivity initially decreases and then begins to increase. An extreme value of the conductivity is achieved at the point

$$\xi_{ex}=\left[\frac{\Theta-\beta\tilde{n}^{\alpha}(h-1)[2-(1-\alpha)h]}{\Theta[(2-\alpha)h-2]}h\right]^{\frac{1}{2-(1-\alpha)h}}-1.$$

Figure 2 gives the results of calculations made using formula (27) which illustrate this behavior of the curve $\sigma=\sigma(\xi)$ (the parameter is $\alpha=0.5$, $n_0/n_m=1/e$). Since α varies as a function of $\gamma(n_m)$ between zero and one, and $h\geqslant 2$, in practice only those cases corresponding to curves *2* and *4* can be achieved.

CONCLUSIONS

We therefore draw the following conclusions on the basis of these results. When an ideal ($\gamma\ll 1$) electron component in a semiconductor is heated under nonequilibrium to temperatures $T\geqslant 0.3\,\text{eV}$ for times $t\leqslant 10^{-11}\,\text{s}$, the electron conductivity of a CCM boundary layer of thickness of the order of v_e/ν_{eff} may far exceed the conductivity in the bulk of the heated sample. This effect can be achieved if ν_{eff} is determined by scattering at impurities and its value is of the order $\nu_{\text{eff}}\sim 10^{15}$–$10^{16}\,\text{s}^{-1}$.

If the CCM electrons form a highly nonideal Coulomb system, the conductivity of the electron boundary layer near the CCM boundary may far exceed the conductivity σ_m^0. As in the case of an ideal electron component, this situation may be achieved if the electron mobility in the CCM is fairly low ($\Theta\propto\nu_{\text{eff}}$).

This effect can be used to achieve efficient heating of CCM boundary layers several angstrom thick by applying a static electric field along its surface. As ν_{eff} increases, this effect should become stronger not only as a result an increase in the boundary conductivity but also because the thermal conductivity of the CCM is proportional to ν_{eff}^{-1} and in this case the transfer of heat to the bulk of the CCM will be slowed.

[1] A. V. Ivlev, K. B. Pavlov, and M. A. Yakovlev, Zh. Tekh. Fiz **64**(9), 51 (1994) [Tech. Phys. **39**, 888 (1994)].

[2] A. V. Ivlev and M. A. Yakovlev, Zh. Tekh. Fiz **65**(4), 142 (1995) [Tech. Phys. **40**, 368 (1995)].

[3] A. V. Ivlev, M. A. Yakovlev, and A. N. Bordenyuk, Zh. Tekh. Fiz **68**(8), 48 (1998) [Tech. Phys. **43**, 921 (1998)].

[4] S. I. Anisimov, Ya. S. Imas, and G. S. Romanov *et al.*, *Action of High-Power Radiation on Metals* [in Russian], Nauka, Moscow (1970), 272 pp.

[5] L. P. Kudrin, *Statistical Plasma Physics* [in Russian], Atomizdat, Moscow (1974), 49 pp.

[6] K. B. Pavlov and M. A. Yakovlev, Izv. Akad. Nauk BSSR Ser. Fiz. No. 1, 84 (1989).

[7] Yu. V. Afanas'ev and A. P. Kanavin, Kvantovaya Élektron. (Moscow) **10**, 2267 (1983) [Sov. J. Quantum Electron. **13**, 1473 (1983)].

[8] I. M. Lifshits, M. Ya. Azbel', and M. I. Kaganov, *Electron Theory of Metals* (Consultants Bureau, New York, 1973; Nauka, Moscow, 1971, 416 pp.]

[9] M. A. Liberman, B. É. Meĭerovich, and L. P. Pitaevskiĭ, Zh. Éksp. Teor. Fiz. **62**, 1737 (1972) [Sov. Phys. JETP **35**, 904 (1972)].

[10] V. F. Zaĭtsev and A. D. Polyanin, *Handbook of Partial Differential Equations* [in Russian], International Education Program, Moscow (1996), 496 pp.

[11] V. L. Ginzburg, *The Propagation of Electromagnetic Waves in Plasmas*, 2nd ed. (Pergamon Press, Oxford, 1970; Nauka, Moscow, 1967, 684 pp.

[12] *Theory of the Inhomogeneous Electron Gas* edited by S. Lundkvist and N. H. March (Plenum Press, New York, 1983; Mir, Moscow, 1987, 400 pp.

[13] A. A. Vlasov, *Statistical Distribution Functions* [in Russian], Nauka, Moscow (1966), 356 pp.

[14] B. Bernu and P. Vieillefosse, Phys. Rev. A **18**, 2345 (1978).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Methods of calculating the band structure and low-energy secondary electron spectroscopy of iridium

O. F. Panchenko

A. A. Galkin Donetsk Physicotechnical Institute, Ukrainian National Academy of Sciences, 340114 Donetsk, Ukraine

(Submitted January 29, 1998)
Zh. Tekh. Fiz. **69**, 94–96 (June 1999)

A theoretical interpretation is put forward for the fine structure of the secondary electron emission spectra of Ir normal to the (111) surface and the total current spectrum of an Ir polycrystal. The calculations took into account the energy dependence of the broadening of the energy band levels, the electron–electron and electron–plasmon contributions to the nonequilibrium electron distribution function, and the isotropic component of the current from the electrons scattered at the surface. It is shown that the fine structure of the secondary electron emission spectrum and the total current spectrum is mainly attributable to the electron structure of the final states into which the electrons enter or from which they are emitted so that the characteristics of the band configuration in the energy band structure can be reconstructed directly from the experimental data. This method can be used to separate bulk effects from surface effects in the secondary electron emission and total current spectra. It is confirmed that the fine structure of the secondary electron emission and total current spectra depends on the geometric structure and the degree of ordering of the crystals. A reduction in the intensity of the fine structure serves as a measure of the defect structure in the surface region of the sample which can be successfully used to monitor the surface state during treatment. © *1999 American Institute of Physics.* [S1063-7842(99)01706-7]

Low-energy secondary emission spectroscopy based on studying the effects accompanying the interaction between a flux of slow primary electrons I_p (of energy $E_p \leqslant 1$ keV) and a crystal surface includes two methods:[1,2] a differential and an integral method. The first method gives the energy distribution curve of the secondary electrons outside the crystal or the secondary electron emission (SEE) spectrum, while the second method gives the curve of the integral (or total) secondary electron current in the sample or the total current spectrum. As well as conventional applications in analyzing elementary excitations and the surface states of solids, both methods are now being used to monitor surface purity and in the technology used to fabricate electrode materials.

The fundamental fine-structure characteristics of the SEE and total current spectrum are mainly attributable to the bulk band structure of the crystal.[3,4] The emission of secondary electrons has been investigated experimentally by many authors, in most cases using polycrystals. Single-crystal targets have been used in comparatively few studies. A fine structure against the background of a cascade maximum in the SEE spectrum was observed for example in the secondary electron energy distribution for Ir (Ref. 5), Ag (Ref. 6), Si (Ref. 7), Pt (Ref. 8), W (Refs. 3 and 9), and Cu (Ref. 10) single crystals. Theoretical analysis of the SEE spectra proved extremely difficult because of the need to allow for a wide range of processes accompanying the interaction between the primary electron flux and the crystal. The theories of secondary electron emission from metals (see, for instance, Refs. 11–14), mostly based on the Sommerfeld model, mainly differ in respect of the method of approximating the cascade processes for different surface models. Moreover, formulas having the form: $J(E) \sim \varkappa(E)\,(E+E_i)^{-s}$ have been obtained for the energy distribution of the secondary electrons (cascade maximum), where $\varkappa(E)$ is the coefficient of transmission of the crystal–vacuum barrier for the secondary electrons and the second cofactor is proportional to the number of secondary electrons of energy E or the distribution function of nonequilibrium secondary electrons $f(E)$. The parameters s and E_i differ in different theories. In Ref. 11, for example, we find $s \approx 2$ and E_i is the energy difference between the vacuum level E_{vac} and the bottom of the conduction band. In practice, E_i is replaced by a fitting value.[6] The experimentally observed fine structure of the SEE spectrum cannot be explained by theories using the free electron model neglecting the influence of the band structure. Calculations of the SEE spectrum made by Christensen and Willis[3] showed that the fine structure is related to the bulk density of states $\rho(E)$; in this case, neglecting the broadening of the energy levels, the authors only analyzed the position of the lines in the spectrum, not their shape and intensity. Kleinherbers *et al.*[10] proposed an interpretation of the fine structure of the SEE spectrum based on the theory of low-energy electron diffraction. The results of this study differ from those of Ref. 3, although they satisfactorily describe the fine structure of the experimental curves. Artamonov *et al.*[15] and Korablev *et al.*[16] showed that the fine structure of the SEE spectrum is determined by the energy dispersion of the high electronic states (higher than E_{vac}) and reflects the band boundaries in the dispersion law of electrons moving in the direction of recording.

Comparatively few studies (see for example, Refs. 2, 4, 17, and 18) have been made of the total current spectrum, i.e., the derivative with respect to E_p of the total current in the sample circuit $I=I_p-I_s$ (where I_s is the current of electrons leaving the sample) provided that the secondary electrons are completely removed. The results of these studies indicate that the fine structure of the total current spectrum is strongly related to the fundamental properties of the material. In accordance with Ref. 2, in the energy range up to 100 eV electron–electron $(e-e)$ scattering predominates with the excitation of interband transitions and the main structures in the total current spectrum reflect the $\rho(E)$ characteristics. Calculations made in Ref. 18 using the dynamic theory of low-energy electron diffraction confirm that the fine structure of the total current spectrum is related to the bulk band structure.

The aim of the present study is the investigate and interpret the fine structure of the SEE and total current spectrum of iridium in terms of the bulk band structure and to develop a method of analyzing the experimental results to obtain the maximum information on the electron dispersion law above E_{vac}. As previously (see, for example, Refs. 7, 9, 19, and 20), for our calculations of the SEE and total current spectrum we considered the scattering of electrons having a given momentum on a crystal where the scattering probability is proportional to the number of final states at a given level E with a given direction of quasimomentum Ω. Allowance was made for energy dependence of the broadening of the energy band levels $\hbar\Gamma(E)-\hbar/\tau(E)$, the electron–electron and electron–plasmon contributions to $f(E)$, and the isotropic component of the current from the electrons scattered at the surface.

This work is based on the real band structure $E_{n\mathbf{k}}$ and $\rho(E)$ of iridium; $E_{n\mathbf{k}}$ appears in the calculations of the spectra in terms of $N(E,\mathbf{\Omega})$ (Fig. 1a), this being the number of energy bands in the direction Ω for which the equality $E=E_{n\mathbf{k}}$ is satisfied. The structure of the iridium energy bands (like that of other $5d$ transition metals) obtained by various authors using different methods generally differs in the energy range above the Fermi level E_F. The bulk band structure of iridium was calculated by Nemoshkalenko *et al.*[21] using the relativistic generalization of the connected plane wave method, by Ray *et al.*[22] using a relativistic Korringa–Kohn–Rostoker method, and by Noffke and Fritsche[23] using the self-consistent method of relativistic linear muffin-tin orbitals. In this last study the energy bands obtained over a wide energy range agree with the experimental values of the photoemission spectra with angular resolution.[24]

Figure 1b gives results of calculating the SEE spectra along the normal to the Ir (111) surface using bulk band structure given by various authors. The background current component — a structureless cascade maximum (peak A), was taken into account by adding the constant C to $N(E,\mathbf{\Omega})$, when the energy structure of the surface region is described by the model of an almost free electron gas. The following values of the parameters were used in the calculations of $J(E,\mathbf{\Omega})$: $C=4$ (gives a cascade maximum whose profile and width are similar to those observed experimentally); $E_{\mathrm{vac}}=E_F+e\varphi$, where $E_F=10.8\,\mathrm{eV}$ and $e\varphi=5.8\,\mathrm{eV}$. The excited-state lifetime $\tau(E)$ was determined from Ref. 25: $\hbar/\tau(E)=E_{\mathrm{pl}}\,(E/E_F-1)^2$, where E_{pl} is the screening parameter, and the coefficient $\varkappa(E)$ for one-dimensional motion is taken from Ref. 26. The filling function of the states $f(E)$, corresponding to multiple electron–electron scattering is obtained for $E-E_F \ll E_p$ by solving the transport equation using a statistical model of electron–electron scattering.[27] The decay of plasmons generated by primary electrons and also by excited electrons in a solid makes its own contribution to $f(E)$ which is obtained from the energy conservation law if the dispersion of the plasmons is neglected.

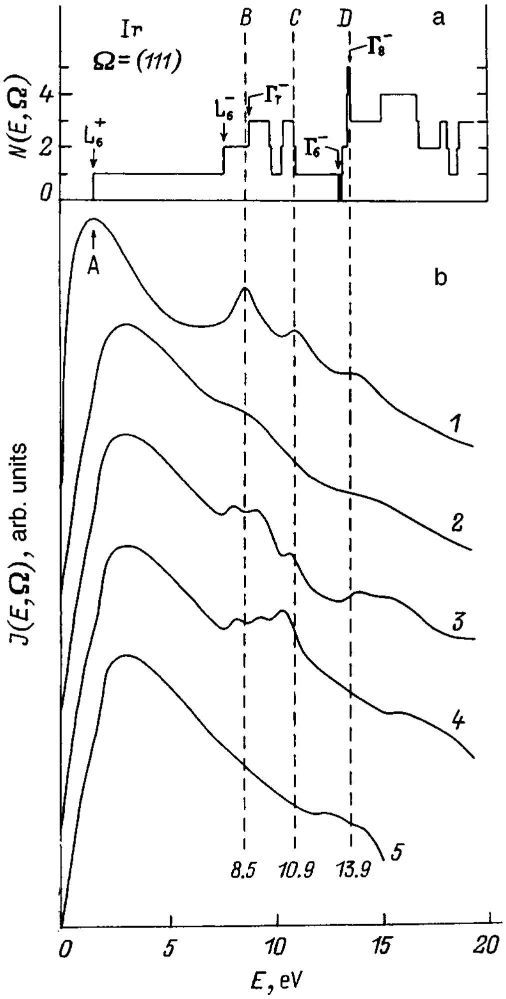


FIG. 1. a — number of electron dispersion law branches for iridium (L_{6^+}, L_{6^-}, Γ_{6^-}, Γ_{7^-}, Γ_{8^-} are its symmetry points), intersecting the level E in the (111) direction; b — secondary electron emission spectrum along the normal to the (111) surface: *1* — experiment from Ref. 5 for $E_p=40$ eV; *2* — theory based on band structure calculations for $E_{\mathrm{pl}}\approx 0.96$ eV; *3–5* — theory based on band structure calculations from Refs. 23, 21, and 22, respectively for $E_{\mathrm{pl}}\approx 0.27$ eV; energy E measured from E_{vac}; the curves are arbitrarily spaced along the ordinate. The vertical dashed lines A–D indicate the main characteristics of the experimental secondary electron emission spectrum.

The best agreement with the fine structure (peaks B, C, and D) in the experimental spectrum (curve *1*) is observed for curve *3* where the broadening parameter E_{pl} which depends on the concentration of s- and d electrons in the electron shell of the atom, was fitted rather than calculated from the general theory of metals (as for curve *2*). This is because

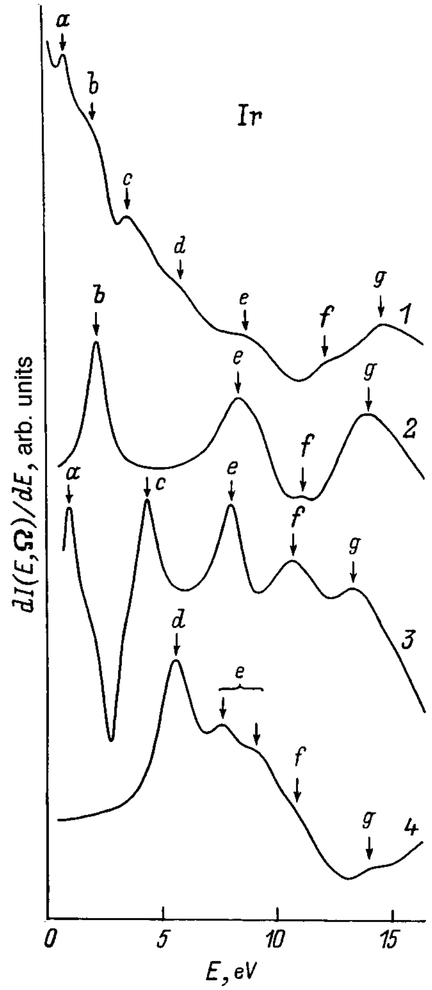


FIG. 2. Total current spectrum of iridium: *1* — experiment for a polycrystal (from Ref. 17); *2–4* — theory for the (111), (110), and (100) faces, respectively for $E_{\mathrm{pl}} \approx 0.96$ eV. The energy E is measured from E_{vac}. The curves are arbitrarily separated along the ordinate. The arrows *a–g* indicate the main features of the experimental total current spectrum.

$\tau(E)$, which is responsible for the broadening of the peaks and was used in the calculations, was obtained near the Fermi surface in Ref. 25 and not in the region of higher excited states. The position and intensity of the fine-structure peaks *C* and *D* on curve *4* differ substantially from similar characteristics of the experimental spectrum because of the approximate nature of the band calculations[21] in the connected plane wave method for $E \geqslant 15$ eV higher than E_F. On the basis of the band structure calculations[22] using the Korringa–Kohn–Rostoker method no fine structure could be identified in the experimental spectrum (curve *5*).

As a result, the total current spectrum was calculated using the bulk band structure[23] and the broadening parameter E_{pl} was calculated from the general theory of metals which did not influence the fine-structure characteristics of the derivative of the total current. The results of the calculations of $dI(E,\mathbf{\Omega})/dE$ (Fig. 2, curves *2–4*) demonstrate (as in Ref. 28 for Pt) the contribution of the various crystallographic planes to the total current spectrum of the polycrystalline sample (curve *1*).

To sum up, these methods can be used to determine experimentally the boundaries of the energy bands $N(E,\mathbf{\Omega})$ in the energy range above E_{vac}, supplementing the photoemission spectral data.

This work was supported by the Ukraine State Fund for Fundamental Research.

[1] A. R. Shul'man and S. A. Fridrikhov, *Secondary Emission Methods of Studying Solids* [in Russian], Nauka, Moscow (1977), 552 pp.

[2] S. A. Komolov, *Integral Secondary-Electron Surface Spectroscopy* [in Russian], Leningrad State University Press, Leningrad (1986), 180 pp.

[3] N. E. Christensen and R. F. Willis, J. Phys. C **12**, 167 (1979).

[4] I. Schäfer, M. Schlüter, and M. Skibowski, Phys. Rev. B **35**, 7663 (1987).

[5] J. U. Mack, E. Bertel, F. P. Netzer, and D. R. Lloyd, Z. Phys. B **63**, 97 (1986).

[6] M. P. Seah, Surf. Sci. **17**, 132 (1969).

[7] O. M. Artamonov, A. G. Vinogradov, O. F. Panchenko *et al.*, Fiz. Tverd. Tela (Leningrad) **31**, 57 (1989) [Sov. Phys. Solid State **31**, 31 (1989)].

[8] B. Lang, Surf. Sci. **66**, 527 (1977).

[9] V. V. Korablev, Yu. A. Kudinov, O. F. Panchenko *et al.*, Fiz. Tverd. Tela (St. Petersburg) **36**, 2373 (1994) [Phys. Solid State **36**, 1290 (1994)].

[10] K. K. Kleinherbers, A. Goldmann, E. Tamura, and R. Feder, Solid State Commun. **49**, 735 (1984).

[11] P. A. Wolff, Phys. Rev. **95**, 56 (1954).

[12] G. F. Amelio, J. Vac. Sci. Technol. **7**, 593 (1970).

[13] M. S. Chung and T. E. Everhart, Phys. Rev. B **15**, 4699 (1977).

[14] M. Rösler and W. Brauer, Phys. Status Solidi B **104**, 161 (1981).

[15] O. M. Artamonov and A. N. Terekhov, Fiz. Tverd. Tela (Leningrad) **28**, 862 (1986) [Sov. Phys. Solid State **28**, 479 (1986)].

[16] V. V. Korablev, Yu. A. Kudinov, and S. N. Sysoev, Fiz. Tverd. Tela (Leningrad) **29**, 702 (1987) [Sov. Phys. Solid State **29**, 402 (1987)].

[17] S. A. Komolov, Izv. Akad. Nauk SSSR, Ser. Fiz. **46**, 1377 (1982).

[18] E. Tamura, R. Feder, J. Krewer *et al.*, Solid State Commun. **55**, 543 (1985).

[19] S. A. Komolov, O. F. Panchenko, and L. K. Panchenko, Fiz. Tverd. Tela (St. Petersburg) **38**, 3172 (1996) [Phys. Solid State **38**, 1733 (1996)].

[20] O. F. Panchenko and V. M. Shatalov, Zh. Tekh. Fiz **63**(10), 145 (1993) [Tech. Phys. **38**, 904 (1993)].

[21] V. V. Nemoshkalenko, Vl. N. Antonov, and V. N. Antonov, Metallofizika (Kiev) **3**(4), 39 (1981).

[22] P. N. Ray, J. Chowdhuri, and S. Chatterjee, J. Phys. F **13**, 2569 (1983).

[23] J. Noffke and L. Fritsche, J. Phys. F **12**, 921 (1982).

[24] J. F. Van der Veen, F. J. Himpsel, and D. E. Eastman, Phys. Rev. B **22**, 4223 (1980).

[25] D. Pines and P. Nozières, *Theory of Quantum Liquids* [Benjamin, New York, 1966; Mir, Moscow, 1967, 384 pp.].

[26] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. [Pergamon Press, Oxford, 1977; Nauka, Moscow, 1974. 752 pp.].

[27] O. F. Panchenko and L. K. Panchenko, J. Electron. Spectrosc. **83**, 21 (1997).

[28] S. A. Komolov, O. F. Panchenko, and V. M. Shatalov, Fiz. Tverd. Tela (St. Petersburg) **34**, 3489 (1992) [Sov. Phys. Solid State **34**, 665 (1992)].

Translated by R. M. Durham

Current in a high-current planar diode with a discrete emitting surface

S. Ya. Belomyttsev, S. D. Korovin, and I. V. Pegel'

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

(Submitted March 2, 1998)
Zh. Tekh. Fiz. **69**, 97–101 (June 1999)

The dependence of the current on emitter size is obtained for a high-current planar diode with a discrete emitting surface. It is shown that if the distance between the emitters appreciably exceeds their size, the dependence of the current on the ratio of the emitter size to the diode gap is a power dependence with an exponent of 3/2. The voltage dependence of the current obeys the ‘‘three-halves’’ law up to higher voltages than that for a planar diode with a homogeneous emitting surface. © *1999 American Institute of Physics.* [S1063-7842(99)01806-1]

INTRODUCTION

Planar vacuum diodes with explosive emission cathodes[1] are used to form high-current relativistic electron beams with large cross sections (tens and hundreds of square centimeters). These beams are widely used in high-power, virtual-cathode, microwave pulse generators to generate x-ray pulses and also for technological applications.

In many applications, especially for the generation of microwave radiation, an important condition is that the beam current and electron energy should remain constant during the pulse, whose duration is usually a few or tens of nanoseconds. In addition, for high-current planar diodes the impedance typically decreases with time.[2] This leads to an increase in current and, as a result of a mismatch between the diode and the pulse source, causes a drop in the diode voltage and the electron energy of the beam (Fig. 1).

The variation of the diode impedance with time is caused by the evolution of the emitting surface at the cathode. One reason for this is the expansion of the cathode explosive emission plasma. Initially, the emitting surface is a set of isolated plasma formations or emission centers. With time the size of the centers increases, leading to an increase in the current extracted from them. For metal and graphite cathodes, the characteristic plasma expansion velocity is $v_{\mathrm{pl}} \sim 2\times10^6$ cm/s (Ref. 3). The mechanism responsible for this expansion is studied in the present paper. Another reason for the evolution of the emitting surface is the change in the number of emission centers with time, whose influence is not considered here.

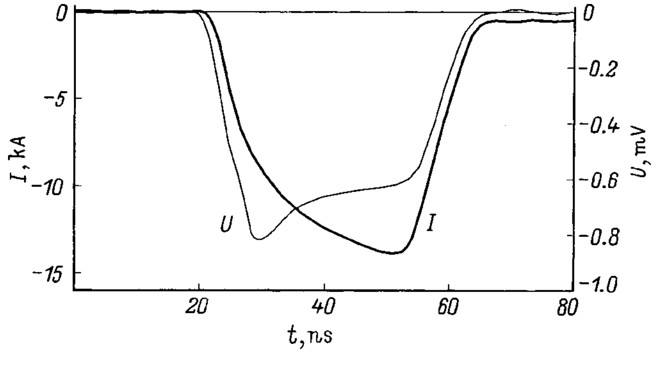


FIG. 1. Typical time dependence of current and voltage in the planar diode of the SINUS-700 high-current, pulse-periodic, electron accelerator using a metal–insulator cathode.

In the present study we investigate the current in a diode with planar electrodes as a function of the size of the emission centers (subsequently called emitters). The solution of this problem is not only of independent interest but is also required to take into account the discreteness of the emitting surface in high-current diodes with cathodes having more complex configurations.

SIMILARITY RELATIONS FOR THE ELECTRON CURRENT IN A DIODE

We shall analyze a steady-state electron flux in a diode of arbitrary configuration having an emitter of unbounded emitting capacity at the cathode. We shall assume that the initial velocity of the emitted electrons is zero. We shall postulate that the particle motion in the gap is nonrelativistic $\Gamma-1 \ll 1$, where $\Gamma = 1 + eU/mc^2$, and U is the diode voltage. We shall assume that no external magnetic field exists and we shall neglect the influence of the self-induced magnetic field on the electron motion. For simplicity, we suppose that the electron beam can be described by the single-flow hydrodynamic model. The parameters of the system, i.e., the electron velocity $\mathbf{v}$, the charge density ρ, and the electric field potential φ are then related by the system of equations

$$(\mathbf{v}\nabla)\mathbf{v} = -\frac{e}{m}\nabla\varphi, \quad \Delta\varphi = -4\pi\rho, \quad \mathrm{div}\,(\rho\mathbf{v}) = 0$$

with the boundary conditions at the cathode $\mathbf{v}|_c = 0$, $\varphi|_c = 0$, $\nabla_t\varphi|_c = 0$ (at the nonemitting surface), $\nabla\varphi|_{\mathrm{em}} = 0$ (at the emitting surface) and at the anode $\varphi|_a = \varphi_0 = U$ (where U is the diode voltage), $\nabla_t\varphi|_a = 0$ (the subscript t denotes the component tangential to the surface).

Denoting by L the characteristic linear dimension of the system, we introduce the dimensionless variables and the differential operators $\tilde{\mathbf{r}} = \mathbf{r}/L$, $\tilde{\mathbf{v}} = \mathbf{v}(e\varphi_0/m)^{-1/2}$, $\tilde{\varphi} = \varphi/\varphi_0$, $\tilde{\rho} = \rho L^2/\varphi_0$, $\tilde{\nabla} = L\nabla$, $\tilde{\Delta} = L^2\Delta$. In the new notation the system of equations and the boundary conditions have the form

$$(\tilde{\mathbf{v}}\tilde{\nabla})\tilde{\mathbf{v}}=-\tilde{\nabla}\tilde{\varphi},\quad \tilde{\Delta}\tilde{\varphi}=-4\pi\tilde{\rho},\quad \operatorname{div}(\tilde{\rho}\tilde{\mathbf{v}})=0, \tag{1}$$

$\tilde{\mathbf{v}}|_c=0$, $\tilde{\varphi}|_c=0$, $\tilde{\nabla}_t\tilde{\varphi}|_c=0$, $\tilde{\nabla}\tilde{\varphi}|_{\mathrm{em}}=0$, $\tilde{\varphi}|_a=1$, $\tilde{\nabla}_t\tilde{\varphi}|_a=0$.

The boundary conditions no longer contain φ_0. Consequently, the functions $\tilde{\mathbf{v}}(\tilde{\mathbf{r}})$, $\tilde{\rho}(\tilde{\mathbf{r}})$, being the solution of Eq. (1), also do not depend on φ_0. The form of these functions is determined only by the shape of the electrodes and does not depend on their absolute dimensions. The current density is

$$\mathbf{j}(\mathbf{r})=I_a(e\varphi_0/mc^2)^{3/2}L^{-2}\tilde{\mathbf{j}}(\tilde{\mathbf{r}})=\sqrt{e/m}\,\varphi_0^{3/2}\tilde{\mathbf{j}}(\tilde{\mathbf{r}}),$$

where $\tilde{\mathbf{j}}=\tilde{\rho}\tilde{\mathbf{v}}$ and $I_a=mc^3/e\approx 17\,\mathrm{kA}$ is the Alfvén current.

The total current in the system is given by

$$I=I_a(e\varphi_0/mc^2)^{3/2}F=\sqrt{e/m}\,\varphi_0^{3/2}F, \tag{2}$$

where the form factor $F=\int_S\tilde{\mathbf{j}}(\tilde{\mathbf{r}})d\tilde{\mathbf{S}}$ is only determined by the relative sizes of the electrodes and remains constant when all the linear dimensions in the diode vary proportionately. This last cross section is taken over any surface containing the beam cross section, such as the emission surface. It follows from expression (2) that the voltage dependence of the current obeys the ''three-halves'' law and the current remains constant when all the linear dimensions in the diode vary proportionately.

It is easy to see that these results apply to a hydrodynamic model with a finite number of fluxes greater than one (a numerical simulation showed that in the particular case of a hemispherical emitter examined below, the electron trajectories intersect each other and the number of fluxes is two).

CURRENT IN A DIODE WITH A SINGLE EMITTER

We shall consider a planar diode with a gap D having a hemispherical emitter of radius R positioned on the cathode plane (Fig. 2a). The emitter has an unbounded emitting capacity. If the transverse dimension of the electrodes considerably exceeds D then, since the emitter shape is defined, the form factor of the system F is merely a function of the parameter R/D. Thus, in the nonrelativistic approximation we have

$$I=\sqrt{e/m}\,\varphi_0^{3/2}F(R/D).$$

In order to determine the form of the function F, we shall use the smallness of the emitter radius compared with the diode gap $R/D\ll 1$. Quite clearly, when the emitter radius tends to zero, the current and thickness of the electron flux also tend to zero and the electron flux weakly perturbs the electric field in the gap. The potential distribution in the diode is close to linear $\varphi(z)\approx\varphi_0 z/D$. Numerical calculations confirm that in the range of R/D ratios of practical importance this is in fact the case (Fig. 3). In a diode with a linear potential distribution the current obviously remains constant when the voltage and length of the diode gap vary proportionately. Consequently, we obtain $\varphi_0^{3/2}D^{-3/2}=\mathrm{const}$ and $F\propto(R/D)^{3/2}$. Hence the current for a single emitter may be written as

$$I=\beta\sqrt{e/m}(RU/D)^{3/2}, \tag{3}$$

where the dimensionless factor β is only determined by the emitter shape.

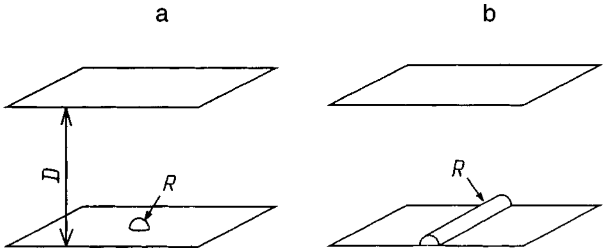


FIG. 2. Diode with point (a) and linear (b) emitters.

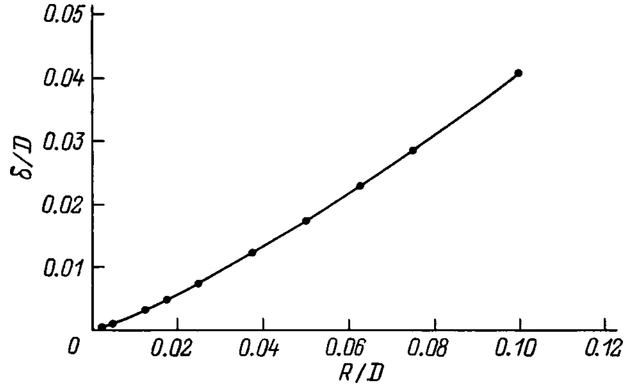


FIG. 3. Shift of the surface with the potential $U/2$ relative to the center of the diode gap as a function of the radius of a hemispherical emitter (δ is the absolute value of the shift). Numerical calculations for $D=4$ cm and $U=500$ kV.

Numerical calculations performed using the SuperSAM program[4] for a hemispherical emitter confirmed this dependence and yielded the coefficient $\beta\approx 0.47$ (Fig. 4).

Similar reasoning yields the conclusion that for ''linear'' emitters (such as semicylindrical, Fig. 2b) the current per unit length of the emitter depends on its radius as $dI/dl \propto R^{1/2}D^{-3/2}$.

We especially note that for small emitters ($R/D\ll 1$), the voltage dependence of the current shows a relativistic deviation from the three-halves law when the voltage is approximately D/R times greater than that for a homogeneous planar diode. This is because the emitter current is determined

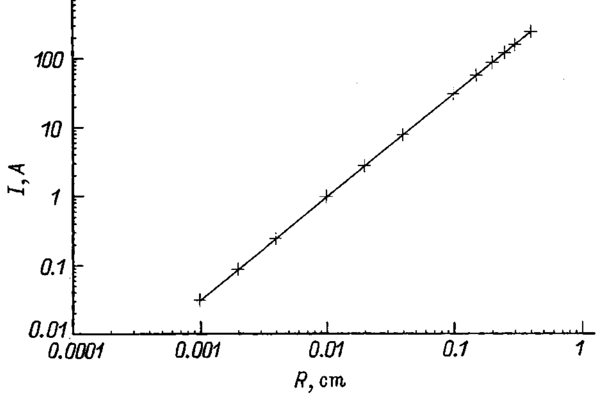


FIG. 4. Current versus radius of a single hemispherical emitter with $D=4$ cm and $U=500$ kV: crosses — numerical calculations, line — calculations using formula (3) for $\beta=0.47$.

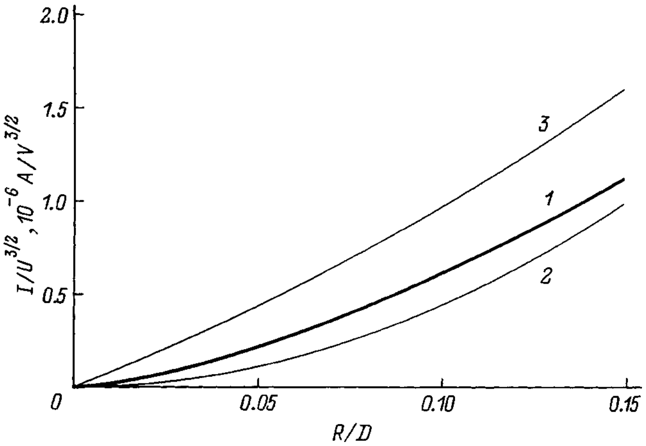


FIG. 5. Perveance of a planar diode with a single hemispherical emitter as a function of the ratio of the emitter radius to the gap: *1* — using formula (3), *2* — from Ref. 5, and *3* — from Ref. 6.

by the space charge in its immediate vicinity whose dimension is of the order R. Even if the total diode voltage U is relativistic, the electron energy in the vicinity of the emitter, which is given by $\sim eUR/D$, may be much lower than the relativistic values. This is responsible for the nonrelativistic law of current behavior. Thus, the condition for validity of expression (3) has the form

$$(\Gamma-1)\frac{R}{D}\ll 1. \tag{4}$$

If the diode voltage is sufficiently high that the electrons become relativistic near the emitter, the voltage dependence of the current tends to linear and the exponent of R is reduced by 1/2. It is easy to see that in this case, for linear emitters the current does not depend on R.

Condition (4) is satisfied in most cases of practical importance, i.e., for plasma expansion velocities of $\sim 10^6$ cm/s, pulse durations between a few and tens of nanoseconds, diode gaps of a few centimeters, and voltages of a few megavolts (Fig. 8).

We stress that the exponent 3/2 in expression (3) is asymptotically exact in the limit of the small ratio R/D. We compare the current obtained from formula (3) with other known results (Fig. 5). The method of specific capacitances[5] gives a quadratic dependence of the current on the parameter R/D. It should be noted that this method has not been sufficiently well substantiated and gives incorrect results for small ratios R/D. An interpolation formula[6] based on the results of numerical calculations gives a linear dependence of the current on R/D for small values of this parameter. Quite clearly, the results of the numerical calculations for $R/D=0.05$ and 0.1 are not completely accurate. It should be noted that Djogo and Gross[6] gave an incorrect representation of the data from Ref. 7 relating to the current of a spherical emitter at the apex of a tip.

CURRENT IN A PERIODIC EMITTING STRUCTURE

Let us assume that emitters of radius R are arranged as a periodic structure separated by the distance p (we shall not specify the shape of the cell). We introduce the notation $x=2R/p$ and $y=p/D$. Thus, for $x=1$ we have a planar diode with continuous uniform emission for which the Child–Langmuir law is satisfied. We write this law in an approximate form which is asymptotically exact in the nonrelativistic and ultrarelativistic limits and gives a deviation of less than 1% from the exact solution[8] in the intermediate range of voltages

$$J_{\text{pd}}=\frac{I_a}{2\pi D^2}(\Gamma^{2/3}-\Gamma^{n})^{3/2},$$

$$n=2/3-2/9^{2/3}\approx 0.204425. \tag{5}$$

We shall now consider another case of practical importance when the distance between the emission centers is considerably less than the diode gap. In this case, provided that the condition $x\ll 1$ is satisfied, the current for each emitter is described by expression (3), i.e., the emitters have little relative influence. If the size of the emitters decreases without bound, the current in the system tends to zero and the space charge of the electron flux in the diode gap has no influence on the impedance of the emitters, i.e., they do not influence each other.

Thus, a general expression which would determine the average current density in the diode over the entire range between $x=0$ and $x=1$ should give expressions (3) and (5) in these limiting cases. This property is obtained for the following formula which matches the solutions for a single emitter and a planar diode with a continuous emitting surface

$$j=j_{\text{pd}}/A,\quad A=1+\frac{f(x)y^{1/2}}{2^{3/2}\beta x^{3/2}}\left(\frac{\Gamma^{2/3}-\Gamma^{n}}{\Gamma-1}\right)^{3/2}. \tag{6}$$

The weighting function $f(x)$ satisfies the conditions $f(0)=1$, $f(1)=0$ and its specific form depends on the shape of the periodic cell containing the emitter.

We shall consider a hexagonal cell. For a numerical simulation this can be approximated by a cylinder of radius $p/2$, having defined the Neuman condition for the electrostatic potential and the specular reflection condition for the electron trajectories at its side surface (Fig. 6). In this way,

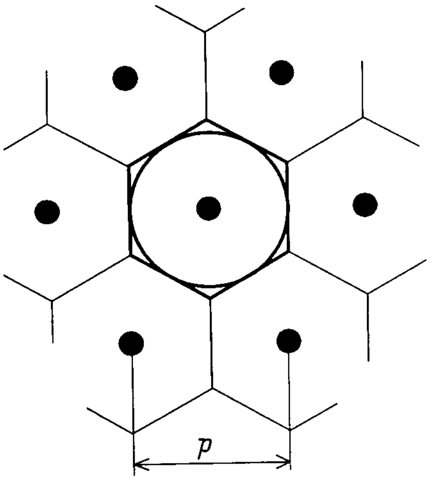


FIG. 6.

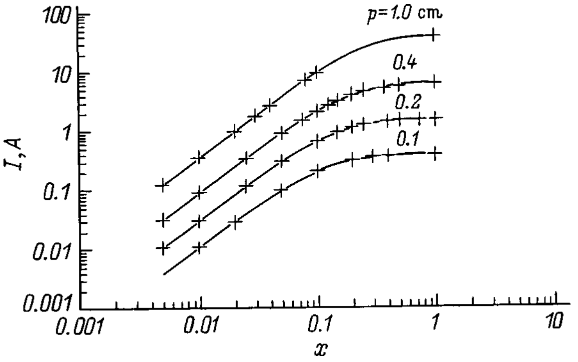


FIG. 7. Current for a hemispherical emitter in a periodic system as a function of the ratio of the emitter diameter to the cell period for $D=4$ cm and $U=500$ kV. Crosses — numerical calculations, curves — calculations using formula (6).

the calculation problem is converted from three-dimensional to axisymmetric two-dimensional. Numerical calculations made using the SuperSAM program showed that in this case, the weighting function can be approximated by $f(x)\approx 1-x^{1-x}$. Then, the dependences $I(x)$ for the current per emitter obtained using formula (6) and as a result of the numerical calculations agree to within 5% (Fig. 7).

Figure 8 gives relative voltage dependences of the perveance of a diode with hemispherical emitters calculated numerically and using formula (6). It is easy to see that the smaller the relative size of the emitter *x*, the higher the voltages for which the nonrelativistic dependence of the current on voltage $I\propto U^{3/2}$ typical of a single emitter is conserved. The difference between the analytic and numerical results at extremely high voltages can be explained by the fact that the condition for no relativistic electron motion ceases to hold near the emitter (quite clearly the constraint for a periodic system of emitters should differ from condition (4) since it contains the parameter p).

The numerical calculations showed that as the parameter *x* decreases, the dependence of the current in a periodic system on the diode gap tends to the form $I\propto D^{-3/2}$ typical of a single emitter.

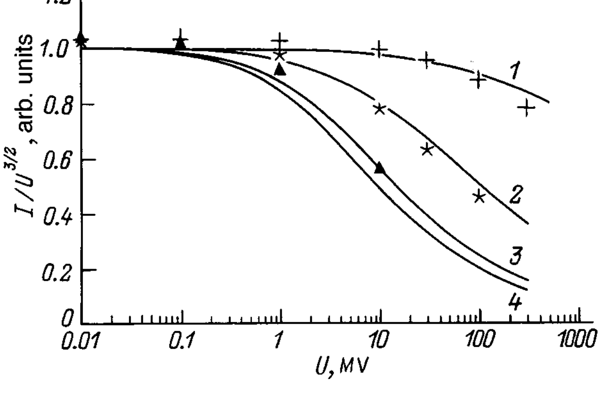


FIG. 8. Comparative voltage dependences of the perveance of a hemispherical emitter in a periodic system with $p=0.1$ cm (*1–3*). Emitter radius, μm: *1* — 5, *2* — 25, and *3* — 100; asterisks — numerical calculations, curves — calculations using formula (6); *4* — relativistic dependence for a homogeneous planar diode (5).

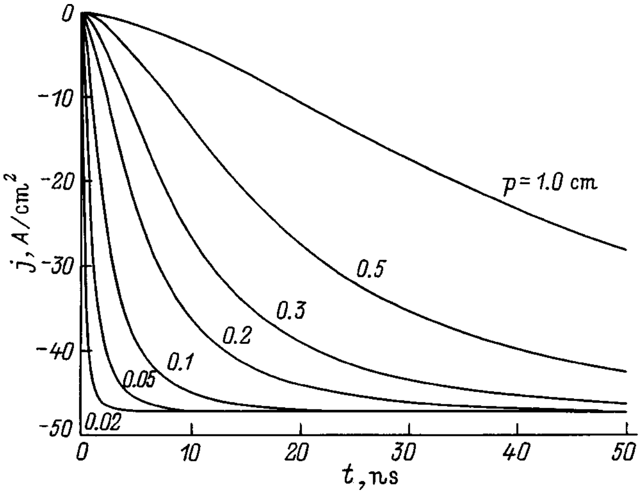


FIG. 9. Time dependences of the current density in a planar diode with a discrete emitting surface calculated using expression (6) for $D=4$ cm and $U=500$ kV. The emitter expansion velocity is 2×10^6 cm/s.

Figure 9 shows calculated time dependences of the average current density for a cathode with hemispherical emitters expanding at the velocity 2×10^6 cm/s for various distances between the emitters. By comparing these dependences with the experimentally measured oscilloscope traces of the current, it is possible to assess the state of the emitting surface at the cathode.

CONCLUSIONS

This theoretical analysis has demonstrated that the microstructure of the emitting surface strongly influences the current in a high-current planar diode.

If the distance between the emitters appreciably exceeds their size, the dependence of the current on the ratio of the emitter size to the diode gap is a power dependence with the exponent 3/2. Moreover, the voltage dependence of the current obeys the three-halves law up to higher voltages than that for a planar diode with a uniform emitting surface.

Quite clearly, the discreteness of the emitting surface and the expansion of the cathode plasma lead to a reduction in the impedance of high-current planar diodes with time. The cathode effects described determine the minimum duration of the resulting current pulse.

However, it should be noted that additional information on the state of the emitting surface at the cathode is required to correctly apply these results to real diodes. The dependences given above were obtained assuming that the number of emission centers remains constant in time and they are equidistant. At the same time, there is no doubt that the emission centers do not appear simultaneously at the cathode. The distribution of the micropoints over the cathode surface and their sizes are not homogeneous. The space charge of a newly formed emission center reduces the electric field intensity in its vicinity,[9] increasing the delay time for the explosion of micropoints situated in this zone. Allowance for the influence of the nonsimultaneous creation of emission centers on the current profile in a high-current diode is a problem for future research.

The authors are grateful to D. I. Proskurovskiĭ for discussions of the results.

[1] S. P. Bugaev, E. A. Litvinov, G. A. Mesyats, and D. I. Proskurovskiĭ, Usp. Fiz. Nauk **115**, 101 (1975) [Sov. Phys. Usp. **18**, 51 (1975)].
[2] N. M. Bykov, V. P. Gubanov, A. V. Gunin *et al.*, in *Proceedings of the Tenth International Pulsed Power Conference*, Albuquerque, NM, 1995, pp. 71–74.
[3] G. A. Mesyats and D. I. Proskurovskiĭ, *Pulsed Electric Discharges in Vacuum* [in Russian], Nauka, Novosibirsk (1984), 256 pp.
[4] D. G. Myakishev, M. A. Tiunov, and V. P. Yakovlev, Int. J. Mod. Phys. A **2B**, 915 (1993).
[5] A. F. Shubin and Ya. Ya. Yurike, Izv. Vyssh. Uchebn. Zaved. Fiz. No. 6, 134 (1975).
[6] G. Djogo and J. D. Gross, IEEE Trans. Plasma Sci. **25**, 617 (1997).
[7] G. A. Mesyats and E. A. Litvinov, Izv. Vyssh. Uchebn. Zaved. Fiz. No. 8, 158 (1972).
[8] H. P. Jory and A. W. Trivelpiece, J. Appl. Phys. **40**, 3924 (1969).
[9] S. Ya. Belomyttsev, S. D. Korovin, and G. A. Mesyats, Pis'ma Zh. Tekh. Fiz **6**, 1089 (1980) [Sov. Tech. Phys. Lett. **6**, 466 (1980)].

Translated by R. M. Durham

Influence of C_{60}-containing additives in lubricant oil on the optimization of wear processes in the boundary friction of metals

D. G. Tochil'nikov and B. M. Ginzburg

Institute of Problems of Mechanical Engineering, Russian Academy of Sciences,
199178 St. Petersburg, Russia

(Submitted February 13, 1998; resubmitted May 20, 1998)
Zh. Tekh. Fiz. **69**, 102–105 (June 1999)

An investigation is made of the influence of fullerene C_{60}-containing additives in lubricant oil on the optimization of the boundary friction process of steel and copper in steel–steel and copper–steel sliding tribocouples. The optimization was characterized by a reduction in the parameter Q introduced in a previous study, which is the probability that any given contact spot between contacting bodies is converted into a debris particle accompanied by a relative shift of the contacting bodies by the average diameter of the contact spot. The experimental data are used to calculate values of the parameter Q for various friction regimes and it is shown that the addition of fullerene to the lubricant oil causes a substantial reduction in Q and for copper this parameter reaches fairly low values similar to those for an arbitrary reference selected previously for classifying solids according to this characteristic. © *1999 American Institute of Physics.* [S1063-7842(99)01906-6]

INTRODUCTION

By applying universal criteria for a comparative assessment of the optimum nature of wear processes, it is possible to compare the antiwear properties of tribocouples and lubricant materials determined in tests under different conditions. Among these universal criteria characterizing the wear process in the boundary friction of solids, mention may be made of the parameter Q proposed in Ref. 1, which is the probability that any contact spot between contacting bodies will be converted into a debris particle accompanied by a relative shift of the contacting bodies by the average diameter of the contact spot $Q=\alpha/n$, where α is the number of contact spots converted into debris and n is the total number of contact spots, whose area is the actual area of the friction spot. We can postulate that an optimum wear regime exists for which Q will have a minimum, which then leads to minimal values of the wear intensity. A reduction in Q during the wear process may be taken as optimization of the wear process.

Kozyrev and Ginzburg[1] proposed a method for the empirical determination of optimum boundary friction regimes for which the wear intensity of the sample material has a minimum and $Q=Q^{\text{opt}}$. For one of the most antifriction materials, babbitt metal, the value of Q^{opt} was $\approx 2\times 10^{-10}$. These authors also showed that Q can be calculated for given test conditions if data are available on the linear wear intensity, the material hardness, and the contact pressure. Published data were used to calculate values of Q for various materials (different types of bronze, a Teflon-4 + graphite + MoS_2 composite, St-45 steel, cast irons, and titanium carbide) for specific test conditions. In all cases, the values of Q were between one and three orders of magnitude higher than 2×10^{-10}. In particular, for St-45 steel in contact with bronze, the authors obtained $Q\approx 10^{-7}$ and for bronze $Q\approx 10^{-8}$. Values of Q = (3–4) $\times 10^{-10}$ were only obtained for high-strength cast irons in contact with tin-plated and chrome-plated piston rings in engines lubricated with engine oils. Thus, the value $Q=2\times 10^{-10}$ can be used for the time being as an arbitrary reference for different materials, in the same way that the hardness of diamond is used as a reference point to estimate the hardness of solids.

However, Ginzburg *et al.*[2–6] showed that the presence of small quantities of fullerene C_{60} in lubricant oil can substantially improve various tribological properties, especially the antiwear properties of steel and copper in steel–steel and steel–copper sliding tribocouples. It would be interesting to determine the extent to which fullerene C_{60} helps to optimize the friction process, i.e., to study how C_{60}-containing additives in industrial oil influence the values of Q for steel and copper and their variations under different test conditions. It would also be interesting to compare these values with the reference point, and these are the aims of the present study.

MATERIALS

Fullerene C_{60} was separated chromatographically (with a 96–98% content) from a fullerene extract (mainly C_{60} and C_{70}) in the V. P. Budtov Laboratory (''Khromotron'' Project in the Russian Scientific and Technical Program ''Fullerenes and Atomic Clusters,'' 1994). The extract was separated from fullerene soot obtained in an electric arc plasma in the G. A. Dyuzhev Laboratory (''Arc'' project under the same program).

The lubricants used were I-40-A commercial industrial oil (I-G-A-68 according to GOST 17479-87), a mixture of I-40 A and 5% fullerene C_{60}, and a mixture of I-40-A and 5% fullerene soot. The fullerene soot contained around 5.5% C_{60} and 1.5% C_{70}. The additives were added mechanically to the oil in the form of finely dispersed powders. Before these lubricant mixtures were used, they were mixed vigorously to

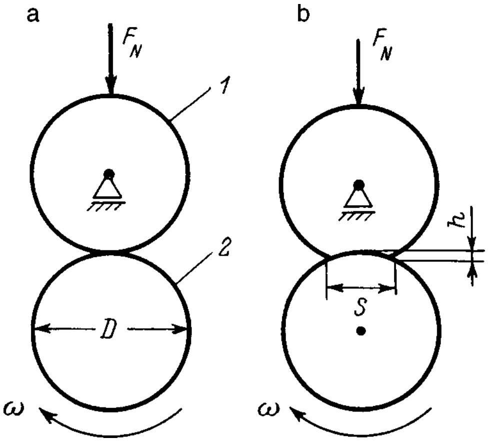


FIG. 1. Schematic showing friction tests: *1* — fixed roller (sample), *2* — moving roller; a — initial state with linear friction contact, b — state of friction contact after formation of wear groove of area *S*.

achieve a more uniform particle distribution in the oil, although the residual inhomogeneity could be one reason for the spread of the data obtained.

TEST METHOD

The investigations were carried out using a standard 2070 SMT-1 roller friction machine using tribocouples comprising a fixed upper roller and a rotating lower roller (Fig. 1). The lower roller, 46 mm in diameter and 16 mm wide, was made of wear-resistant 18Kh2NChMA steel (GOST 4543-71) and rotated at a speed $\omega = 400\,\mathrm{min}^{-1}$, which corresponded to a linear sliding velocity of 1 m/s. The sample being tested was either the upper fixed roller (50 mm in diameter and 10 mm wide) made of St-45 steel or 0.25 mm copper foil attached to this roller.

Various normal loads F_N were successively applied to the samples. The tests began with a linear friction contact at a load $F_N = 100\,\mathrm{N}$ (Fig. 1a). The duration of the testing at each load was 300 s and lubrication was provided by a single drop of oil ($\sim 0.05\,\mathrm{g}$) before each change in load. After every 300 s the carriage of the friction machine was raised and the geometric parameters of the wear spot were determined using a measuring magnifier. These values were used to calculate the average area of the wear spot S, the depth of the wear groove h (Fig. 1b), and the bulk wear V_W (Ref. 7). The area of the wear spot S_e determined at the end of the tests under a given load was taken as the area of the initial frictional contact S_i for the next load. Between three and fifteen tests were carried out for each test variant (for each load and fixed type of lubricant) depending on the spread, and the arithmetical means of the measured values were then calculated. In order to monitor the reproducibility after every 10–12 tests, tests were carried out using I-40 A base oil.

The value of Q was calculated using a formula proposed in Ref. 1

$$I_h = 3(P_N Q/H_B)\cdot(g/d), \tag{1}$$

where I_h is the linear wear intensity of the sample, P_N is the nominal contact pressure, H_B is the Brinell hardness, g/d is the ratio of the average height of a debris particle to the average diameter of the contact spot or the diameter of the debris particle, if the spot has been converted into a debris particle.

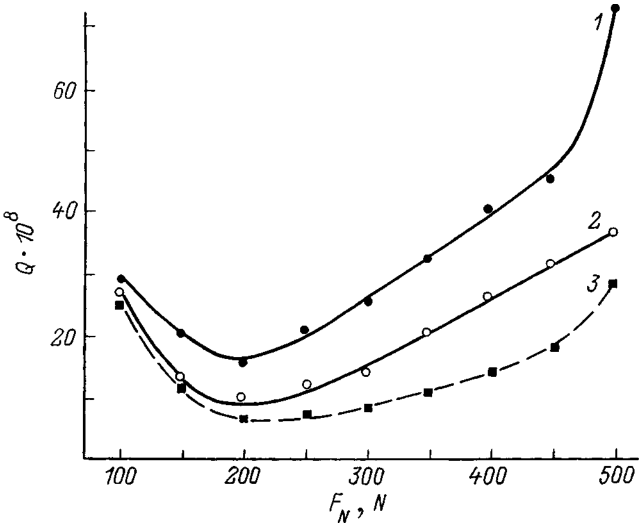


FIG. 2. Dependence of the parameter *Q* on the load for a steel–steel tribocouple. Lubrication with I-40 A base oil (*1*), I-40 A + 5% fullerene soot (*2*), and I-40 A + 5% C_{60} (*3*).

We subsequently assumed that $g/d \approx 1$, i.e., the debris is lumpy, as is frequently observed for metals.

The linear wear intensity for a linear initial contact is defined as $I_h = h/L$, where h is the depth of the wear groove and L is the friction path during testing. The friction path was defined as $L = \pi D \omega t$, where D is the diameter of the moving roller and t is the test time. The nominal contact pressure P_N was defined as F_N/S_{av}, where S_{av} is the average nominal contact area during the test. Special tests showed that our formula $S_{\mathrm{av}} \approx S_i + 0.8(S_e - S_i)$ more accurately describes the change in the contact area than the arithmetical mean of the initial and final areas, since the contact area varies nonlinearly with time, increasing rapidly in the initial instants of application of a new load and then varying more

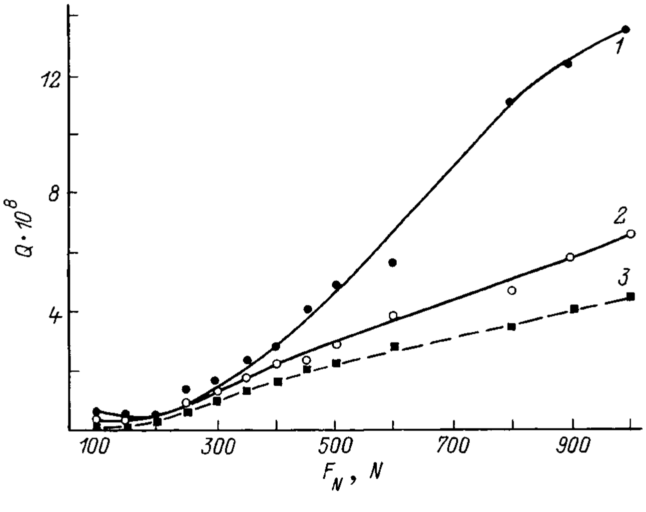


FIG. 3. Dependence of the parameter *Q* on the load for a copper–steel tribocouple. Lubrication with I-40 A base oil (*1*), I-40 A + 5% fullerene soot (*2*), and I-40 A + 5% C_{60} (*3*).

TABLE I. Minimum and maximum values of the linear wear intensity I_h, the parameter Q, and the contact pressures P_N over the entire range of loads for steel and copper in steel–steel and copper–steel sliding tribocouples.

		Steel					Copper			
	Range of average values of P_N, N/mm^2	Linear wear intensity I_h, 10^{-8}		Probability Q, 10^{-8}		Range of average values of P_N, N/mm^2	Linear wear intensity I_h, 10^{-8}		Probability Q, 10^{-8}	
Oil		minimum	maximum	minimum	maximum		minimum	maximum	minimum	maximum
I-40 A	26–38	0.8	4	16	80	60–80	0.15	3	0.5	14
I-40 A+5% fullerene soot	27–50	0.3	2	8	40	80–90	0.08	2	0.3	7
I-40 A+5% C_{60}	28–53	0.3	2	7	30	90–120	0.05	2	0.15	5

smoothly. The Brinell hardness for St-45 steel was taken as 2000 N/mm^2 and that for copper as 800 N/mm^2 (Ref. 8).

Taking into account all the constant values of the parameters in formula (1) given above, the values of Q can be calculated using the formulas

$$Q=2.2\times10^{-3}hS_{\mathrm{av}}/F_N \quad \text{for a steel–steel tribocouple,} \tag{2}$$

$$Q=0.9\times10^{-3}hS_{\mathrm{av}}/F_N \quad \text{for a copper–steel tribocouple,} \tag{3}$$

where the values of h are in millimeters, S_{av} is in square millimeters, and F_N is in newtons.

The relative mean square error of the arithmetical means of Q was 7–10%.

TEST RESULTS

Note that a specific feature of friction tests is that in the elastic contact region, the tribological characteristics depend not only on the pressure applied to the contact, but also to a greater extent on the absolute value of the applied force.[8–10] As the applied force increased, the contact area also increased so that the contact pressure did not change very significantly although the values of Q varied substantially. Figures 2 and 3 give the results of calculating Q as a function of the applied force for tests on steel–steel and copper–steel tribocouples, respectively and Table I gives the maximum and minimum values of the parameter Q and the linear wear intensity, more familiar to tribologists, over the entire range of applied loads. The range of average pressures on the contact is also given.

We shall examine the behavior of these curves and how they change when fullerenes are added to the oil for steel–steel tribocouples (Fig. 2). At the initial run-in stage ($F_N=100$ N) the values of Q for the different oils differ negligibly. When the load is increased from $F_N=100$ to 200 N, a fall in Q is observed in all cases, i.e., so-called run-in or optimization of the wear process takes place. However, for the fullerene-containing oils the optimization process is more intensive. For example, for the base oil in the range of loads used, the value of Q is halved, for oil with soot additives it is reduced threefold, and for oil with fullerene C_{60} additives it is reduced fourfold. When the load is then increased to 300 N, the value of Q remains almost constant for oil with C_{60} additives, whereas for the other oils it increases appreciably. At loads of 300–500 N a substantial increase in the values of Q is observed and for the base oil a sharp increase can be identified near 500 N, which is typical of scoring. However, for the oil with added C_{60} the value of Q at these high loads is 2.7 times lower than that for the base oil.

The lowest value of Q for a steel–steel tribocouple is observed at 200 N and is $\approx 7\times10^{-8}$ for pure C_{60} additives,, which is more than two orders of magnitude higher than the standard value for babbitt metal. A slightly reduced effect is achieved with added fullerene soot: the lowest value of Q under these conditions is $\approx 9\times10^{-8}$. The addition of fullerenes to the oil reduces the linear wear intensity by a factor of 2–3 and its lowest value is 0.3×10^{-8} (see Table I).

For the copper–steel tribocouples the curve $Q(F_N)$ only exhibited a weak minimum (a 20% reduction) for the base oil (Fig. 3). For the fullerene-containing additives the run-in is very fast for the first friction contact at a 100 N load and thus no fall in Q could be recorded. Nevertheless, the value of Q for the fullerene-containing oil at this load is 2.5–4.5 times lower than that for the base oil (see Table I). Subsequently, an increase in Q is observed for all loads but in all cases, the fullerene-containing additives give a significant reduction in Q compared with the base oil. Compared with the steel–steel tribocouple, the copper–steel tribocouple exhibits better antiwear properties under all conditions and the lowest values of I_h and Q achieved for pure C_{60} additives are 5×10^{-10} and 15×10^{-10}, respectively (see Table I). This last figure is of the same order of magnitude as the reference value for babbitt metal.

To sum up, there is no doubt that fullerene additives in oils help to optimize the friction surfaces of steel and copper and reduce their wear. The parameter Q characterizing the surface optimization has an extremely low value for copper close to the arbitrary reference value characterizing the minimum possible wear.

This work was carried out under the Russian Scientific-Technical Program ''Fullerenes and Atomic Clusters'' (''Tribol'' Project).

[1] Yu. P. Kozyrev and B. M. Ginzburg, Zh. Tekh. Fiz **68**(4), 48 (1998) [Tech. Phys. **43**, 392 (1998)].

[2] B. M. Ginzburg *et al.*, Pis'ma Zh. Tekh. Fiz **21**(22), 62 (1995) [Tech. Phys. Lett. **21**, 933 (1995)].

[3] B. M. Ginzburg *et al.*, Pis'ma Zh. Tekh. Fiz **21**(23), 34 (1995) [Tech. Phys. Lett. **21**, 964 (1995)].
[4] B. M. Ginzburg, D. G. Tochil'nikov, and V. P. Bulatov, Trenie Iznos **18**, 235 (1997).
[5] B. M. Ginzburg, V. A. Krasnyĭ, Yu. P. Kozyrev *et al.*, Pis'ma Zh. Tekh. Fiz **23**(15), 1 (1997) [Tech. Phys. Lett. **23**, 575 (1997)].
[6] B. M. Ginzburg *et al.*, Trenie Iznos **18**, 523 (1997).
[7] V. P. Bulatov *et al.*, Trenie Iznos. **15**, 1009 (1994).
[8] I. V. Kragel'skiĭ, M. M. Dobychin, and V. S. Kombalov, *Principles of Friction and Wear Calculations* [in Russian], Mashinostroenie, Moscow (1977), 440 pp.
[9] I. N. Nesterova and A. F. Gusev, Tech. Phys. Lett. **16**, 1152 (1995).
[10] B. M. Ginzburg *et al.*, Pis'ma Zh. Tekh. Fiz **21**(1), 41 (1995) [Tech. Phys. Lett. **21**, 18 (1995)].

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Determination of the temperature distribution in liquids and solids using holographic interferometry

V. V. Bat'kovich, O. N. Budenkova, V. B. Konstantinov, O. L. Sadov, and E. A. Smirnova

A. F. Ioffe Physicotechnical Institute, 194021 St. Petersburg, Russia

(Submitted February 27, 1998)
Zh. Tekh. Fiz. **69**, 106–111 (June 1999)

It is shown that holographic interferometry can be applied to solve two problems: heating of a glass plate by a complex heat source and nonisothermal flow of a submerged jet around a wedge. The process of isolating and numbering the skeletal lines on the interferograms is automated and direct calculations are made of the temperature fields. © *1999 American Institute of Physics.* [S1063-7842(99)02006-1]

INTRODUCTION

Two processes are considered. The first involves the flow of a submerged jet around a wedge where the water temperature in the jet differs from the temperature of the water filling the container, which also varies with time. The second involves the propagation of heat in a glass plate heated by a complex heat source. Numerical or analytic solutions of these problems present difficulties because of the complex boundary conditions. In the first process, for instance, allowance must be made for the reflection of the jet from the walls because of the bounded volume. In the second problem, even if we know the nature of the heat release of the heating element, we need to estimate the quantity of heat entering the plate. It is technically difficult to insert thermocouples into these objects and would distort the real process so we used holographic interferometry.

The quantitative characteristics of a process investigated by holographic interferometry are obtained by analyzing the interference fringes.

The intensity distribution on the interference pattern has the form[1]

$$I(x,y)=I_1(x,y)+I_2(x,y)+2\sqrt{I_1 I_2}\cos\varphi(x,y),$$

where $I_1(x,y)$ is the intensity of the wave scattered by the object before the onset of the process at point (x,y), $I_2(x,y)$ is the intensity of the wave scattered by the object during the process at point (x,y), and $\varphi(x,y)$ is the phase difference between the first and second waves at the point (x,y).

The spatial frequency of the fringes varies as a result of a change in the refractive index of the object. A shift of the fringes may also be caused by a change in the path length of the beam in the medium as a result of its expansion or compression. The shift of the interference fringes relative to their initial position is used to calculate the change in the phase distribution of the wave scattered by the object which is related to the interesting characteristics of the process

$$k=\frac{\Delta\varphi}{2\pi}=\frac{1}{\lambda}\left(\int_S n(x,y,z)\,dl-l_0 n_0\right), \tag{1}$$

where k is the number of fringes by which the interference pattern was shifted, $\Delta\varphi$ is the phase difference at different times, λ is the wavelength of the radiation used, $n(x,y,z)$ is the refractive index at the point (x,y,z), S is the beam trajectory, l_0 is the path length of the beam in the object, and n_0 is the refractive index in the absence of any inhomogeneity.

The fringe shift can be traced by measuring the coordinates of the characteristic points of each interference fringe. The characteristic points are those points at which the intensity has extreme values, i.e., the phase difference[2]

$$\varphi(x,y)=\pi n;\quad n=0,\pm 1,\pm 2,\ldots. \tag{2}$$

A set of characteristic points of one type belonging to one interference fringe is a line of zero thickness, i.e., a skeletal or core line.

On an ideal interference pattern it is easy to isolate the characteristic points. In reality, however, the intensity distribution on the interference pattern reflects the presence of various types of noise in addition to the useful signal and may be described as

$$I_r(x,y)=\omega(I_1^2+I_2^2+2I_1I_2\cos\varphi(x,y))+N(x,y),$$

where ω is a function describing the nonlinearity of the response of the recording system relative to the optical intensity, and $N(x,y)$ is the noise recorded together with the signal (background noise).

The presence of noise in the measuring system leads to the appearance of spurious local extrema located at points (x,y) where condition (2) is not satisfied.

If the statistical characteristics of the noise and the useful signal are known, we can construct an optimum filter which when applied to the initial image will give a signal as close as possible to the useful signal according to the criterion of minimum rms error. A priori information on the noise characteristic of the measuring system is not generally available and attempts to estimate the frequencies corresponding to the useful signal directly from the image spectrum have been unsuccessful. Thus, most studies concerned with the automatic processing of interference patterns have examined various methods of filtering the image, i.e., modifications of known filters and the development of new ones. The processing of interference patterns has been automated most comprehensively only in specific cases.[3,4]

Here we present an algorithm for obtaining quantitative information on a particular process using interference patterns which includes preliminary processing of the initial image, isolating the skeletal lines, and calculating the shift of the interference fringes which can be used as the basis for a Cu program.

PROCESSING OF INTERFERENCE PATTERNS

The initial image (interference pattern) is represented as an $M \times N$ numerical matrix B. Each element has values between 0 and 255 which correspond to the minimum and maximum intensities.

In order to correctly isolate the characteristic points, the image was preliminarily processed using a moving average filter[5] where the intensity at a particular point is replaced by some average of the values of its neighboring elements. This filter was used because of its simplicity and also because the following analysis did not require a knowledge of the true intensity. By repeatedly applying the filter (between three and five times), it was possible to avoid most of the spurious local maxima and minima present in the image as a result of noise. The result of applying the moving-average filter was written into a new matrix B_1.

From among the various filters specially developed for processing interference patterns, we selected a spin filter.[6] When a spin filter is applied, a direction is selected for each image point and the intensity at this particular point is replaced by some average of the intensities at neighboring points in this direction. Applying a spin filter to the interference patterns obtained in our measuring system also yielded good results but with slower convergence.

The process of isolating the skeletal lines included several stages. The first involved isolating the characteristic points at which the intensity has an extreme value. On an ideal one-dimensional sinusoidal surface the point at which the sine has a maximum is nonextremal only in one direction. When determining the maxima and minima on a real interference pattern, we assumed that because of the presence of noise such a point can be nonextremal in more than one direction.

The intensity $a_{i,j}$ at a point having the matrix coordinates (i,j) was compared with the intensities at other points within a window of dimensions $n \times n$ centered on the point (i,j), along four principal directions, vertical, horizontal, and two diagonals. If $a_{i,j}$ was higher than the intensities at all other points in any two directions, a pixel with the coordinates (i,j) in a new matrix B_2 was assigned the value 255 (maximum) or if its intensity was lower, this pixel was assigned the value 0 (minimum). In other cases it was assigned the value 128 (background value). This procedure was repeated for each matrix point B_1.

The image produced at the first stage of filtering could not completely avoid spurious extrema and the isolated characteristic points did not form continuous lines of zero thickness. The skeletal lines were isolated from the set of characteristic points at the second stage. A cutting algorithm,[5] which can only be applied to connected regions, was used to isolate the core lines. Gaps in which connectedness was impaired were filled artificially: each background pixel of the image was given values of 0 or 255 if the values of more than two neighboring pixels were 0 or 255, respectively.

After the cutting procedure, the remaining characteristic points formed the core lines. In addition to the skeletal lines corresponding to the true intensity extrema on the interference pattern, the image also exhibited noise in the form of isolated points, short lines, short branches on lines, and also line discontinuities.

The third stage involved determining the assignment of each point of a specific line for convenience in the subsequent processing of the image. We assumed that points belong to the same skeletal line if they have the same type of extremeness (minimum or maximum) and are located at a distance shorter than some predefined value. If the points were positioned at a distance of one pixel or more, the gap was filled with the same extrema. In the program, points belonging to the same line were held in linked lists. This data structure simplified the procedure for removing short branches, whose length does not exceed some predefined length and short lines, whose length does not exceed some predefined value.

After we had performed these procedures, some of the skeletal lines remained broken and thus we carried out an additional line connecting process. The discontinuities were eliminated as follows: parts of the lines were interconnected if the straight line linking their ends did not intersect a line made up of the other type of extremum. At this point the procedure for isolating the skeletal lines was completed. An important advantage of this method of isolating the skeletal lines is that it is independent of the configuration and orientation of the fringes.

In order to calculate the shift of the fringes relative to their initial position, these were numbered on the initial interference pattern and then suitably numbered on the subsequent ones. A necessary condition for the successful operation of the fringe-numbering program is that there should be at least one continuous fringe passing through the entire image. This fringe was numbered zero. Fringes positioned to the right were given positive numbers in order while those to the left were given negative numbers, after which the numbers of the corresponding fringes became the values of the points assigned to the skeletal lines. Points not assigned to skeletal lines were given fractional values obtained from interpolating the fringe numbers over the entire field. The shift of each interference fringe was calculated simply by subtracting the two matrices containing the fringe numbers: $B_{3,0}$ (at the initial time) and $B_{3,t}$ (at the instantaneous time t).

EXPERIMENTS

The technique described was used to process various interference patterns obtained using a holographic interferometer with recording on a photothermoplastic carrier and recording of the information on a television. The image was fed to a monitor which allowed the evolution of the process to be observed in real time. At the times of interest the image

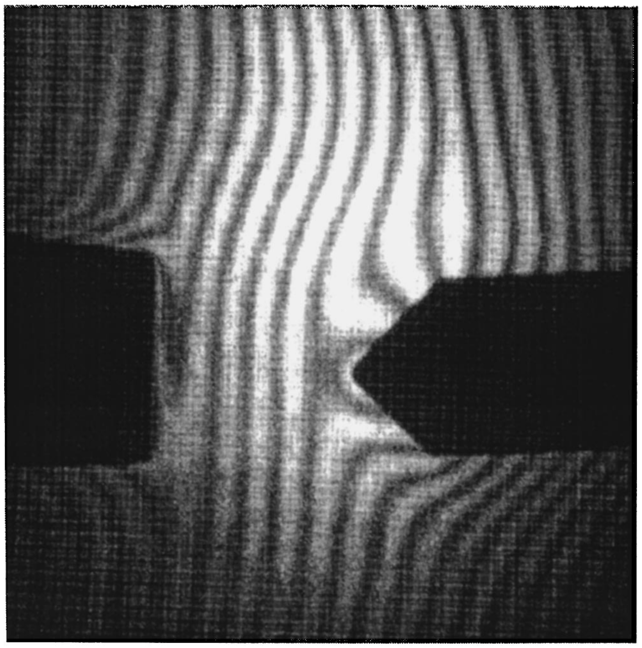


FIG. 1.

was fed into a computer using a frame grabber The size of the field being studied was 3×3 cm and the field resolution was 512×512 pixels.

The following parameters were defined to analyze the interference patterns: the moving-average filter was used three times, the size of the window for isolating the characteristic points was 7×7, points located at a distance of less than ten pixels were assumed to belong to one fringe, the length of the fringes being removed was less than 11 pixels, and the length of the branches being removed was less than five pixels.

In order to solve the formulated problems we assumed that the object was homogeneous in the direction of illumination and formula (1) then becomes

$$k=\frac{1}{\lambda}(n(x,y)l-n_0(x,y)l_0).$$

FLOW OF A SUBMERGED JET AROUND A WEDGE

Water initially begins to flow into a liquid-filled cell containing a wedge. The initial temperatures of the water jet flowing through a nozzle inside the cell, the wedge, and the nozzle are different. The temperature of the jet decreases during the process. The bounded volume leads to appreciable mixing of the water in the cell and the fluxes reflected from the walls begin to have a negative influence on the temperature field. The initial interference pattern is shown in Fig. 1.

Figure 2 shows isolated skeletal lines: the flow velocity of the jet is 5 cm/s, the size of the cell 5×3×1.5 cm, and the nozzle diameter 0.4 cm.

The change in the refractive index was calculated using the formula

$$\Delta n(x,y)=\frac{\lambda k(x,y)}{l_0}. \tag{3}$$

The bulk temperature field was calculated for times when return flows have a significant influence. The results are plotted in Figs. 3 and 4.

TEMPERATURE FIELD IN GLASS PLATE

Initially heat begins to be released in a heating element, its nature being unknown. The released heat is dissipated partially into the ambient medium, some enters the glass plate and then propagates further by heat conduction and radiation. According to our estimates, the shift of the fringes caused by a change in the refractive index is two orders of magnitude lower than that caused by expansion of the glass so that the change in the refractive index was neglected. The temperature field in the glass plate was calculated using the formula

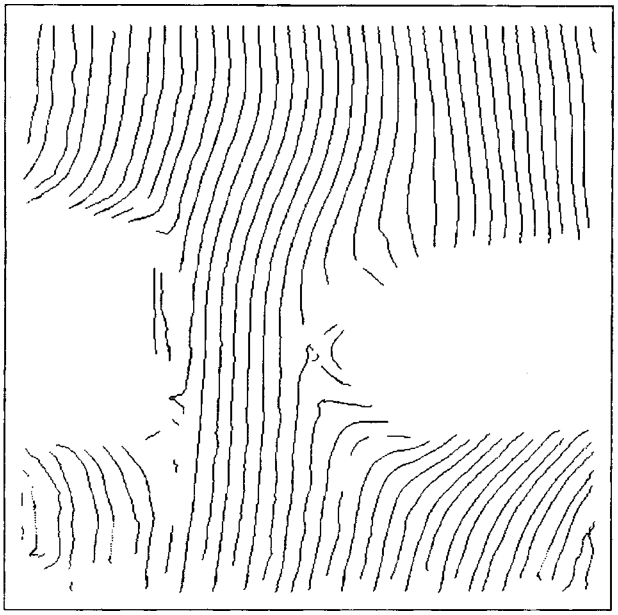


FIG. 2.

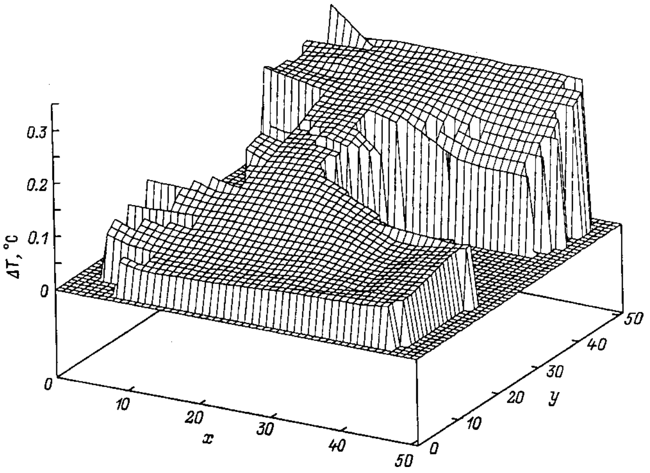


FIG. 3.

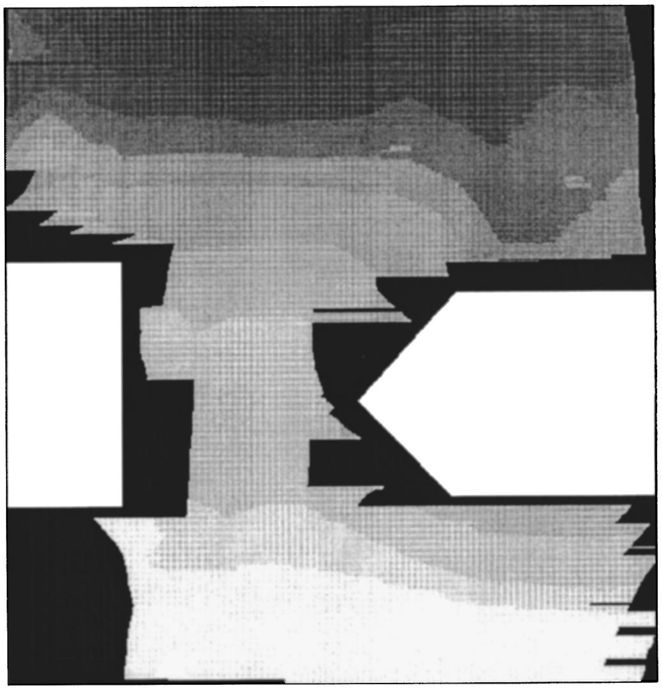


FIG. 4.

$$\Delta T = \frac{\lambda k(x,y)}{\alpha l_0},$$

where α is the coefficient of linear expansion of the glass.

The calculated temperature field in the plate is shown in Figs. 5 and 6.

DISCUSSION OF RESULTS

The type of filtering of the image and the number of times it should be applied is the most controversial aspect of this algorithm. Since the optical system contains scattering centers, speckle noise occurs which leads to the formation of many spurious maxima and minima on the interference pattern. If the optical characteristics of the system are known, the characteristic speckle size is given by[7]

$$b_s = 1.22\,\frac{(1+M)\lambda f}{D},$$

where f is the focal length of the objective, M is the magnification coefficient, and D is the aperture size.

The spatial frequency of the speckle structure is $\nu_s = 1/b_s$ but since the speckle noise is multiplicative, a filter cannot be constructed directly in this frequency range. However, a moving-average filter and a spin filter, used many times, are effective. For interference patterns for which the period of the fringes is small and the noise is negligible, it is best to use a spin filter: the frequent intensity maxima and minima (corresponding to the fringes) do not disappear with the spurious extrema. The other parameters used to process the interference patterns were selected as a function of the fringe period. The error in calculations of the characteristics of a physical process using interference patterns is mainly attributable to the error in isolating the skeletal lines. Despite the stage-by-stage filtering, an isolated skeletal line may pass through points whose coordinates do not satisfy condition (2). Since the presence of these points can be attributed to speckle noise, we can postulate that these are randomly scattered around the true skeletal line. The error can be estimated if the shape of the fringes on the interference pattern is already known. In these experiments, straight parallel fringes were initially observed on the interference pattern. The average slope of the lines was calculated for an image with selected skeletal lines. The error was estimated using the formula[8]

$$\Delta = t_{0.95}\sigma,$$

where $t_{0.95} = 2.5$ is Student's coefficient for 0.95 confidence and σ is the rms deviation calculated using the formula

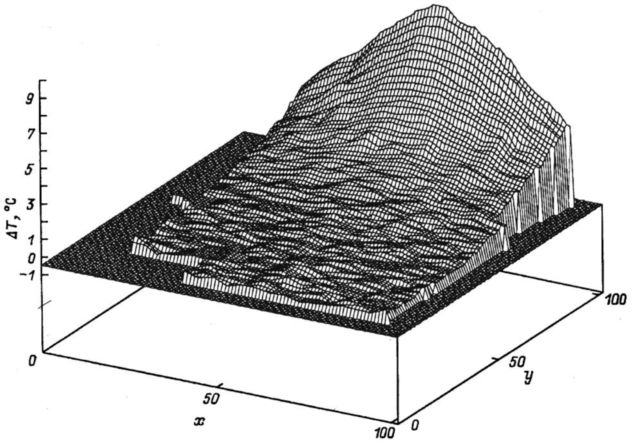


FIG. 5.

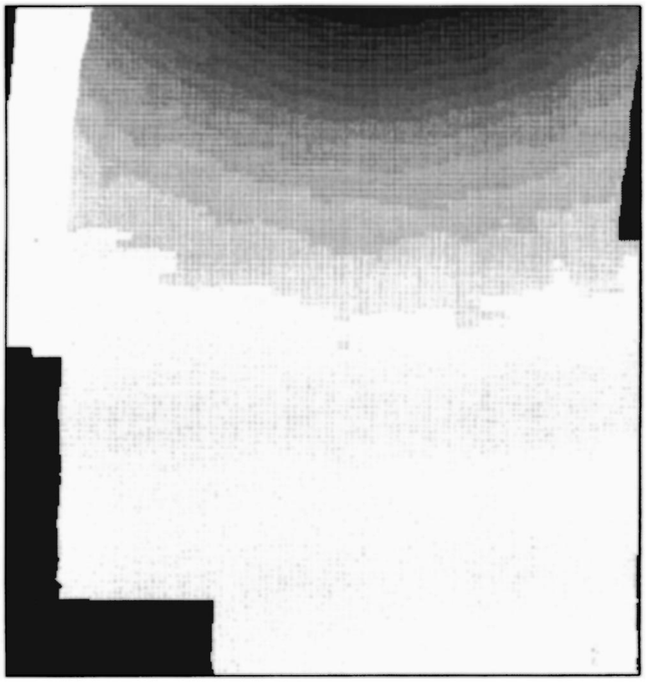


FIG. 6.

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x_i})^2+(y_i-\overline{y_i})^2}{n}},$$

where n is the number of points, x_i, y_i are the coordinates of points assigned to isolated skeletal lines, and $\overline{x_i}, \overline{y_i}$ are the coordinates of the corresponding lines.

Estimates indicate that the error in measuring the change in the refractive index as a result of isolating the skeletal lines is 8%.

Since the formulated problems were not solved analytically, the results were assessed qualitatively.

The temperature distribution in the liquid at the very beginning of the process can reveal characteristic effects for this process: the initial section of the jet, its reflection from the walls, and the thermal boundary layer near the wedge. In Fig. 3 the temperature field is highly asymmetric. This can be caused by tilting of the nozzle and by natural convection of the heated liquid.

If we now consider the propagation of heat in the glass, the temperature maximum is situated near the edge of the plate with the heating element, around the geometric center of the heat source. With increasing distance from this source, the temperature decreases and reaches its initial value. The heating zone of the glass and also the maximum increase as the heating time increases.

Errors in calculating the change in the refractive index are associated with the assumption that the object is homogeneous along the line of observation. These errors cannot be estimated without making additional measurements.

The precision of this method of holographic interferometry in determining the temperature can be calculated as follows. For a fringe shift of 1/100 and an object of thickness 1 cm formula (3) gives $\Delta n = 0.63\times10^{-6}$, which corresponds to a temperature change of 6.6×10^{-3} °C. For the estimates we use data for water in the temperature range 20–25 °C.[9]

CONCLUSIONS

Holographic interferometry has significant advantages over other methods of investigations since it is a noncontact method and introduces no perturbations into the object. It can also be used to calculate the characteristics of a process at almost any observable point with a high degree of accuracy. However, the process of obtaining quantitative information by analyzing interference fringes is laborious and at this stage some of the accuracy characteristic of the method is lost. This study makes it possible to use holographic interferometry to obtain the quantitative characteristics of some processes.

Advantages of this algorithm are that few parameters are required and some of them can be predetermined for a specific optical system and do not require subsequent correction.

To conclude, the authors would like to thank L. M. Malkhasyan and V. M. Levushkin for assistance with these studies.

[1] Yu. I. Ostrovskiĭ, M. M. Butusov, and G. V. Ostrovskaya, *Holographic Interferometry* [in Russian], Nauka, Moscow (1977), 336 pp.

[2] L. A. Borynyak, A. V. Loginov, P. M. Mednis, and V. N. Sarnadskiĭ, Avtometriya No. 4, 59 (1993).

[3] *Scientific-Technical Report ''Methods of Obtaining and Primary and Mathematical Processing of Interference Patterns of Supersonic Gasdynamic Objects''* [in Russian], A. F. Ioffe Physicotechnical Institute, St. Petersburg (1993), 99 pp.

[4] A. M. Ushakov, *Digital Signal Analysis and Its Applications* [in Russian], Nauka, Moscow (1981), pp. 99–125.

[5] T. Pavlidis, *Algorithms for Graphics and Image Processing* [Computer Science Press, Rockville, MD, 1982; Radio i Svyaz', Moscow, 1986, 400 pp.].

[6] Q. Yu, X. Li, and K. Andresen, Appl. Opt. **33**, 3705 (1994).

[7] N. A. Fomin, *Speckle Interferometry of Gas Flows* [in Russian], Nauka i Tekhnika, Minsk (1989), 168 pp.

[8] D. J. Hudson, CERN Report (unpublished) [Geneva (1964); Mir, Moscow (1970), 29 pp.].

[9] *Handbook of Chemistry and Physics*, 37th ed. (Chemical Rubber Publishing Co., Boca Raton, Florida, 1955–1956), 3156 pp.

Translated by R. M. Durham

BRIEF COMMUNICATIONS

Generation of electromagnetic waves by relativistic electrons in a cavity with crossed radial electrostatic and axial magnetic fields under plasma resonance conditions

Yu. V. Kirichenko

National Scientific Center, Kharkov Physicotechnical Institute, 310108 Kharkov, Ukraine

(Submitted September 29, 1998)
Zh. Tekh. Fiz. **69**, 112–114 (June 1999)

A theoretical analysis is made of the conditions for generation of electromagnetic waves by a thin cylindrical layer of relativistic electrons rotating in crossed axial magnetic and radial electrostatic fields in a cylindrical cavity. A dispersion equation is obtained to describe the interaction between waves and electrons under plasma resonance conditions. The dependence of the growth rates on the relativistic factor and the magnetic field are studied.

[S1063-7842(99)02106-6]

Interest in studying the dynamics of charged particle motion in a cylindrically symmetric electrostatic field has been stimulated by various practical applications of systems such as Geiger–Müller counters, electrical filters for pollutant gases, ion–plasma pumps, high-pressure gas-discharge meters, millimeter wave generators, and so on. The possibility of energy exchange between nonrelativistic electrons rotating in crossed fields and an electromagnetic wave was analyzed theoretically in Refs. 1 and 2 where the authors showed that one mechanism responsible for the generation of electromagnetic waves is plasma resonance. An increase in the electron velocity to relativistic levels could substantially increase the oscillation frequency as far as the submillimeter range. A longitudinal magnetic field increases the force keeping the electrons in an orbit which can increase not only the electron velocity but also their density. This last fact can increase the oscillator power and the wave growth rates.

We shall consider a cylindrical metal cavity unbounded along the z axis (using the r, φ, z cylindrical system of coordinates) along which a magnetic field B_0 is directed. A cylindrical layer of electrons rotates about the cylinder axis on which a metallic charged filament is located. The internal radius of the cavity is b and the filament radius is a $(a\ll b)$. The relativistic electrons are confined on equilibrium circular orbits by the crossed magnetic field B_0 and the radial electrostatic field of the filament $E_0(r)$. We shall neglect the constant self-induced magnetic and electric fields of the electron layer. We shall assume that the system is homogeneous along the z axis. We shall determine the dependence of all the variables on φ and the time t using the factor $\exp[i(m\varphi-\omega t)]$, where $m\neq 0$ is an integer and ω is the complex frequency. The analysis will be performed in the hydrodynamic approximation. The equilibrium unperturbed electron velocity is given by

$$v_0(r)=\frac{1}{2}\omega_c r+\frac{1}{2}q\left(\omega_c^2 r^2+\frac{4E_0(r)er}{m_e\gamma}\right)^{1/2}, \tag{1}$$

where $\omega_c=eB_0/m_e\gamma c$, $\gamma=1/(1-v_0^2(r)/c^2)^{1/2}$, c is the velocity of light, $-e<0$, and m_e are the electron charge and mass, and $q=\pm 1$.

The unperturbed density of the cylindrically symmetric electron layer $n_0(r)$ is nonzero between the surfaces $r=r_-$ and $r=r_+$, i.e., $n_0(r)=0$ for $r\leqslant r_-$ and $r\geqslant r_+$ $(a<r_-<r_+\ll b)$. The approximation of a thin electron layer implies that

$$\frac{\delta r}{r_-}\ll 1, \tag{2}$$

where $\delta r=r_+-r_-$.

Using the linearized Lorentz equation and equation of continuity, together with the Maxwell equations we obtain a differential equation in the region $r_-<r<r_+$ for E_φ which is the azimuthal component of the wave field **E**

$$\begin{aligned}\frac{d}{dr}&\left\{\frac{c^2}{r\omega_{r1}}\left(1-\frac{\Omega^2}{\gamma^3 w_r}\right)\frac{d}{dr}(rE_\varphi)+\frac{\omega_d\Omega^2(mc^2-v_0\omega r}{\gamma^3 r w_r w_{r1}\omega_m}E_\varphi\right\}\\ &=-\frac{\Omega^2\omega_g(mc^2-v_0\omega r)}{\gamma\omega_m r^2 w_r w_{r1}}\frac{d}{dr}(rE_\varphi)\\ &-\left\{1-\frac{\Omega^2}{\gamma^3 w_r}+\frac{\Omega^2\omega_d}{\gamma^3 w_r\omega_m^2}\left(v_0'-\frac{v_0}{r}+\frac{\Omega^2\omega_g\omega_m^2}{\gamma w_r w_{r1}}\right)\right\}E_\varphi,\end{aligned} \tag{3}$$

where $w_r=\omega_m^2+w_{r0}$, $w_{r1}=\omega^2-m^2c^2/r^2-\omega_m^2\Omega^2/\gamma w_r$, $w_{r0}=(\omega_g+v_0/r-v_0')\omega_d$, $\omega_d=2v_0/r+(\gamma^2-2)\omega_c+ev_0E_0\gamma/m_ec^2$, $\omega_g=\omega_c-2v_0/r+ev_0E_0/m_e\gamma c^2$, $\Omega^2(r)=4\pi n_0(r)e^2/m_e$, and $\omega_m=\omega-mv_0/r$.

A plasma resonance occurs when the frequency of the electromagnetic field is close to the frequency of the natural longitudinal local oscillations of the electrons in the laboratory frame, i.e., subject to the condition

$$\gamma^3 w(\omega) \simeq \Omega^2(r). \tag{4}$$

Condition (4) yields an approximate expression for the real part of the resonance frequency

$$\mathrm{Re}\,(\omega) \simeq \omega_p\,,$$

where

$$\omega_p = \frac{m v_0}{r} \pm \Delta\omega_p\,, \quad \Delta\omega = \left(\frac{\Omega^2(r)}{\gamma^3} - w_{r0} \right)^{1/2}. \tag{5}$$

Integrating Eq. (3) over r and neglecting terms higher than the first order of smallness with respect to the parameter $\delta r / r_-$, we obtain boundary conditions for E_φ and $d(rE_\varphi)/dr$ (a similar method was used in Refs. 1–6). Matching E_φ and $d(rE_\varphi)/dr$ at the layer boundaries with the corresponding values in vacuum and taking into account the resonance condition (4), we obtain a very complex dispersion equation. In order to analyze this equation at least qualitatively, we express the Bessel J_m and Neumann N_m functions of the arguments $x_1 = \omega a/c$ and $x = \omega r_-/c$ contained therein as a series in powers of the arguments, confining ourselves to the first terms of the expansion. In the functions $J_m(x_2)$ and $N_m(x_2)$ $(x_2 = \omega b/c)$ we confine ourselves to the first terms of the asymptotic expansion. In accordance with the condition (2), we shall seek a solution of the dispersion equation in the form $\omega = \omega^{(0)} + \omega^{(1)}$, $|\omega^{(1)}| \ll \omega^{(0)}$, where $\omega^{(0)}$ are the natural frequencies of the cavity in the absence of electrons. The dispersion equation describing the plasma resonance in a homogeneous electron layer finally has the form

$$(\delta\omega)^2 + (i\nu - \Delta_{\nu p}) + \Delta_p^2 = 0, \tag{6}$$

$$\Delta_p^2 = -\frac{c^2 \pi x_0^{2\bar{m}-1} \omega_p [\bar{m} \mp |v_0| \gamma^2 r_- \Delta\omega_p / c^2]^2 \varepsilon^2(r_-)\, \delta r}{r_-^2 2^{2\bar{m}+2} \bar{m}!^2 \eta^{4\bar{m}}\, \mathrm{Re}(\omega_m)\, b \gamma^5} \times (\omega_d d_0 - \eta^{2\bar{m}} - 1)^2, \tag{7}$$

$$d_0 = \frac{\mathrm{sign}\,(v_0)\, \bar{m} (\eta^{2\bar{m}} - 1)}{[\bar{m} \mp |v_0| \gamma^2 r_- \cdot \Delta\omega_p / c^2]\, \mathrm{Re}\,(\omega_m)}, \tag{8}$$

$$\Delta_{\nu p} \simeq \omega^{(0)} - \omega_p + \Delta_\nu\,, \tag{9}$$

where $\delta\omega = \omega - \omega_p$, $x_0 = k_0 r_-/c$, $k_0 = \omega^{(0)}/c$, $\varepsilon(r_-) = \Omega(r_-) c / r_-$, $\eta = r_-/a$, $\bar{m} = |m|$; the term $i\nu$ in Eq. (6) takes into account phenomenologically the losses caused by absorption in the walls and radiation from the cavity, and the term Δ_ν in Eq. (9) allows for the frequency shift caused by these losses; in formulas (7) and (8) the $\mp$ signs correspond to the $\pm$ signs in formula (5).

Since $\mathrm{Im}\,(\omega) = \mathrm{Im}\,(\delta\omega)$, the condition $\mathrm{Im}\,(\delta\omega) > 0$ corresponds to oscillations which increase in time. These solutions of Eq. (6) are obtained for $\Delta_p^2 > 0$ or, as follows from Eq. (7), when

$$v_0 > v_{\mathrm{ph}}\,, \quad v_{\mathrm{ph}} = \frac{\mathrm{Re}(\omega) r_-}{m}, \tag{10}$$

v_{ph} is the phase velocity of the wave near the electron layer.

It follows from Eqs. (6) and (7) that when condition (10) is satisfied, the wave may become damped. However, unlike Cherenkov resonance,[1,2] the decrement of this wave is not equal to the growth rate. In the most general case, it follows from Eq. (6) that for $\Delta_{\nu p}^2 < 4\Delta_p^2$, when the difference $|\omega^{(0)} - \omega_p|$ or the frequency shift Δ_ν are small, unstable solutions may be obtained in the absence of losses $(\nu = 0)$. However, when $\Delta_{\nu p}^2 > 4\Delta_p^2$ the losses cause instability. If $|i\nu - \Delta_{\nu p}| \gg |\Delta_p|$, the growth rate is given by $\mathrm{Im}(\omega) = \nu\Delta_p^2/(\nu^2 + \Delta_{\nu p}^2)$ and has a maximum in terms of ν.

On considering an inhomogeneous electron layer, we shall assume that the function $\Omega^2(r)$ is Gaussian with a maximum at $r = r_m$. The plasma resonance condition (4) will be satisfied for two values of the radial coordinate $r = r_1$ and $r = r_2 (r_- < r_1 < r_m < r_2 < r_+)$. Finally, the growth rate is given by

$$\mathrm{Im}(\omega) = -\nu - \mathrm{sign}(\mathrm{Re}(\omega_m)) \times \frac{\pi^2 h x_0^{2\bar{m}-1} \omega_p [\bar{m} \mp |v_0| \gamma^2 r_- \cdot \Delta\omega_p / c^2]^2}{2^{2\bar{m}+1} \bar{m}!^2 b \gamma^2 \eta^{4\bar{m}}} \times (\omega_d d_0 - \eta^{2\bar{m}} - 1)^2, \tag{11}$$

where

$$h = \frac{\Omega^2(r_1)}{\partial\Omega^2/\partial r|_{r_1} - \partial(\gamma^3 w_r)/\partial r|_{r_1,\omega_p}} - \frac{\Omega^2(r_2)}{\partial\Omega^2/\partial r|_{r_2} - \partial(\gamma^3 w_r)/\partial r|_{r_2,\omega_p}}. \tag{12}$$

In our particular case when $|\partial\Omega^2/\partial r|_{r_{1,2}} \gg |\partial(\gamma^3 w_r)/\partial r|_{r_{1,2}\omega_p}$, it can be seen from Eq. (11) and (12) that oscillation accompanying plasma resonance in an inhomogeneous electron layer occurs when condition (10) is satisfied. Unlike the case of a homogeneous layer, the instability is a threshold process and occurs when the second term in Eq. (11) is greater than the first. For a homogeneous layer expression (5) gives the formula for the frequency of the generated wave

$$\mathrm{Re}(\omega) \simeq \frac{c}{r_-} \left\{ \frac{\bar{m}(\gamma^2 - 1)^{1/2}}{\gamma} - \frac{1}{\gamma^2} (\varepsilon^2 \gamma + \gamma^2 (\gamma^2 - 1) + [(\gamma^2 - 1)^{1/2} - \mathrm{sign}(v_0) \varepsilon_1]^2)^{1/2} \right\}, \tag{13}$$

where $\varepsilon_1 = \omega_c r_- \gamma / c$.

For an inhomogeneous electron layer the frequency at which the resonance condition (4) is satisfied depends on the coordinate r which varies in the range $r_- \leq r \leq r_m$ (or $r_m \leq r \leq r_+$). Hence the generated frequencies will lie in the range

$$\frac{c}{r_-}\left\{\frac{\overline{m}(\gamma^2-1)^{1/2}}{\gamma}-\frac{1}{\gamma^2}(\varepsilon_m^2\gamma+\gamma^2(\gamma^2-1)+[(\gamma^2-1)^{1/2}-\mathrm{sign}(v_0)\varepsilon_1]^2)^{1/2}\right\}\leqslant\mathrm{Re}(\omega)\leqslant\frac{c}{r_-}\times\left\{\frac{\overline{m}(\gamma^2-1)^{1/2}}{\gamma}-\frac{1}{\gamma^2}(\gamma^2(\gamma^2-1)+[(\gamma^2-1)^{1/2}-\mathrm{sign}\,(v_0)\varepsilon_1]^2)^{1/2},\right\}, \qquad (14)$$

where $\varepsilon_m=\max_r\varepsilon(r)$.

It can be seen from conditions (13) and (14), where condition (10) is taken into account, that generation can only occur when $\overline{m}>1$. For large γ the range (14) becomes narrower and the generated frequency will be

$$\mathrm{Re}\,(\omega)\simeq\frac{c(\overline{m}-1)}{r_-}. \qquad (15)$$

The frequency (13) also tends toward (15) for large γ. The growth rate for an inhomogeneous layer, which is proportional to (h/b), is smaller than that for a homogeneous layer, which is proportional to $(\delta r/b)^{1/2}$. In the first case, however, the generation has the advantage that the resonance conditions are not violated when the radius of the electron orbit varies.

We also analyzed the case of an axisymmetric wave when $m=0$. We found that for both homogeneous and inhomogeneous electron layers $\mathrm{Im}\,(\omega)<0$ which corresponds to damped oscillations.

The author would like to thank V. V. Dolgopolov for useful discussions of the results.

[1] I. Alexeff and F. Dyer, Phys. Rev. Lett. **45**, 351 (1980).
[2] V. V. Dolgopolov, M. V. Dolgopolov, Yu. V. Kirichenko *et al.*, in *Proceedings of the Seventh International Crimean Conference on ''Microwave Engineering and Telecommunication Technologies''* [in Russian], 1997 Vol. 2, pp. 487–488.
[3] V. V. Dolgopolov, M. V. Dolgopolov, and Yu. V. Kirichenko, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **40**(12), 16 (1997).
[4] K. N. Stepanov, Zh. Tekh. Fiz **35**, 1002 (1965) [Sov. Phys. Tech. Phys. **10**, 773 (1965)].
[5] K. N. Stepanov, Zh. Tekh. Fiz **35**, 1349 (1965) [Sov. Phys. Tech. Phys. **10**, 1048 (1965)].
[6] V. V. Dolgopolov and A. Ya. Omel'chenko, Zh. Éksp. Teor. Fiz. **18**, 1089 (1950).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Vibrations of twinning-boundary atoms

O. M. Ostrikov

Mozyr State Pedagogical Institute, 247760 Mozyr, Belarus, Russia

(Submitted March 30, 1998)
Zh. Tekh. Fiz. **69**, 115–118 (June 1999)

A classical method is used to obtain dispersion equations for the vibrations of twinning-boundary atoms. © *1999 American Institute of Physics.* [S1063-7842(99)02206-0]

1. Twinning is one of the main types of plastic deformation in crystals. Hence, its study from the scientific viewpoint is of major importance because for many materials (such as Bi, Zn, Sb, TiAl, and silicon iron) plastic deformation in specific crystallographic directions is accomplished solely by twinning. Despite a vast amount of experimental data,[1–3] the problem of constructing a logical theory of crystal twinning has yet to be resolved.

The dynamic characteristics of twinning boundaries are decisive factors in studying the physical relationships determining the twinning-induced plastic deformation of solids. In the present paper we study the dynamics of the vibrational motion of twinning-boundary atoms from the point of view of the microscopic theory of twins.[2]

2. At present, an expression describing the vibrations of a chain of atoms of like mass M has now become a classical result[4] in solid-state physics and in this form it is used to simulate a one-dimensional crystal lattice

$$\omega=2\sqrt{\frac{\beta}{M}}\left|\sin\left(\frac{aq}{2}\right)\right|, \tag{1}$$

where ω is the frequency of the atomic vibrations, q is the wave vector, a is the distance between the atoms, and β is a constant.

Using the method applied to obtain formula (1), we can calculate the dependence of ω on q for an atom situated at the twinning boundary A_1A_2 (Fig. 1).

3. We shall analyze a crystal consisting of atoms of like mass M. We shall assume that each atom only interacts with neighboring atoms. In the harmonic approximation the potential interaction energy of these atoms is related to the distance x between them by

$$U(x)\sim x^2. \tag{2}$$

The interaction force f between the atoms is therefore directly proportional to x.

We relate the origin of the Cartesian coordinate system to a twinning-boundary atom (Fig. 2). The position of this atom on the X axis will be denoted by x_n. The positions of atoms *1* and *2* on this axis will be denoted by $x_{n-1}\cos\alpha$ and $x_{n+1}\cos\alpha$, where α is the twinning angle. The force acting on the twinning-boundary atom will then be given by

$$f_n=f_{n+1}+f_{n-1}=\beta(x_{n+1}\cos\alpha+x_{n-1}\cos\alpha-2x_n), \tag{3}$$

where f_{n+1} is the force acting on the twinning-boundary atom from atom *1* (Fig. 2) and f_{n-1} is the force from atom *2*.

In this case, the equation of motion for the twinning-boundary atoms will have the form

$$M\ddot{x}_n=\beta(x_{n+1}\cos\alpha+x_{n-1}\cos\alpha-2x_n). \tag{4}$$

We shall seek a solution of Eq. (4) in the form

$$x_n=x_0\mathbf{e}^{i(\omega t-naq)}, \tag{5}$$

where x_0 is the amplitude of the vibrations and n is the number of the atom.

The position of neighboring atoms numbered $n+1$ and $n-1$ on the x axis at time t can obviously be given by

$$x_{n+1}=x_0\mathbf{e}^{i(\omega t-(n+1)aq)}, \tag{6}$$

$$x_{n-1}=x_0\mathbf{e}^{i(\omega t-(n-1)aq)}. \tag{7}$$

Substituting Eqs. (5)–(7) into Eq. (4) gives a solution which we write as the dependence of ω on q

$$\omega^2=\frac{2\beta}{M}(1-\cos(aq)\cos\alpha) \tag{8}$$

or

$$\omega=\pm2\sqrt{\frac{\beta}{M}\left(\sin^2\left(\frac{aq}{2}\right)+\cos(aq)\sin^2\frac{\alpha}{2}\right)}. \tag{9}$$

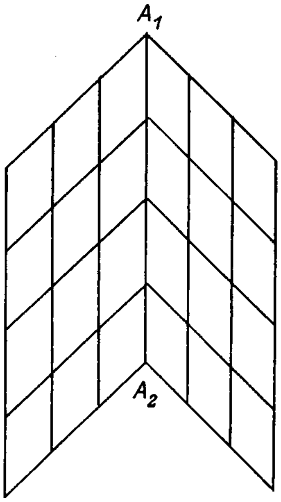


FIG. 1. Schematic of A_1A_2 twinning boundary.

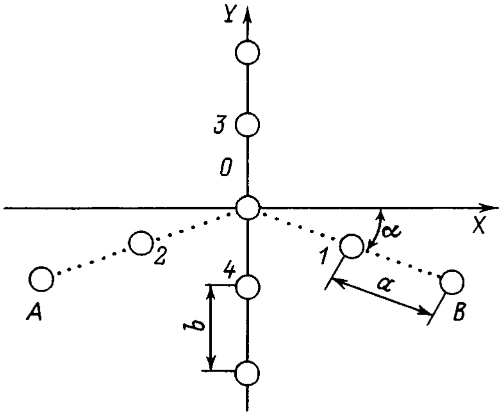


FIG. 2. Twinning-boundary atom surrounded by neighboring atoms.

A comparison of Eqs. (1) and (9) indicates that these formulas differ in respect of the term in the radicand: $\cos(aq)\sin^2(\alpha/2)$. In the limiting case when $\alpha=0$, expression (9) gives formula (1).

Figure 3 gives the dependences

$$\frac{\omega}{2}\sqrt{\frac{M}{\beta}}=\frac{\omega}{2}\sqrt{\frac{M}{\beta}}(aq)$$

for the vibration of a twinning-boundary atom along the X axis for various twinning angles α between zero and 40°. Table I gives the values of $\sin^2(\alpha/2)$ in Eq. (9) for given values of α.

TABLE I.

α	$\sin^2\frac{\alpha}{2}$	$\sin\alpha$
0	0.000	0.0
10	0.008	0.2
20	0.030	0.3
30	0.070	0.5
40	0.100	0.6

It can be seen from Fig. 3 that an increase in the twinning angle α leads to ‘‘truncation’’ of the low frequencies and reduces the frequency range of the vibrations of the twinning-boundary atoms.

4. Along the x axis, the nearest atoms *1–4* will exert the force

$$\begin{aligned}F_n&=F_{n+1}+F_{n-1}+2F\\&=\beta(y_{n+1}+y_{n-1}(1+2\sin\alpha)-2y_n),\end{aligned} \tag{10}$$

on the twinning-boundary atom, where F_{n+1} is the force acting on the twinning-boundary atom from atom *3*, F_{n-1} is the force from atom *4*, and F is the force from atoms *1* or *2*.

In this case, the equation of motion has the form

$$M\ddot{y}_n=\beta(y_{n+1}+y_{n-1}(1+2\sin\alpha)-2y_n), \tag{11}$$

whose solution will be sought in the form

$$y_n=y_0\mathbf{e}^{i(\Omega t-nbq)}, \tag{12}$$

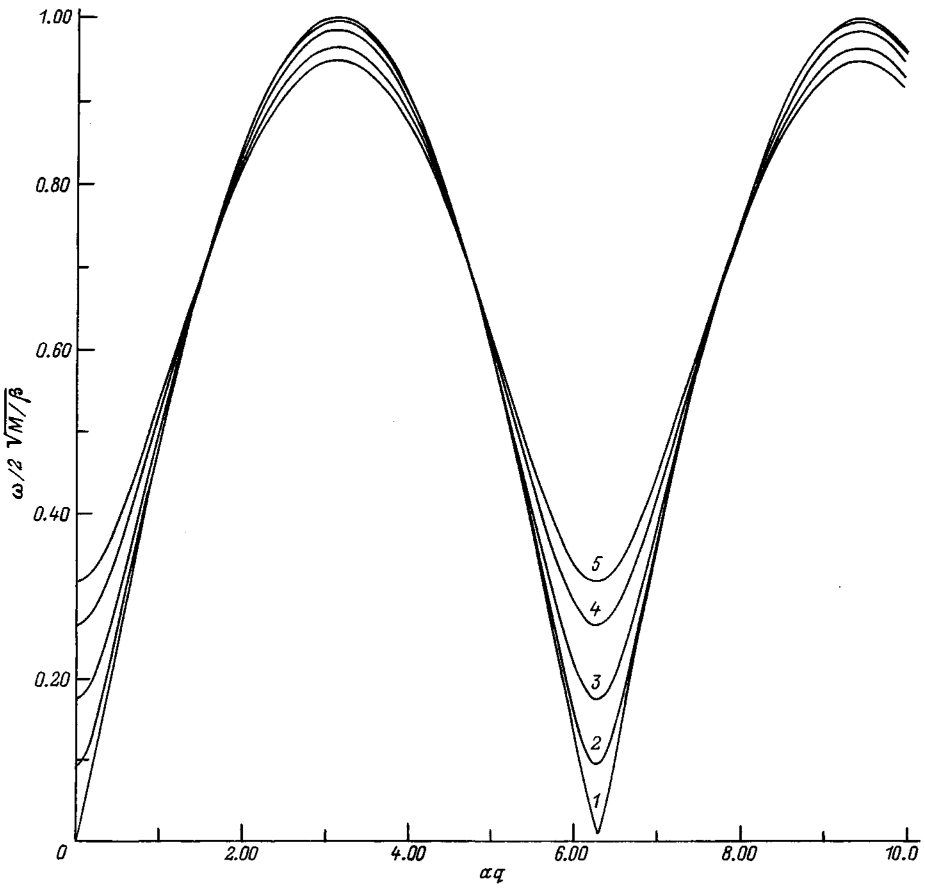


FIG. 3. Dependences of $\frac{\omega}{2}\sqrt{\frac{M}{\beta}}$ on aq: $\alpha=0$ (*1*), 10 (*2*), 20 (*3*), 30 (*4*), and 40° (*5*).

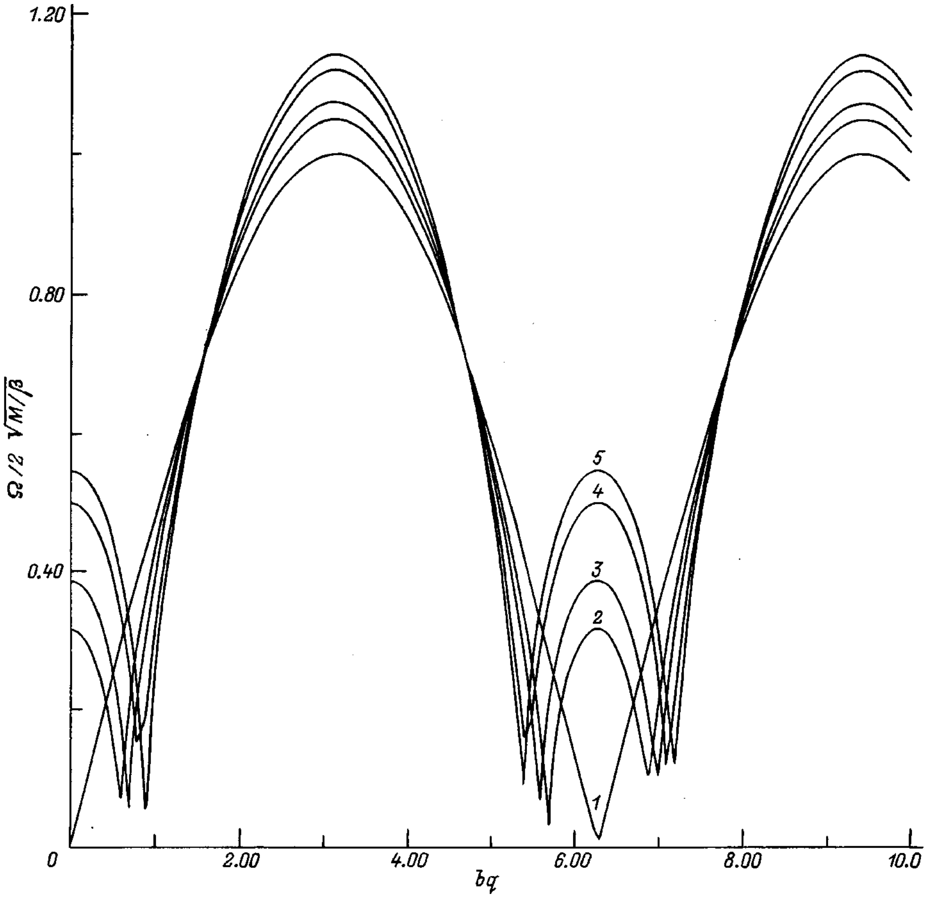


FIG. 4. Dependences of $\frac{\Omega}{2}\sqrt{\frac{M}{\beta}}$ on bq: *1–5* — as in Fig. 3.

where Ω is the frequency of the atomic vibrations along the y axis and b is the distance between the atoms along this axis.

In this case, the solution of Eq. (11) may be written

$$\Omega^2=\frac{2\beta}{M}\left(2\sin^2\left(\frac{bq}{2}\right)-\cos(bq)\sin\alpha\right) \tag{13}$$

or

$$\Omega=\pm 2\sqrt{\frac{\beta}{M}\left(\sin^2\left(\frac{bq}{2}\right)-\frac{1}{2}\cos(bq)\sin\alpha\right)}. \tag{14}$$

This relation differs from Eq. (1) in respect of the term in the radicand: $-(1/2)\cos(bq)\sin\alpha$. As in the previous case [see formula (9)], Eq. (14) gives formula (1) for $\alpha=0$.

Figure 4 shows a graph of

$$\frac{\Omega}{2}\sqrt{\frac{M}{\beta}}=\frac{\Omega}{2}\sqrt{\frac{M}{\beta}}(bq).$$

In this case, an increase in the twinning angle leads to the appearance of an additional maximum at the boundaries of the Brillouin zone and increases the range of vibration frequencies of the twinning-boundary atom.

Note that when plotting the graph in Fig. 4, we imposed the constraint that the radicand in Eq. (14) should be greater than zero.

5. To sum up, a classical method has been used to calculate the dispersion dependence of the vibrations of twinning-boundary atoms. An increase in the twinning angle leads to a decrease in the frequency range of the vibrations of a twinning-boundary atom perpendicular to the twinning boundary and increases this range for vibrations parallel to the twinning boundary.

[1] M. V. Klassen-Neklyudova, *Mechanical Twinning of Crystals* [Consultants Bureau, New York, 1964; USSR Academy of Sciences Press, Moscow, 1960].

[2] A. M. Kosevich and V. S. Boĭko, Usp. Fiz. Nauk **104**, 201 (1971) [Sov. Phys. Usp. **14**, 286 (1971)].

[3] V. S. Savenko, O. M. Ostrikov, A. I. Pinchuk, and S. D. Shavreĭ, *Abstracts of Papers presented at the Fourth International Conference on ''Action of Electromagnetic Fields on the Plasticity and Strength of Materials''* [in Russian], 1996, p. 21.

[4] C. Kittel, *Introduction to Solid State Physics*, 5th ed. [Wiley, New York (1976); Nauka, Moscow (1978), 792 pp.].

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Nonlinear surface magnetostatic waves in a ferrite semiconductor structure

A. S. Kindyak

Minsk Research Institute of Radio Materials, 220024 Minsk, Belarus

(Submitted February 9, 1998)
Zh. Tekh. Fiz. **69**, 119–121 (June 1999)

A theoretical analysis is made of the propagation of a nonlinear surface magnetostatic wave in a planar ferrite semiconductor structure as a function of the carrier concentration in the semiconductor layer. It is shown that for certain concentrations the surface magnetostatic wave is unstable with respect to longitudinal perturbations and may propagate perpendicular to the magnetic field in the form of solitons. *© 1999 American Institute of Physics.* [S1063-7842(99)02306-5]

Recently considerable interest has been shown in studying envelope solitons of magnetostatic spin waves in thin ferromagnetic films.[1–4] So far, the conditions for the existence and formation of solitons have been studied theoretically[1,2] and the propagation of magnetostatic waves pulses in the form of solitons has also been observed experimentally.[3] However, nonlinear surface magnetostatic waves (SMSWs) in ferrite semiconductor structures have been studied considerably less.

In the present paper we show that it is theoretically possible for SMSW solitons to exist in a planar ferrite semiconductor structure. The SMSWs propagate along the Y axis perpendicular to an external magnetic field H_0 directed along the Z axis parallel to the surface of the structure.

The initial equations for studying nonlinear surface spin waves in a ferrite semiconductor structure are the Maxwell equations, the Landau–Lifshitz equation in the magnetostatic approximation for a ferrite layer, and also the Maxwell equations and the Lorentz equation for a semiconductor layer.

During the propagation of SMSWs, the role of nonlinearity is manifested as a dependence of the frequency and phase velocity on the wave amplitude. Assuming that the amplitude φ is a slowly varying function, the nonlinear dispersion equation for the SMSWs may be formally written in the form:[1] $G(\omega,k,|\varphi|^2)=0$, where ω is the carrier frequency and $k\equiv k(\omega)$ is the SMSW wave vector. Since the nonlinearity is assumed to be weak, the deviation of the frequency ω from ω_0 to $\omega_0+\Omega$, $\Omega\ll\omega_0$ and the changes in k compared with the wave vector k_0 of the linear equation will be small. After expanding $k(\omega)$ about $k_0(\omega_0)$ as a series in terms of $(\omega-\omega_0)$, we obtain the nonlinear Schrödinger equation[1,4]

$$i\frac{dA}{dy}-\frac{\beta_2}{2}\frac{d^2A}{dt^2}+\gamma|A|^2A=0, \tag{1}$$

where $\beta_2=d^2k/d\omega^2|_{\omega_0}$ is the dispersion of the linear group velocity calculated at the point $k_0(\omega_0)$ and $\gamma=dk/d|A|^2|_{A=0,\omega_0}$ is the nonlinear coefficient.

A necessary condition for the formation of an envelope soliton is that the Lighthill criterion[5] $\beta_2\gamma<0$ is satisfied.

The dispersion equation for a linear SMSW propagating in a ferrite semiconductor structure has the form[6]

$$e^{-2kd}=\frac{(\mu^-+k)(\delta\mu^++k)\tanh(k_st)+(\mu^++k)\sqrt{\delta}}{(\mu^+-k)(\delta\mu^--k)\tanh(k_st)+(\mu^--k)\sqrt{\delta}}. \tag{2}$$

Here $\mu^+=k(\mu\pm\mu_a)$, $k_s^2=\delta k^2$, $k=k_1+ik_2$, d and t are the thickness of the ferrite and semiconductor layer, respectively, μ and μ_a are the diagonal and off-diagonal elements of the magnetic permeability tensor of the ferrite allowing for losses. The value of δ determines the characteristics of interaction between the SMSW and the semiconductor layer and in this case, may be written in the form

$$\delta=1-\frac{\omega\varepsilon_{zz}}{k^2c^2}, \tag{3}$$

where $\varepsilon_{zz}=\varepsilon_{zz}^e+\varepsilon_{zz}^h$, $\varepsilon_{zz}^{e(h)}$ is an element of the permittivity tensor of the semiconductor in the magnetic field H_0

$$\varepsilon_{zz}^e=\varepsilon_s\left(1-\frac{\omega_p^2}{\omega^2+\nu^2}+i\frac{\omega_p^2\nu}{\omega(\omega^2+\nu^2)}\right), \tag{4}$$

where ε_s is the permittivity of the semiconductor caused by the contribution of the lattice, ν is the electron collision frequency, $\omega_p=(4\pi e^2N/m^*\varepsilon_s)^{1/2}$ is the plasma frequency, and N is the electron concentration.

For the case $\delta=1$ Eq. (2) gives the Damon–Eschbach equation. In order to determine the coefficient of nonlinearity γ we shall assume that for small deviations of the magnetization from the equilibrium state

$$M_z\cong M_0\left(1-\frac{|M_x|^2+|M_y|^2}{2M_0^2}\right),$$

where M_0 is the saturation magnetization and M_i are the components of the variable magnetization $(i=x,y,z)$.

Then, in the limit $kd\ll1$, we have $M_z\cong M_0-M_0|\varphi|^2$ and $\omega_m\cong\omega_m(1-|\varphi|^2)$ (Refs. 1 and 4). Substituting these expressions into Eq. (2) we obtain

$$\gamma=-\frac{dG}{d|A|^2}\bigg/\frac{dG}{dk},$$

$$\frac{dG}{d|A|^2}=AD_a-BC_a,$$

$$A=\frac{1}{2}[\delta+\sqrt{\delta}+(\sqrt{\delta}-\delta)\exp(-2k_s t)],$$

$$B=(1-\delta)(1-\exp(-2kd))/4,$$

$$D_a=\omega_m(\omega_n+\omega_m(1-\exp(-2kd))/2),$$

$$C_a=\omega_m(\omega_h-\omega)/2,$$

$$\frac{dG}{dk}=A_kD+AD_k-(B_kC+BC_k),$$

$$D=\omega_0^2-\frac{\omega_m^2}{4}(1-e^{-2kd}),$$

$$C=\omega_0^2+\frac{\omega_m}{2}(\omega+\omega_h)(1-e^{-2kd}), \tag{5}$$

where A_k, B_k, C_k, D_k, and δ_k are derivatives with respect to k, and $\omega_0^2=\omega_h^2+\omega_h\omega_m$.

Using

$$\frac{d\omega}{d|A|^2}=v_g\frac{dk}{d|A|^2}$$

and substituting $\delta=1$, $\delta_k=0$, $k=0$, and $\omega=\omega_0$ into Eq. (5), we obtain $d\omega/d|A|^2=\omega_m\omega_h/2\omega_0$, which agrees with the data given in Ref. 1 to within a factor.

We know that solitons can occur as a result of longitudinal and transverse modulational instability. The first is observed as self-modulation and the second as self-focusing. Self-focusing of SMSWs in a ferrite film without losses was studied by Zvezdin and Popkov,[1] who showed that an SMSW propagating perpendicular to the magnetic field is stable with respect to longitudinal perturbations. Boardman *et al.*[2] showed that an SMSW propagating at an angle to the magnetic field is unstable with respect to longitudinal perturbations, which leads to the formation of SMSW envelope solitons.

It can be seen from the data plotted in Fig. 1 for ferrite semiconductor films that a SMSW is unstable with respect to longitudinal perturbations even when it propagates normal to the magnetic field. Calculations were made neglecting the losses for values of the parameters corresponding to those of iron yttrium garnet films and a CdSe semiconductor: $4\pi M_0=1750$ G, $H=960$ Oe, hole mobility $\mu_h=50\ \mathrm{cm^2/V\,s}$, electron mobility $\mu_e=650\ \mathrm{cm^2/V\,s}$, $d=0.002$ cm, and $t=0.0002$ cm.

Figure 1 shows that for an isolated ferrite film ($\delta=1$) the SMSW is stable with respect to longitudinal perturbations. In the presence of a semiconductor layer, a region of existence of envelope solitons appears from low carrier concentrations and this extends over the entire region of existence of the SMSW as the concentration increases.

Allowance for losses in the ferrite and the semiconductor yields more complex frequency dependences of the SMSW group velocity dispersion and the nonlinear coefficient for various carrier concentrations in the semiconductor (Fig. 2). Figure 2 gives the calculated products of the real parts of β_2 and γ. It can be seen that at low concentrations the Lighthill criterion is satisfied over a narrow range in the low-frequency part of the spectrum. An increase in concentration causes the region of existence of SMSW solitons to disappear. A further increase in concentration leads to the reap-

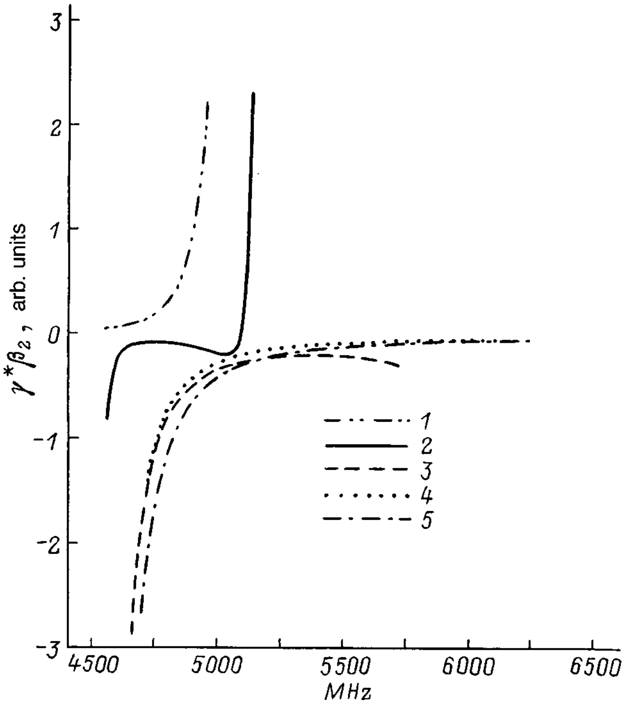


FIG. 1. Product of group velocity dispersion and nonlinear coefficient as a function of frequency for an isolated ferrite film (*1*) and a ferrite–semiconductor structure with electron concentrations of 10^{16} (*2*), 10^{17} (*3*), 10^{18} (*4*), and $10^{19}\ \mathrm{cm^{-3}}$ (*5*).

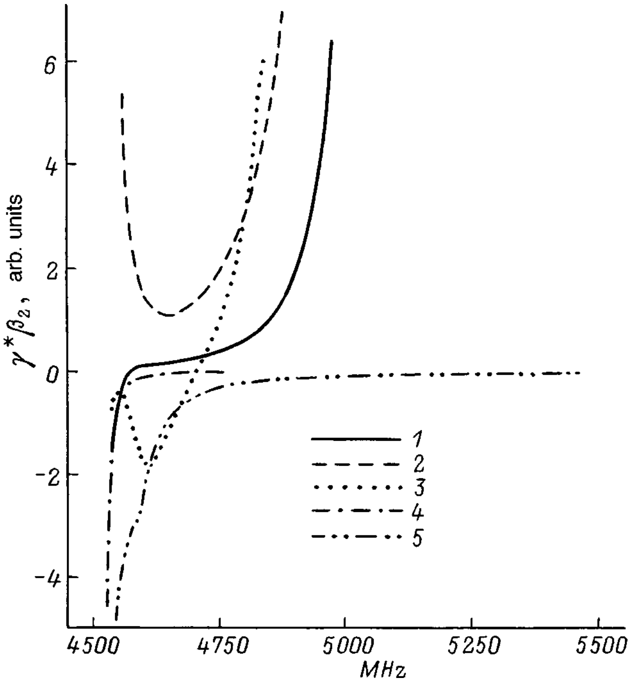


FIG. 2. Product of group velocity dispersion and nonlinear coefficient as a function of frequency for a ferrite–semiconductor structure with electron concentrations of 10^{16} (*1*), 10^{17} (*2*), 10^{18} (*3*), 10^{19} (*4*), and $10^{20}\ \mathrm{cm^{-3}}$ (*5*).

pearance of this region which then expands over almost the entire SMSW frequency range as N increases. Note that Fig. 2 gives values for which $k_2 > -20\,\mathrm{cm}^{-1}$. Thus, these data show that in planar ferrite semiconductor structures an SMSW is unstable relative to longitudinal perturbations.

We shall estimate the threshold SMSW power needed for soliton formation which is defined as $|\varphi|^2 = -\beta_2/\gamma\tau$, where τ is the pulse duration. Using an asymptotic expression for the SMSW power[1,4] $P = \pi L d^2 \omega M^2 |\varphi|^2$ (where L is the structure width), we find that for the parameters used and $N = 10^{18}\,\mathrm{cm}^{-3}$, $\omega = 4.6\,\mathrm{GHz}$, and $\tau = 8 \times 10^{-7}$ s the threshold power for soliton formation is $P = 17\,\mathrm{mW}$.

This work was supported by the Belarus Republican Fund for Fundamental Research, Project No. 96-142.

[1] A. K. Zvezdin and A. D. Popkov, Zh. Éksp. Teor. Fiz. **84**, 606 (1983) [Sov. Phys. JETP **57**, 350 (1983)].

[2] A. D. Boardman, Yu. V. Gulyaev, and S. A. Nikitov, Zh. Éksp. Teor. Fiz. **95**, 2140 (1989) [Sov. Phys. JETP **68**, 1238 (1989)].

[3] B. A. Kalinikos, N. G. Kovshikov, and A. N. Slavin, Zh. Éksp. Teor. Fiz. **94**(2), 159 (1988) [Sov. Phys. JETP **67**, 303 (1988)].

[4] A. D. Boardman, S. A. Nikitov, and N. A. Waby, Phys. Rev. B **48**, 13 602 (1993).

[5] M. J. Lighthill, J. Inst. Appl. Math. **1**, 269 (1965).

[6] A. S. Kindyak, Zh. Tekh. Fiz **64**(11), 99 (1994) [Tech. Phys. **39**, 1143 (1994)].

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Efficiency of microwave energy storage in a cavity

S. N. Artemenko

Research Institute of Nuclear Physics at Tomsk State University, 634050 Tomsk, Russia

(Submitted February 20, 1998)
Zh. Tekh. Fiz. **69**, 122 (June 1999)

A refinement is made to calculations of the microwave energy storage efficiency in a cavity.

The optimum input coupling coefficient β corresponding to the maximum energy storage efficiency η for given input pulse length t_i is usually determined using the well-known formula

$$\beta=(2.512\tau_p)/t_i-1, \tag{1}$$

which is derived from the expression for the storage efficiency[1]

$$\eta=4\beta\tau_p(1-\exp(-t_i(1+\beta)/2\tau_p))^2/((1+\beta)^2 t_i), \tag{2}$$

where τ_p is the power time constant of the cavity.

Differentiating η with respect to $x=t_i/\tau_p$, expression (2) gives the condition for maximum η in terms of x for given β

$$\exp(z/2)=1+z, \tag{3}$$

where $z=(1+\beta)x$, which yields formula (1).

For instance, for $\beta=1$ formula (1) gives the well-known optimum $t_i=1.256\tau_p$ for which the storage efficiency is approximately 0.41. However, this does not imply that when the input pulse length is $1.256\tau_p$, the maximum efficiency corresponds to the critical coupling $\beta=1$. It is easy to see from formula (2) that in this case, the maximum efficiency is obtained for β satisfying the equation

$$\exp(z/2)=1+z\beta/(\beta-1). \tag{4}$$

In this case, the solution of Eq. (4) is $\beta\approx 2.5$ and the storage efficiency is then ≈ 0.515.

Figure 1 gives the storage efficiency as a function of β for $x=1.256\tau_p$ and as a function of x for $\beta=1$.

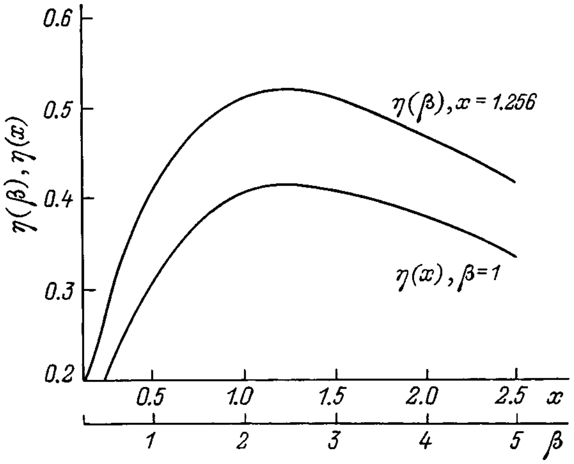


FIG. 1. Storage efficiency as a function of input coupling coefficient β for a fixed normalized input pulse length x and as a function of pulse length x for fixed coupling β.

Thus, formula (1) can only be used to find the optimum input pulse length for a given input coupling coefficient and formula (4) must be used to determine the optimum coupling for a given input pulse length to avoid errors.

---

[1] S. V. Baraev and O. P. Korovin, Zh. Tekh. Fiz **50**, 2465 (1980) [Sov. Phys. Tech. Phys. **25**, 1444 (1980)].

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Loading characteristics of an electrostatic coaxial lens

L. P. Ovsyannikova and T. Ya. Fishkova

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

(Submitted March 4, 1998)
Zh. Tekh. Fiz. **69**, 123–126 (June 1999)

The position of the image is calculated as a function of the position of the object (tubular beam) for an electrostatic system consisting of two coaxial cylinders and an end diaphragm with an annular slit. © *1999 American Institute of Physics.* [S1063-7842(99)02506-4]

In an earlier study[1] we determined the operating conditions of a coaxial cylindrical lens (CCL) consisting of two cylinders and a planar diaphragm on the charged particle entry side which focuses a tubular beam onto the axis of the lens. The cardinal elements of the image space were calculated numerically. For these we obtained simple numerical formulas which give an accuracy better than 25% and are convenient for engineering calculations.[2]

The aim of the present study is to calculate the loading characteristics of a CCL (the position of the image as a function of the position of the object). In addition, the authors show that a CCL with a tubular beam focused onto the axis satisfies the classical lens formula which is well-known in optics as the Newton formula.

A coaxial cylindrical lens with the front end closed by a diaphragm and the rear end open belongs to a class of lenses for which the cardinal elements of the object and image space differ in magnitude. This is because no field exists in the object space whereas a field does exist in the image space and at a distance from the lens greater than its aperture, it becomes homogeneous.[3] The calculations were made numerically using a program developed by the authors for electrostatic electron-optics systems possessing rotational symmetry. The CCL geometry was selected on the basis of previous studies made by the authors: the ratio between the radii of the external and internal cylinders was $R/\rho=2$–100, the gap between the input diaphragm and the external cylinder was $s/R=0.1$ and the length of the lens was $l=R$. A working voltage V is applied to the external cylindrical electrode with the internal electrode and the diaphragm grounded. For this geometry and power supply the tubular beam is focused onto the axis of the lens over the entire range of variation of the distance between the rear edge of the lens and the image.

Figure 1 shows a schematic of the CCL and its cardinal elements for the object space as a function of the excitation of the lens $\gamma=eV/\varepsilon$, where e is the particle charge and ε is the energy. We used numerical calculations to obtain empirical formulas for the cardinal elements of the object space: the focal length f_0, the position of the principal plane h_0, and the initial radius r_i of the family of central beam trajectories on the side of the image space, about which the focusing takes place

$$f_0/R=2[(1-\rho/R)/\gamma-0.3],$$

$$h_0/R=0.3\ (\gamma-1+1.1\sqrt{\rho/R}),$$

$$r_1/R=0.2(1.5\rho/R-1)\gamma+0.4\sqrt{\rho/R}+0.65. \qquad (1)$$

The crosses in Fig. 1 give the values calculated using the formulas (1). It can be seen that these differ negligibly from the numerical calculations.

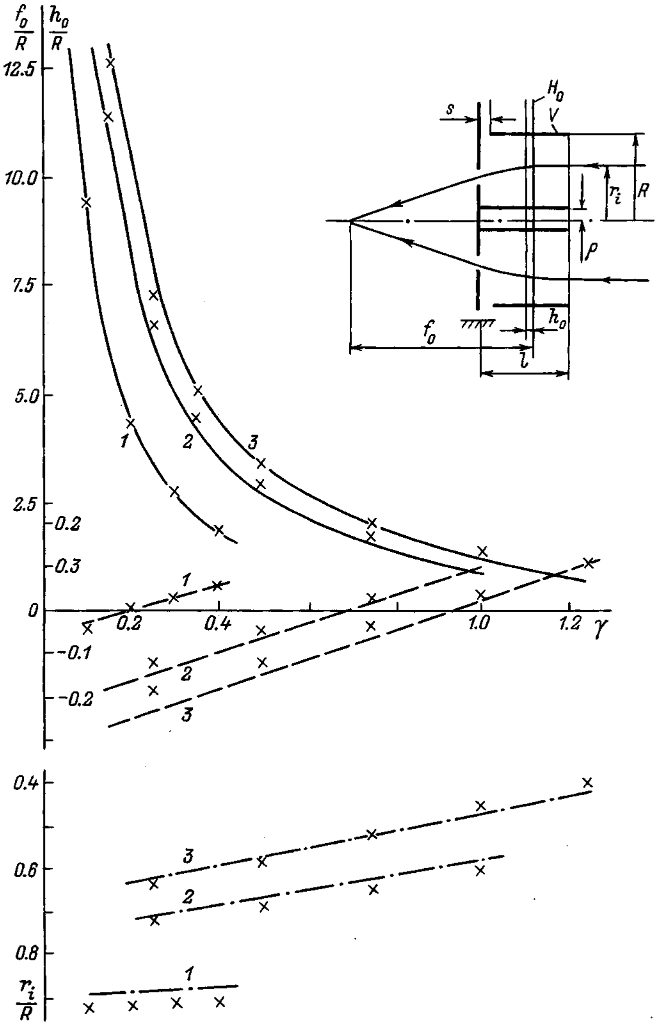


FIG. 1. Cardinal elements of the object space of a CCL of length $l=R$: focal length (solid curves), position of principal plane (dashed curves), and entry radius of axial beam trajectories (dot-dash curve) for various ratios of the radii of the external and internal cylinders: $R/\rho=2$ (*1*), 10 (*2*), and 100 (*3*).

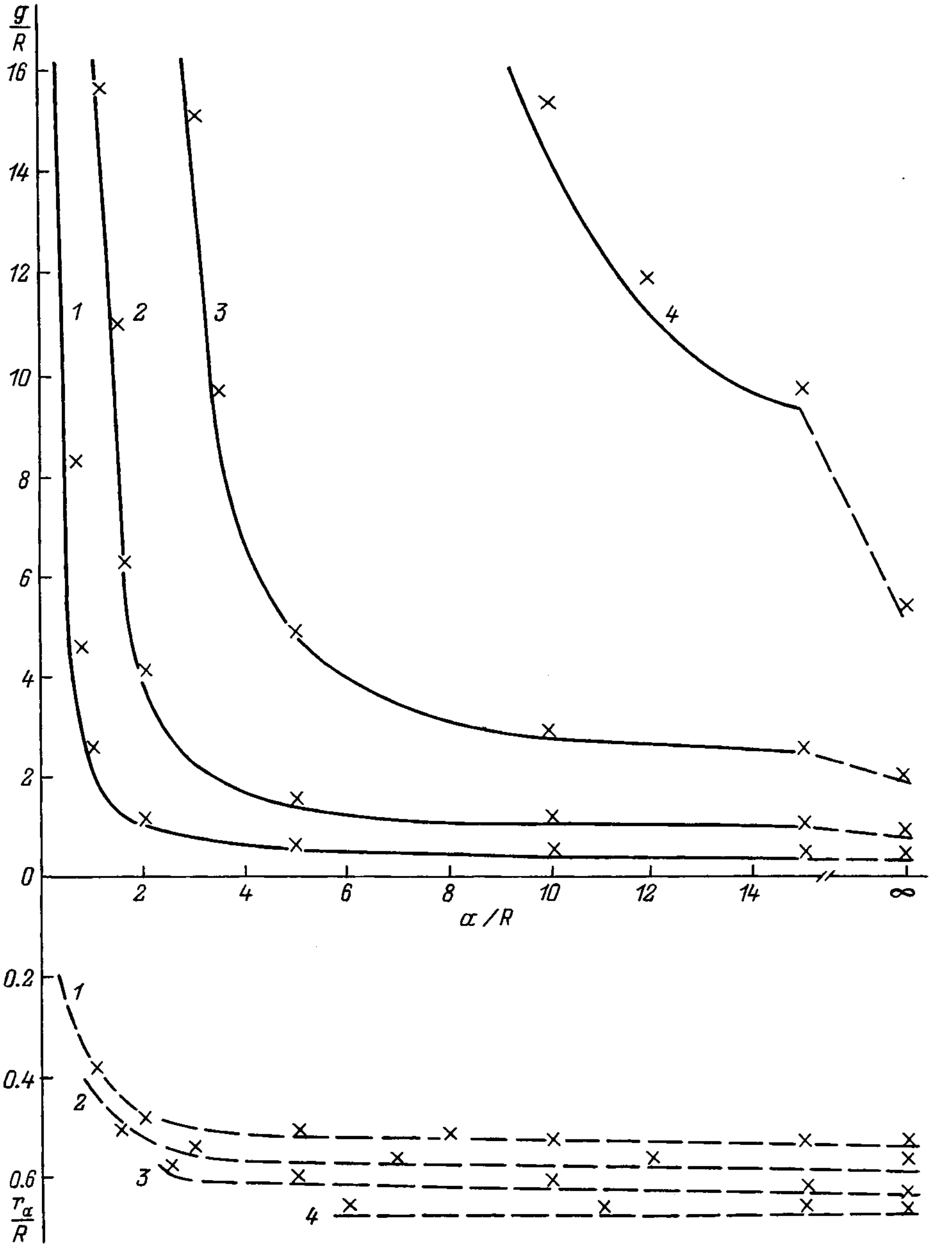


FIG. 2. Position of the image (solid curves) and radius of the central beam trajectories on entry to the lens (dashed curves) as a function of the position of the object for various excitations of a lens having $R/\rho = 10$ and $l = R$: $\gamma = 1.0$ (*1*), 0.75 (*2*), 0.5 (*3*), 0.25 (*4*).

It is of practical interest to obtain the loading characteristics of the lens which determine the relationship between the position of the object and its image. By varying the distance between the object and the entrance to the lens a for a specific lens power, we determined the distance between the exit and the image g for axis–axis focusing. It should be noted that in this case, the radius of the central beam trajectories at the entrance to the lens, about which the beam is focused, depends on the position of the object. This radius r_a was also determined numerically and the results of the calculations are given in the form of graphs and tables.

Figure 2 gives the loading characteristics of a CCL with the typical geometry $R/\rho = 10$ and $l = R$ (solid curves) and also the corresponding values of the radius r_i (dashed curves). Table I (upper line) gives the loading characteristics for a CCL with $R/\rho = 2$ and $l = R$ while Table II (upper line) gives those for $R/\rho = 100$ and $l = R$. All the geometric parameters are given in units of the external cylinder radius R.

TABLE I. $R/\rho = 2$, $l = R$.

γ	0.1		0.2		0.3		0.4	
a	g	r_a	g	r_a	g	r_a	g	r_a
∞	7.9	0.88	3.3	0.84	!.8	0.82	1.1	0.80
	8.0	0.89	3.2	0.86	1.7	0.83	0.90	0.80
15	20.1	0.88	5.0	0.84	2.4	0.82	1.4	0.79
	21.1	0.89	4.7	0.86	2.1	0.83	1.1	0.80
10	64.0	0.86	6.3	0.84	2.8	0.82	1.6	0.79
	83.3	0.88	5.9	0.86	2.4	0.83	1.2	0.80
5	···	···	17.5	0.83	4.8	0.82	2.3	0.78
			18.3	0.84	3.8	0.82	1.6	0.80
3	···	···	···	···	15.2	0.78	4.2	0.77
					9.2	0.81	2.5	0.78
2	···	···	···	···	···	···	13.0	0.74

TABLE II. $R/\rho=100,\ l=R$.

γ	0.25		0.5		0.75		1.0	
a	g	r_a	g	r_a	g	r_a	g	r_a
∞	6.7	0.58	2.5	0.56	1.1	0.52	0.46	0.48
	5.8	0.56	2.2	0.52	0.94	0.48	0.49	0.45
15	14.3	0.58	3.5	0.56	1.4	0.51	0.59	0.48
	11.8	0.55	2.9	0.51	1.2	0.48	0.58	0.44
10	29.4	0.58	4.2	0.56	1.6	0.51	0.66	0.48
	21.7	0.54	3.4	0.51	1.3	0.47	0.62	0.43
5	⋯	⋯	8.6	0.54	2.3	0.50	0.90	0.46
			6.8	0.49	1.9	0.46	0.78	0.42
3	⋯	⋯	∞	0.51	4.8	0.47	1.4	0.44
				0.47	3.1	0.43	1.1	0.40
2	⋯	⋯	⋯	⋯	10.7	0.44	1.9	0.42
					8.3	0.41	1.6	0.38
1	⋯	⋯	⋯	⋯	⋯	⋯	7.7	0.33
							8.6	0.31

It should be noted that substituting these numerical values of the focal lengths of the object and image spaces into the classical formula reveals good agreement with the numerical values plotted on the graph and given in the tables. Exceptions are the operating regimes where the focal length is smaller than the lens aperture, when the focus falls within the strong field. In this case, the Newton formula cannot be applied, as is the case for all other lenses.

On the basis of the numerical calculations, we obtained a simple empirical formula for the radius of the central beam trajectories

$$r_a=r_0-0.05[3+1/Ln(\rho/R)]/a, \qquad (2)$$

where r_0 is the initial radius of the central trajectories of a beam parallel to the axis on entry to the lens ($a\rightarrow\infty$) for which we derived an empirical formula in Ref. 2.

The values of r_a calculated using formula (2) are indicated by the crosses in Fig. 2 and are given in Tables I and II (lower lines). The maximum deviation between the empirical values and the numerical calculations does not exceed 10%.

It is of considerable practical interest to demonstrate the possibility of using the Newton formula in which empirical expressions are substituted for the cardinal elements of the CCL. We performed these calculations for the coaxial lens geometry indicated above using a refined empirical formula for the focal length of the image space

$$f_i/R=1.9(1-\rho/R)/\gamma-1.0. \qquad (3)$$

The results of the calculations are presented in Tables I and II (lower lines) and are shown by the crosses in Fig. 2. It can be seen that in the region of existence of a real image, the Newton formula is satisfied to within at least 25% (Fig. 2). The accuracy of determining the position of the image for a given position of the object deteriorates as the object approaches the focus of the lens because of the specific features of the Newton formula.

To sum up, we have calculated the loading characteristics of an electrostatic lens consisting of two coaxial cylinders and an adjacent planar input diaphragm under conditions of point-to-point focusing. The calculations were made by computer and using empirical formulas. It was shown that the classical lens formula is satisfied for a CCL having this geometry.

[1] L. P. Ovsyannikova and T. Ya. Fishkova, Pis'ma Zh. Tekh. Fiz **22**(16), 39 (1996) [Tech. Phys. Lett. **22**, 660 (1996)].

[2] L. P. Ovsyannikova and T. Ya. Fishkova, Zh. Tekh. Fiz **67**(12), 81 (1997) [Tech. Phys. **42**, 1444 (1997)].

[3] L. P. Ovsyannikova, S. V. Pasovets, and E. V. Shpak, Nucl. Instrum. Methods Phys. Res. A **298**, 344 (1990).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Surface flow of liquid in an electric field

V. A. Semenov

Perm State University, 614600 Perm, Russia

(Submitted March 5, 1998)
Zh. Tekh. Fiz. **69**, 127–128 (June 1999)

An analytic solution is proposed for a model problem which demonstrates the occurrence of a surface flow of weakly conducting liquid in an electric field which was previously observed experimentally by the author. © *1999 American Institute of Physics.* [S1063-7842(99)02606-9]

Berezhnov and Semenov[1] reported results of an experimental investigation of a new type of surface flow observed at the free surface of a weakly conducting liquid in a static electric field. They established that unlike the electrocapillary motion of mercury droplets,[2] in this effect the surface velocity is proportional to the square of the voltage between the electrodes. In the present paper the existence of this effect is demonstrated analytically.

We shall assume that an infinite air bubble of radius R is located inside an infinite cylinder of radius b filled with a weakly conducting liquid (planar problem). The axes of the cylinder and the bubble are the same. We shall assume that at the surface of the cylinder the potential distribution is described by $U_0 \cos\varphi$, where φ is the angle in the polar coordinate system (r,φ) measured from the x axis, perpendicular to the cylinder axis, and U_0 is the characteristic potential difference. Since the conductivity of air and liquid differ, a free charge forms at the free surface of the liquid which can be determined by solving the following problem in the approximation of weakly conducting air:

$$\Delta u_j = 0, \quad j = 1,2,$$

$$r = b: \quad u_2 = U_0 \cos\varphi,$$

$$r = R: \quad u_1 = u_2,$$

$$\sigma \frac{\partial u_2}{\partial r} = -\sigma_s \Delta_s u_2 + \mathrm{div}_s(\mathbf{v}_s \tau) - \nabla_s \sigma_s \nabla_s u_2,$$

$$4\pi\tau = \frac{\partial u_1}{\partial r} - \varepsilon \frac{\partial u_2}{\partial r}. \tag{1}$$

Here u_1 and u_2 are the potential in the bubble and in the liquid, σ and ε are the conductivity and the permittivity of the liquid, respectively, σ_s is the surface conductivity, Δ_s is the surface Laplacian, $\mathbf{v}_s$ is the surface velocity, ∇_s is the surface gradient, and τ is the free surface charge. When current flows, the component of the electric field strength tangential to the free surface is nonzero. As a result, a force will act on the free surface charge which is responsible for the onset of surface flow. We shall find the steady-state velocity of this flow in the approximation of low Reynolds numbers assuming that the free surface is nondeformable and the viscosity of the air is low. For this case we have the following equations:

$$\mathrm{div}\,\mathbf{v} = 0, \quad -\nabla p + \eta \Delta \mathbf{v} = 0,$$

$$r = b: \quad \mathbf{v} = 0,$$

$$r = R: \quad v_n = 0, \quad \sigma_{i,k} n_k = \tau E_i, \tag{2}$$

where $\mathbf{v}$ is the velocity in the liquid, E_i is the strength component tangential to the free surface, η is the viscosity of the liquid, $\sigma_{i,k}$ is the viscous stress tensor, p is the pressure, and n is the unit vector of the normal.

After dedimensionalizing ($r \to R$, $u \to U_0$, $v \to U_0^2/4\pi R\eta$, $\tau \to U_0/4\pi R$, $\sigma_s \to \sigma_s/\sigma R$, $b \to b/R$) assuming that the surface current induced by the Ohmic conductivity is much smaller than the current produced by the surface flow, we have

$$\Delta u_j = 0, \quad j = 1,2,$$

$$\mathrm{div}\,\mathbf{v} = 0, \quad \Delta \mathbf{v} = \nabla p,$$

$$r = b: \quad u_2 = \cos\varphi, \quad v_r = 0, \quad v_\varphi = 0,$$

$$r = 1: \quad u_1 = u_2, \quad \frac{\partial u_2}{\partial r} = \beta v_\varphi \nabla_\varphi \tau, \tag{3}$$

$$\tau = \frac{\partial u_1}{\partial r} - \varepsilon \frac{\partial u_2}{\partial r},$$

$$v_r = 0, \quad \frac{1}{r}\frac{\partial v_r}{\partial \varphi} + \frac{\partial v_\varphi}{\partial r} - \frac{v_\varphi}{r} = \tau E_\varphi,$$

$$\beta = \frac{U_0^2}{16\pi^2 R^2 \eta\sigma}. \tag{4}$$

Solving the problem (3) and (4) by expanding as a series in powers of β ($\beta \ll 1$) we find the velocity at the free surface of the liquid

$$v_\varphi^{(s)} = v_0 \sin 2\varphi + (v_1 \sin 2\varphi + v_2 \sin 4\varphi)\beta + \ \ldots,$$

$$v_0 = \frac{b^2(b^2-1)}{2\varepsilon(b^2+1)^3},$$

$$v_1 = -\frac{b^4(b^2-1)^2(\varepsilon b^2 + \varepsilon - 1 + b^2)}{4\varepsilon^3(b^2+1)^7},$$

1063-7842/99/44(6)/2/$15.00

$$v_2=\frac{b^4(b^2-1)(\varepsilon b^6+\varepsilon+2b^6-2)(b^{10}+3b^8+6b^6-6b^4-3b^2-1)}{16\varepsilon^3(b^2+1)^6(b^6+1)(b^8+2b^6+4b^4+2b^2+1)}. \tag{5}$$

The solution (5) indicates that the surface velocity is proportional to the square of the characteristic potential difference U_0, which agrees qualitatively with the experimental results obtained in Ref. 1.

[1] V. V. Berezhnov and V. A. Semenov, Pis'ma Zh. Tekh. Fiz **22**(5), 92 (1996) [Tech. Phys. Lett. **22**, 223 (1996)].
[2] V. G. Levich, *Physicochemical Hydrodynamics* [in Russian], Academy of Sciences of the USSR Press, Moscow (1952), 553 pp.

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Influence of an internal getter in silicon on the parameters of Au–Si structures

V. K. Kiselev, S. V. Obolenskiĭ, and V. D. Skupov

N. I. Lobachevskiĭ Nizhniĭ Novgorod State University, 603600 Nizhniĭ Novgorod, Russia

(Submitted March 30, 1998)
Zh. Tekh. Fiz. **69**, 129–131 (June 1999)

It is shown that internal gettering of impurities and defects in Au–Si structures can enhance their reliability and stabilize their characteristics.
[S1063-7842(99)02706-3]

Silicon–gold structures are extensively used as the basis for fabricating photoconverters (diodes and controllable switches) and devices for measuring the physicochemical parameters of external media, such as gas sensors. As in other types of semiconductor devices, the electrophysical, functional, and reliability characteristics of Au–Si structures depend very strongly on the presence of accidental (background) impurities and other crystallographic defects in the silicon substrate. Various gettering procedures are used to reduce the concentration of impurities and defects, of which internal gettering is held to be the most promising method for silicon.[1] As yet, the positive capabilities of this comparatively new method, especially for enhancing the radiation resistance of devices, have not yet been fully identified. The aim of the present study is to investigate the influence of internal gettering on the photoelectric properties of Au–Si structures bombarded by moderate-energy protons.

The structures used for the investigations consisted of twenty elements arranged in an interdigital topology on dislocation-free Czochralski-grown KDB-12 (001) silicon crystals by thermal deposition of a gold layer followed by brazing. A 2 μm layer of galvanic gold was grown on the electrodes to lower the resistance. The width of the interelectrode region was 50 μm. Data obtained by infrared absorption spectrometry indicate that the oxygen content in the initial substrates did not exceed $1.4\times10^{18}\,\mathrm{cm}^{-3}$ and the average microdefect density established from selective etching patterns in a Sirtl solution was $1.4\times10^{5}\,\mathrm{cm}^{-2}$. The x-ray diffraction patterns of the initial crystals revealed a dominant (004) allowed peak according to the extinction conditions in silicon together with ''forbidden'' (002) and (006) reflections whose intensity is related to that of the main peak as $I_{(002)}/I_{(004)}=3\times10^{-2}$ and $I_{(006)}/I_{(004)}=1.1-2.3\times10^{-2}$. The appearance of these forbidden peaks indicates that the initial samples contain regions of microstresses having a component normal to the reflecting surface.

Before the gold was deposited, one batch of substrates underwent internal gettering using a method described by Skupov.[2] This involved irradiating the initial samples with alpha particles from a ^{210}Po radionuclide source ($E=4.5\,\mathrm{MeV}$, $\Phi=10^{11}$–$10^{12}\,\mathrm{cm}^{-2}$) to activate nucleation centers and grow gettering inclusions during the subsequent three-stage annealing in a dry nitrogen atmosphere: 6 h at 1373 K, 12 h at 1073 K, and 8 h at 1373 K. The other batch of

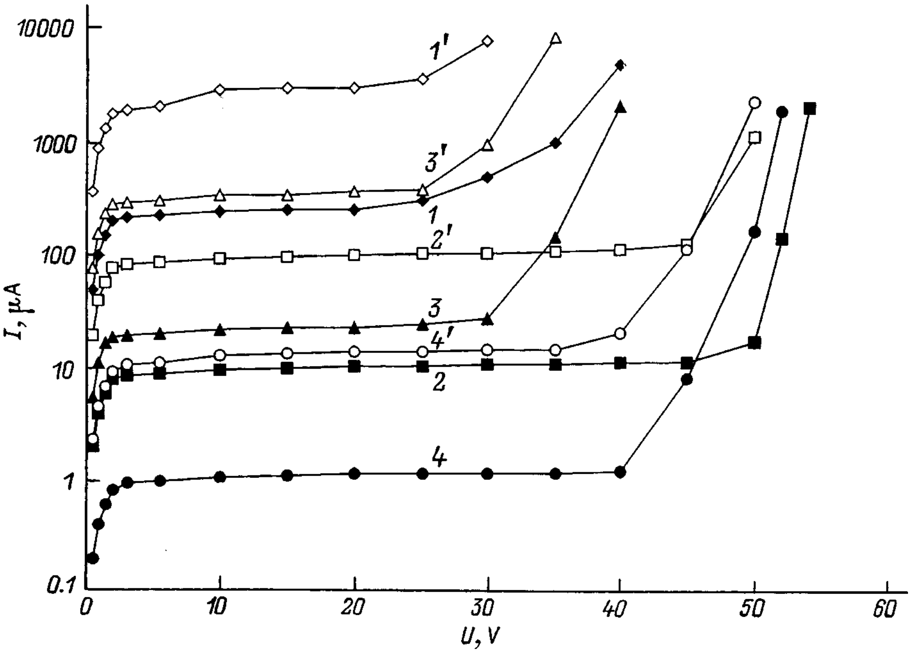


FIG. 1. Current–voltage characteristics of Au–Si structures without illumination (*1–4*) and with illumination (*1′–4′*): *1, 1′* — control samples, *2, 2′* — after gettering, *3, 3′* — control samples after proton irradiation, and *4, 4′* — after gettering and proton irradiation.

1063-7842/99/44(6)/2/$15.00

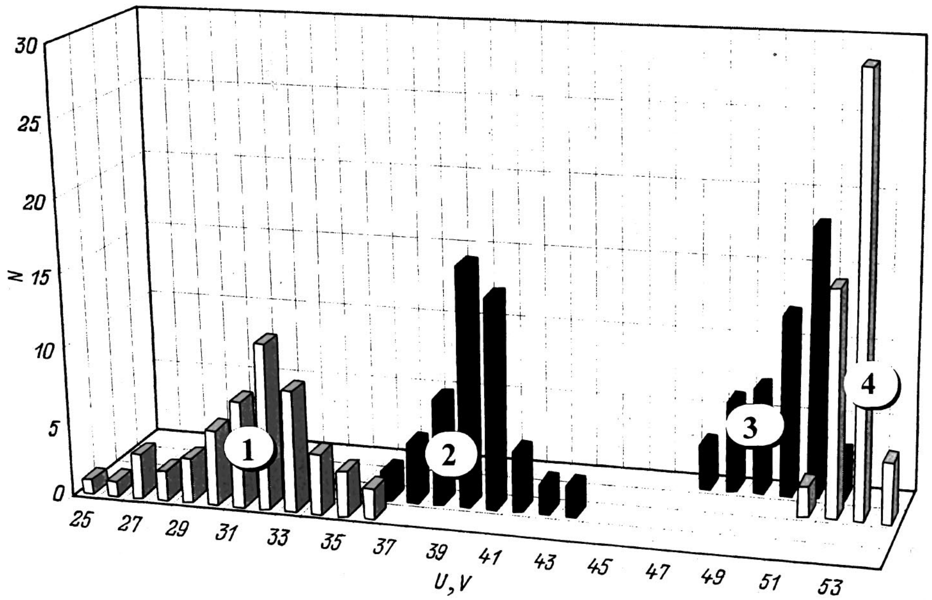


FIG. 2. Histogram showing the distribution of the structures in terms of breakdown voltage (without illumination): *1* — control samples, *2* — control samples after proton irradiation, *3* — after gettering, *4* — after gettering and proton irradiation; *N* is the structure number.

substrates was not annealed and the structures on them served as controls. It should noted that when the electrophysical characteristics were measured for various control samples, some stabilization of the resistance was observed, which showed up as an abrupt increase in current from a few to hundreds of microamperes when the voltage at the contacts was increased to 10–15 V. This effect indicates some instability of the impurity-defect composition of the initial substrates and the Au–Si transition region which is not completely eliminated by conventional thermal brazing.

Some of the structures in each batch were irradiated by a proton dose of $6.25\times 10^{11}\,\mathrm{cm}^{-2}$ at increasing energies of 30, 60, and 90 keV to suppress surface conduction channels. Before and after irradiation we recorded the current–voltage characteristics without illumination and under illumination by 0.85 μm 40 mW radiation.

The experiments showed that as a result of internal gettering in the surface zone 7–10 μm deep, the microdefect density decreased to $9\times 10^{3}\,\mathrm{cm}^{-2}$ and the density dispersion decreased to $2.9\times 10^{3}\,\mathrm{cm}^{-2}$. The (002) and (006) forbidden peaks disappeared on the x-ray diffraction patterns, i.e., the level of microstresses in the crystals was reduced. In contrast to the control samples, no instability of the electrical parameters was detected at the initial time of measurement for any of the structures formed on the gettered substrates.

Figure 1 gives the current–voltage characteristics of these structures, which show that after gettering the resistance increases and the structures reveal a more abrupt transition to the prebreakdown state, both without illumination and under photoexcitation of nonequilibrium carriers. These observations correlate with data on the dissolution of microdefects under gettering and suggest that the main channels for current flow and microplasma formation in the initial silicon are growth clusters and clusters of intrinsic point defects and impurities formed during the aggressive chemical treatment of the substrates, as well as their surrounding impurity atmospheres. Indirect confirmation of this may be provided by the increase in the resistance of the structures after proton irradiation, when additional gettering processes may be activated by radiation point defects and elastic waves.[3] Measurements using batches of fifty structures revealed that after gettering, the breakdown voltage increases by an average of 62% whereas after irradiation of the control samples it increases by 22% (Fig. 2). Some reduction in the breakdown voltage after proton irradiation of gettered structures is clearly attributable to the electrical activity of the incorporated radiation defects whose influence, however, is not so appreciable (less than 2%) as the gettering. This is also evidenced by the fact that gettered and irradiated structures are stable under the action of short-term (up to 1 s) loads of 55–60 V unlike the irradiated control samples.

These results suggest that internal gettering of impurities and defects in the substrates can stabilize the parameters and enhance the functional reliability of silicon–metal structures, including that under irradiation.

[1] G. Z. Nemtsev, A. I. Pekarev, and Yu. D. Chistyakov, Mikroelektronika **12**, 432 (1983).

[2] V. D. Skupov, *Abstracts of Papers presented at First All-Russia Conference on Materials Science and Physicochemical Principles of Fabrication Technologies for Doped Silicon Crystals* [in Russian], Moscow, 1996, p. 127.

[3] P. V. Pavlov, Yu. A. Semin, V. D. Skupov, and D. I. Tetel'baum, Fiz. Tekh. Poluprovodn. **20**, 503 (1986) [Sov. Phys. Semicond. **20**, 315 (1986)].

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Sulfide passivation of a textured interface of a gallium arsenide surface-barrier photovoltaic cell

N. L. Dmitruk, O. Yu. Borkovskaya, and I. B. Mamontova

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

(Submitted April 2, 1998)
Zh. Tekh. Fiz. **69**, 132–134 (June 1999)

A comparative study is made of the influence of sulfide passivation in an aqueous $Na_2S \cdot 9H_2O$ solution on the photoconversion parameters of solar radiation in Au–GaAs barrier structures as a function of the character of the microrelief and the pretreatment of the GaAs surface. A quasigrating and a dendritic surface microrelief were produced by anisotropic chemical etching. It is shown that this type of GaAs surface treatment is potentially useful for enhancing the efficiency of a photovoltaic cell stored for several years. A possible mechanism is discussed for the processes leading to changes in the structure parameters. *© 1999 American Institute of Physics.* [S1063-7842(99)02806-8]

A GaAs surface microrelief produced by anisotropic chemical etching can reduce the reflection coefficient of light in the visible and in the near ultraviolet and infrared to a few percent, provided that its morphology and depth are suitably optimized.[1] This is particularly important when such a surface is used as the interface in solar cells formed by a metal–semiconductor contact since even thin (<10 nm) layers of metal (Au) increase the reflection coefficient to 45–55% (from 30–35% for GaAs).[2] Depending on the microrelief morphology and the thickness of the metal film, the photosensitivity of these structures can be enhanced compared with that of planar structures, between 1.5–2 and several tens of times in different parts of the spectrum.[3] However, the presence of pyramidal protrusions and acutely angled faces of a dendritic relief intensifies the field in these parts of the surface, induces a thermal-field current component, and effectively reduces the barrier height,[4] which should be observed as a reduction in the shunt resistance R_{sh} and a drop in the open-circuit voltage V_{oc} of the solar cell. Quite clearly, these effects may be suppressed by means of an intermediate passivating layer. Here we examine the possibility of suppressing these effects by sulfide passivation of a GaAs surface in an $Na_2S \cdot 9H_2O$ aqueous solution, bearing in mind its positive influence on the photosensitivity of these structures,[5] and also the presence of an etching phase during this treatment[6,7] which may slightly smooth the acute-angled relief.

We made a comparative study of the loading optical current–voltage characteristics and photocurrent spectra of Au–GaAs barrier structures fabricated by vacuum deposition of semitransparent Au layers on planar and microrelief GaAs surfaces with a natural layer of oxide or with the oxide removed by etching. Some of the wafers were also passivated in a 2N $Na_2S \cdot 9H_2O$ aqueous solution (20 s) followed by rinsing in distilled water. The substrate temperature during deposition of the metal was ~110 °C. In order not to complicate the comparative analysis, no antireflection coatings or contact grid were used and the thickness of the metal film was varied between 15 and 25 nm and was determined by multiangular ellipsometry using quartz satellite plates. The structures were fabricated using single-crystal *n*-GaAs wafers doped to between 1×10^{16} and $3\times10^{17}\,\mathrm{cm}^{-3}$ and (100)-oriented epitaxial n–n^+-type films ($n=1\times10^{15}$–3×10^{16} cm^{-3}). A dendritic microrelief was fabricated by etching in concentrated HNO_3 (Ref. 1) and a variable-period quasigrating microrelief was obtained using $2HF:2H_2SO_4:1H_2O_2$ etchant.[5] The oxide was removed by etching in an HCl solution.

Figure 1a shows spectra of the short-circuit photocurrent I_{sc} per incident photon for Au–GaAs diode structures with different surface microrelief and different pretreatment. It can be seen that the effect of the sulfide passivation depends on the initial state of the surface. The maximum increase in I_{sc} is observed for structures with a planar or microrelief surface from which the oxide layer was not removed by etching (curves *1*, *1′*, *2*, and *2′*). For relief surfaces with the oxide removed, which gives the maximum I_{sc} sulfide passivation has a considerably weaker influence on I_{sc}, mainly in the short-wave part of the spectrum (curves *3*, *3′*, *4*, and *4′*). By analyzing the spectra of the internal quantum efficiency of these diode structures using a model considered in Ref. 3 taking account of the light transmission spectrum in an air–metal–intermediate oxide layer–semiconductor structure calculated using the method described in Ref. 8, we can determine the recombination parameters of the structure, i.e., the ambipolar carrier diffusion length, the surface recombination velocity S, and the velocity of above-barrier transport of majority carriers V_n into the metal. We established that sulfide passivation reduces the velocity S of the Au–GaAs interface by one or two orders of magnitude for structures with a natural oxide layer and by a factor of 1.2–4 for structures with a textured surface and the oxide layer removed. In this last case, V_n also decreases which is observed as an enhancement of the photosensitivity in the short-wavelength part of the spectrum. Figure 1b shows the influence of sulfide passivation on the optical current–voltage characteristics of

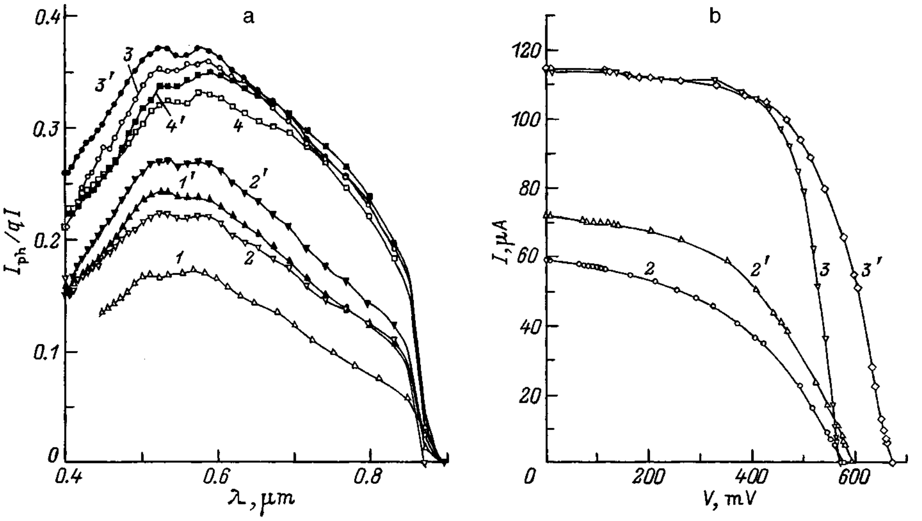


FIG. 1. Spectra of short-circuit photocurrent (a) and optical current–voltage characteristics (b) of Au–GaAs structures with a planar surface (*1, 1'*), a quasigrating microrelief (*2, 2', 3, 3'*) and dendritic microrelief (*4, 4'*) after preliminary removal of the oxide (*3, 3', 4, 4'*) and sulfide passivation (*1'–4'*); *n*-GaAs single-crystal wafer with $N_d = 3\times10^{16}\,\mathrm{cm}^{-3}$, and gold film 20 nm thick.

Au–GaAs structures with a quasigrating microrelief measured using a sun simulator under conditions corresponding to zero atmospheric mass. It can be seen that sulfide passivation improves the characteristics of the solar cell in structures with and without an oxide layer. Quite remarkably, in this last case we observe an increase in V_{oc} which is not associated with a drop in the surface recombination velocity. Figure 2 shows the optical current–voltage characteristics of structures with a dendritic surface microrelief from which the oxide layer has been removed. We can see that in this case, sulfide passivation improves the fill factor of the characteristic and gives a corresponding increase in the efficiency of the solar cell η. The equivalent circuit parameters (series resistance R_s and shunt resistance R_{sh}) of the microrelief structures determined from the optical characteristics and also the change in the solar energy conversion parameters of the solar cell as a result of sulfide passivation ($I_{\mathrm{scs}}/I_{\mathrm{sc0}}$, $U_{\mathrm{ocs}}/U_{\mathrm{oc0}}$, η_s/η_0) are presented in Table I. It can be seen that sulfide passivation of a textured GaAs surface can substantially enhance the efficiency of a surface-barrier photovoltaic cell by increasing the fill factor, V_{oc}, and R_{sh} and reducing R_s.

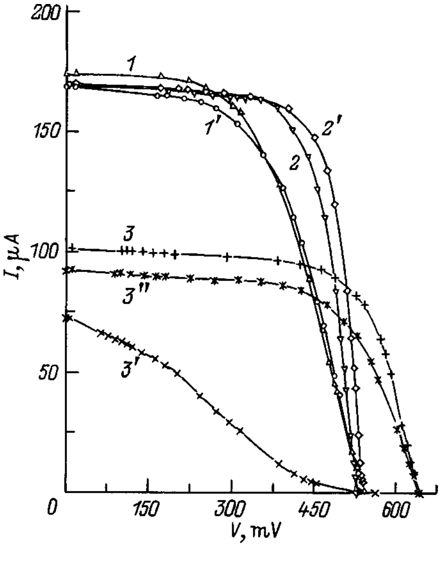


FIG. 2. Optical current–voltage characteristics of Au–n–n^+-GaAs structures with a dendritic surface microrelief (*1, 1', 2, 2'*), including after sulfide passivation (*2, 2'*), and a planar GaAs surface with a natural layer of oxide (*3, 3', 3''*): *1'–3', 3''* — six months after *1–3*; *3''* — after treating the structure in an HCl solution; $N_d = 7.4\times10^{15}\,\mathrm{cm}^{-3}$, gold layer 16.5 nm thick. The increase in efficiency as a result of the microrelief is $\eta_1/\eta_3 = 1.16$ and as a result of the microrelief plus passivation $\eta_2/\eta_3 = 1.46$.

The investigations showed that this GaAs surface treatment prevents the degradation of the characteristics of Au–GaAs photovoltaic cells, which remain almost constant for between two and four years. Figure 2 shows the characteristics of structures after different treatment of a microrelief surface obtained at six-monthly intervals. For comparison we also show the characteristics of planar structures with a natural layer of oxide on the GaAs surface which deteriorate more rapidly, mainly as a result of a drop in R_{sh}. Studies using photodiode structures having different areas showed that the shunt resistance R_{sh} is inversely proportional to the diode diameter, i.e., is mainly caused by leakage across its perimeter. We can hypothesize that as under annealing,[8] the local stoichiometry of the interface plays a major role in the degradation of these structures with time. The natural oxide layer of GaAs is characterized by excess gallium in the form Ga_2O_3 (Ref. 9), whereas the arsenic is in the form As_2O_3 (up to 35%) as in the unoxidized state, and the free arsenic possesses fairly high mobility.[8] Since the conditions at the edge of the diode are more favorable for its emergence at the surface, the oxide Ga_2O_3 may predominate here with mainly gallium diffusing through it. We postulate that this results in the formation of channels for the interdiffusion of gallium and gold atoms at room temperature (for example, as a result of recombination-stimulated processes under illumination). These channels may shunt the photovoltaic cell and ultimately result in its degradation. Treatment of this diode in an HCl solution, as a result of which the oxide becomes thinner and severely depleted in gallium around the perimeter of the diode,[9] almost restores the optical characteristics (curve *3''*). For planar structures, R_s is proportional to the diode area whereas for textured structures its value is substantially

TABLE I.

Surface state	R_s, kΩ	R_{sh}, kΩ	FF	I_{scs}/I_{sc0}	U_{ocs}/U_{oc0}	η_s/η_0
1) Relief—quasigrating	3.74	63.3	0.45	–	–	–
with oxide + passivation	3.16	86.5	0.49	1.22	1.03	1.38
2) Relief—quasigrating,	1.08	58.5	0.62	–	–	–
oxide removed + passivation	1.45	84.5	0.62	1.0	1.19	1.18
3) Dendritic relief, oxide	0.81	53.5	0.56	–	–	–
removed + passivation	0.20	103.0	0.72	1.01	1.19	1.32

lower, i.e., it is mainly determined by the conductivity of the intermediate layer. This drop in the resistance as a result of treatment in an HCl solution and/or sulfide passivation of the GaAs surface correlates with the decrease in the thickness of the intermediate layer.[9,10] Thus, the increase in V_{oc} observed as a result of sulfide passivation cannot be attributed to an increase in the thickness of the intermediate layer but is caused by a change in the surface structure which in particular, leads to a reduction in the strong-field effects in the current flow.[4]

To sum up, sulfide passivation has a positive influence on the parameters of Au–GaAs photovoltaic cells with a textured surface not only by reducing the density of surface electronic states which determine the surface recombination velocity at the interface, which increases I_{sc} but also by changing the structure and chemical composition of the intermediate layer, which is observed as a decrease in R_s and an increase in R_{sh}, V_{oc}, and the fill factor. The proposed treatments can appreciably enhance the efficiency of the photovoltaic cell and also improve its degradation resistance.

[1] T. Ya. Gorbach, E. V. Pidlisnyĭ, and S. V. Svechnikov, Optoélektron. Poluprovodn. Tekh. No. 13, 34 (1988).

[2] P. A. Iles, J. Vac. Sci. Technol. **14**, 1100 (1977).

[3] O. Yu. Borkovskaya, N. L. Dmitruk, and O. N. Mishchuk, Fiz. Tekh. Poluprovodn. **25**, 487 (1991) [Sov. Phys. Semicond. **25**, 294 (1991)].

[4] O. Yu. Borkovskaya, N. L. Dmitruk, T. Ya. Gorbach, and O. N. Mishchuk, Élektron. Tekh. Ser. 2, No. 5, 50 (1989).

[5] N. L. Dmitruk, O. Yu. Borkovskaya, O. I. Mayeva *et al.*, Proc. SPIE **2999**, CT No. 12, 384 (1997).

[6] V. L. Berkovits, V. M. Lantratov, T. V. L'vova *et al.*, Fiz. Tekh. Poluprovodn. **28**, 428 (1994) [Semiconductors **28**, 260 (1994)].

[7] N. L. Dmitruk, O. V. Fursenko, and O. Yu. Borkovskaya, Optoélektron. Poluprovodn. Tekh. No. 27, 115 (1994).

[8] Z. Liliental-Weber, R. Gronsky, J. Washburn *et al.*, J. Vac. Sci. Technol. B **4**, 912 (1986).

[9] V. V. Nemoshkalenko, V. G. Aleshin, L. G. Gassanov *et al.*, Poverkhnost' No. 2, 88 (1983).

[10] H. Sugahara, M. Oshima, H. Oigawa *et al.*, J. Appl. Phys. **69**, 4349 (1991).

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Instability of the photoluminescence of porous silicon

A. M. Orlov and A. V. Sindyaev

Ulyanovsk State University, 432700 Ulyanovsk, Russia

(Submitted April 2, 1998)
Zh. Tekh. Fiz. **69**, 135–137 (June 1999)

Results of studies of the photoluminescence of porous silicon with different prehistories have revealed the mechanism and nature of the instability of the luminescence properties of freshly prepared samples. It was established that the initial quenching and subsequent rise of the photoluminescence is attributable to the intermediate formation of silicon monoxide (photoluminescence degradation) and subsequent additional oxidation to form SiO_2 (photoluminescence rise). Ultraviolet laser irradiation accelerates this process by a factor of 200–250 compared with passive storage of the samples in air. Plasma-chemical treatment in an oxygen environment merely results in a subsequent rise in the photoluminescence as a result of the formation of monoxide on the porous silicon surface. A kinetic model is proposed for this process. *© 1999 American Institute of Physics.* [S1063-7842(99)02906-2]

The luminescence properties of porous silicon in the visible were first detected less than a decade ago.[1] However, the unique nature of this effect, caused by the quantum dimensions of the silicon filaments and the nature of the chemisorbed atoms, provided the main impetus for its intensive study.

Here we report the main results of investigations to specify the mechanism and nature of the evolutional changes to the photoluminescence properties of porous silicon. Samples were prepared from [111]-oriented, boron-doped, dislocation-free *p*-Si wafers ($p = 2.6$–$3.1\,\Omega\cdot$cm. Electrochemical etching was carried out at an anode current density of 7–14 mA/cm^2 using a 1:1 mixture of 48% HF and ethyl alcohol. Each silicon wafer was first subjected to chemical etching for 2 min in a 1:2 mixture of 48% HF:HNO_3 acids. Photoluminescence was excited by an LGN-409 ultraviolet He–Cd laser ($\lambda = 325$ nm) at a radiation power of 3.0 ± 0.1 mW. The calculated power density of the ultraviolet flux taking into account the irradiated area, was 133.6 mW/cm^2. Each of the six–eight sections of sample was exposed to the continuous (up to 8 h) action of ultraviolet radiation with 14–16 h between successive irradiations. The processes taking place at the porous silicon walls were assessed from changes in the starting intensity $I_0(\tau)$ of the luminescent samples stored passively in air for time τ and from the changes in $I(t)$ during the continuous ultraviolet irradiation process. Repeated irradiation of any section was eliminated by accurately setting the preselected coordinates of a two-coordinate microscope stage.

Typical changes in the photoluminescence intensity as a function of the irradiation time for freshly prepared porous silicon samples and samples held previously in air are shown by curves *1* and *2* in Fig. 1. It can be seen that regardless of the preliminary holding time in air, the photoluminescence of the porous silicon always exhibits extreme changes. Rapid quenching at the initial stage of irradiation is accompanied by a subsequent smooth rise with a tendency to reach a steady-state level. Moreover, the time at which the extreme (minimum) point is reached, identified 15–20 min after ultraviolet irradiation, is almost independent of the storage time in air. At the same time, the numerical value of the starting intensity depends strongly on the storage time in air. As a result, the steady-state luminescence level can be either lower or higher than the initial value.

We established (curve *1* in Fig. 2) that the change in I_0 accompanying prolonged storage of the sample in air is similar to the change in the photoluminescence under continuous laser irradiation (curves *1* and *2* in Fig. 1). However, the length of the similar sections and the time shifts of the extremum points are 200–250 times shorter under laser irradiation. This indicates that the chemical processes taking place at the porous silicon walls are identical and may be represented by series-parallel reactions of the type

$$\mathrm{SiH}_x+\frac{x+2}{4}\,\mathrm{O}_2\Rightarrow\mathrm{SiO}+\frac{x}{2}\,\mathrm{H}_2\mathrm{O},$$

$$\mathrm{SiH}_x+\frac{1}{2}\,\mathrm{O}_2\Rightarrow\mathrm{SiO}+\frac{x}{2}\,\mathrm{H}_2, \qquad (1)$$

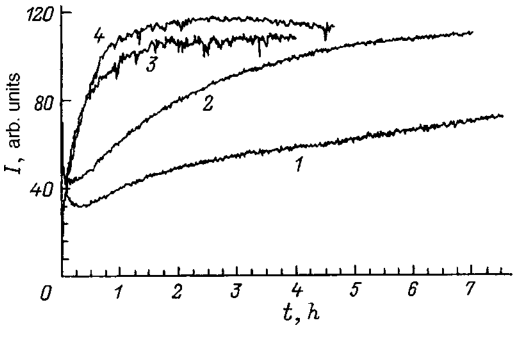


FIG. 1. Change in the photoluminescence of porous silicon during continuous laser irradiation after various holding times in air: $\tau_1 = 47.22$ h, $\tau_2 = 188.2$ h (the porous silicon did not undergo plasma chemical treatment), $\tau_3 = 22.2$ h, $\tau_4 = 4.5$ min (freshly prepared samples underwent plasma chemical treatment for 5 min). The subscripts 1–4 to τ_1 correspond to the numbering of the curves.

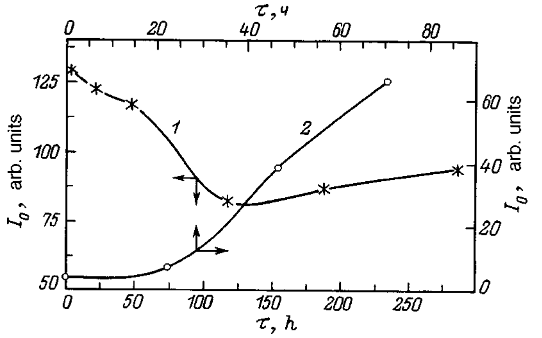


FIG. 2. Photoluminescence starting intensity of porous silicon as a function of the holding time in air: *1* — after anode etching and *2* — after additional plasma chemical etching for 5 min.

$$(R-\mathrm{O})+\frac{1}{2}\,\mathrm{O}_2 \Rightarrow (R-\mathrm{O}_2),$$

$$(R-\mathrm{O})+\frac{1}{2}\,\mathrm{H}_2 \Rightarrow (R-\mathrm{OH}), \tag{2}$$

where $x=1$ or 2 and R describes the chemical bond of the molecule with the silicon crystal lattice.

If the observed extreme changes, including those during the passive storage of porous silicon in air, can be attributed to the expulsion of adsorbed hydrogen by oxygen and subsequent additional oxidation of the porous silicon walls, at the first stage preceding the extreme value $I_{\min}$, reaction (1) should have priority and at the second stage reaction (2). This presupposes that the photoluminescence of porous silicon with an appropriate set of adsorbed atoms should satisfy the constraint $I_{\mathrm{SiO}_2} \geqslant I_{\mathrm{SiH}_x} \geqslant I_{\mathrm{SiO}}$, although the contribution of SiO and SiO_2 to the photoluminescence has not yet been determined. Thus, we carried out additional investigations using a high-frequency (13.65 MHz), low-pressure (2.7 Pa) oxygen plasma. The samples were subjected to plasma chemical treatment using a Plazma-600 device whose reaction chamber was a quartz cylinder with external capacitor plates. The plasma was excited and sustained by a generator having an output power up to 600 W with an anode current of 0.6–0.7 A and a grid current of 50 mA.

The plasma chemical treatment has a radical influence on the photoluminescence of the porous silicon, reducing the starting value I_0 approximately 26 times (curves *1* and *4* in Fig. 1) and completely eliminating the degradation section on the $I(t)$ curves for all the plasma treatment times being studied (1–40 min). This shows that the initial oxidation of the porous silicon walls is completed during the plasma chemical treatment, with the natural exception of reaction (1) which, in our view, is mainly responsible for quenching the photoluminescence at the initial stage of ultraviolet irradiation of the untreated samples. This is convincingly confirmed by the results of an Auger spectral analysis1) if plasma chemical treatment for 1 min gives 35 at. % silicon, 28.3% carbon, and 36.7% oxygen, after treatment for 40 min almost all the bonds at the surface are occupied by oxygen which corresponds to the stoichiometric composition of SiO monoxide (50.3 at. % O and 49.7% Si).

Complete filling of the surface bonds with oxygen should stabilize the light-emitting properties of the porous silicon.[2,3] However, the clearly defined rise on the $I(t)$ and $I_0(\tau)$ curves can only be caused by additional chemical processes, which must primarily include additional oxidation of silicon (2) and the possible adsorption of hydrogen or $\mathrm{H_2O}$ molecules to form OH^- hydroxyl groups which, according to infrared spectroscopy, are always present at the surface of quantum filaments[4] and are not subject to ultraviolet damage $E(h\nu)=3.82$, $E(\mathrm{Si{-}O})=4.4$, and $E(\mathrm{O{-}H})=4.8$ eV (Refs. 5 and 6).

We shall analyze these effects from the point of view of the kinetics of the processes. Let us assume that C_Σ is the total concentration of surface bonds, C_n is the concentration of nonradiative centers (which we associate with SiO), and C_1 and C_2 are the concentrations of radiative centers identified with SiO_2 and SiOH. Since $C_\Sigma=C_n+C_1+C_2=\Sigma C_i$ $=\mathrm{const}$, the rates of the parallel reactions (2) may be given by the obvious equations

$$\frac{dC_1}{dt}=k_1C_n=k_1(C_\Sigma-C_1-C_2), \tag{3}$$

$$\frac{dC_2}{dt}=k_2C_n=k_2(C_\Sigma-C_1-C_2), \tag{4}$$

where k_i are the constants of the appropriate reactions (2).

Assuming that $C_1+C_2=C$ and $k_1+k_2=k$, we express Eqs. (3) and (4) in the more compact form after summation

$$\frac{dC}{dt}=kC_n=k(C_\Sigma-C). \tag{5}$$

The solution of this equation for $C_n(0)=C_{0n}=C_\Sigma$ $-C(0)$ and $C(\infty)=C_\Sigma$ has the form

$$C=C_\Sigma-C_{0n}\exp(-kt). \tag{6}$$

The photoluminescence intensity is proportional to the total concentration C. Hence we have $I(t)=PC=P(C_1$ $+fC_2)$, where P is the proportionality factor and f allows for the difference in the luminescence properties of C_1 and C_2. By subsequently differentiating this equation and taking into account Eqs. (3) and (4), we obtain

$$I(t)=I_0+\frac{k_1+fk_2}{k_1+k_2}PC_{0n}\{1-\exp[-(k_1+k_2)t]\}$$
$$=F-G\exp(-k). \tag{7}$$

Here F and G are the corresponding constants, the subscript "0" satisfies the initial conditions $t=0$, and $I_0=P(C_{01}$ $+f\,C_{02})$. If C_{02}, and therefore k_2 are zero, we find

$$I(t)=PC_\Sigma-PC_{0n}\exp(-k_1t). \tag{8}$$

Equation (7) shows good agreement with the experiment if for curves *3* and *4* (Fig. 1) we have $F=107.38$ and 114.02, $G=75.68$ and 100.79, $k=2.29$, and $2.86\,\mathrm{s}^{-1}$.

To conclude, the changes in the photoluminescence of porous silicon with time are caused by the expulsion of adsorbed hydrogen by oxygen. This process is initially accompanied by the formation of silicon monoxide and quenching of the photoluminescence. Subsequent additional oxidation of the silicon to give SiO_2 leads to a rise in the photoluminescence. Ultraviolet laser irradiation accelerates this process many times.

This work was financed by a Goskomvuz Grant (''Chemical Technologies'' section) and by the Scientific and Technical Program ''Conversion and High Technologies'' (No. 101-1-2).

---

1)The authors would like to thank B. M. Kostishko for assistance with the Auger spectral analysis.

[1] L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).

[2] N. H. Zoubir, M. Vergnat, T. Delatour *et al.*, Thin Solid Films **255**, 228 (1995).

[3] A. G. Cullis, L. T. Canham, and P. D. G. Calcott, J. Appl. Phys. **82**, 909 (1997).

[4] T. Hilliard, D. Andsager, L. Abu Hassan *et al.*, J. Appl. Phys. **76**, 2423 (1994).

[5] *Plasma Processing for VLSI*, edited by N. G. Einspruch and D. M. Brown, Vol. 8 of VLSI Electronics [Academic Press, New York, 1984; Mir, Moscow, 1987, 472 pp.].

[6] *Short Chemical Encyclopedia*, edited by I. L. Knunyants, G. Ya. Bakharovskiĭ, A. I. Busev *et al.*, Vol. 2 [in Russian], Sovet-skoe Éntsiklopediya, Moscow (1963), 1088 pp.

Translated by R. M. Durham

TECHNICAL PHYSICS VOLUME 44, NUMBER 6 JUNE 1999

Photosensitivity of GaAs : *N*(GaP : *N*)/GaAs(GaP) heterojunctions in linearly polarized radiation

V. I. Ivanov-Omskiĭ and Yu. V. Rud'

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

V. Yu. Rud'

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

(Submitted April 7, 1998)
Zh. Tekh. Fiz. **69**, 138–142 (June 1999)

Results are presented of investigations of the photoelectric properties of nitrided layer/GaAs (GaP) heterojunctions prepared by plasma treatment of GaAs and GaP crystals in the presence of nitrogen ions. The heterojunctions exhibited broad-band photosensitivity relative to the intensity of the natural radiation. It was established that when linearly polarized radiation is obliquely incident on the surface of nitrided layers, polarization photosensitivity occurs which is controlled by the angle of incidence Θ and increases proportionately as Θ^2. The spectral dependences of the induced photopleochroism are attributed to the antireflecting properties of the wide-gap layers. Nitrided-layer heterojunctions can be used as broad-band photoanalyzers for linearly polarized radiation. © *1999 American Institute of Physics.* [S1063-7842(99)03006-8]

INTRODUCTION

The present authors recently established that treating the surface of gallium arsenide and gallium phosphide single crystals with an rf discharge in a nitrogen atmosphere leads to the formation of wide-gap layers in the surface region which are produced by arsenic or phosphorus atoms being substituted by nitrogen.[1] This substitution leads to the appearance of short-wavelength photoluminescence at the fundamental absorption depth of the initial semiconductors together with broad-band photosensitivity relative to the intensity of the natural radiation as far as 3.8 eV. Here we report results of measurements of the photosensitivity in linearly polarized radiation using heterojunctions formed in the surface region of GaAs and GaP crystal by wide-gap GaAs:N and GaP:N layers, which for conciseness we shall subsequently call *N*-layers.

EXPERIMENTAL METHOD

The heterojunctions were fabricated using (100)- or (111)-oriented *n*-type GaAs and GaP wafers having a free electron concentration of $\cong 10^{17}\,\mathrm{cm}^{-3}$ at $T = 300\,\mathrm{K}$. The specular surface of the initial semiconductors was prepared by mechanical then chemical polishing. The wafers were inserted in an evacuated chamber with a residual pressure of around 10^{-5} Torr into which a mixture of hydrogen and nitrogen was then admitted and an rf plasma was ignited. A chemical potential gradient formed at the GaAs and GaP–plasma interface, causing the solid-phase substitution of arsenic and phosphorus atoms in the substrates by nitrogen. The substrate temperature during the plasma treatment process was controlled in the range 50–500°. These As(P)→N plasma substitution regions produced uniformly colored layers with a mirror-smooth outer surface in the surface regions of the initial crystals. The layers exhibited strong adhesion to GaAs and GaP. Uniformly colored, dark yellow layers were obtained as a result of plasma substitution on the (100) and (001) GaAs planes whereas under similar conditions for the GaP wafers, the layers were reddish lilac. These differences can be attributed to the different atomic composition and thicknesses of the layers.

The prepared heterostructures were mounted on a Fedorov table which could control the angular coordinates of the samples to within at least 30′. In these experiments the light was incident on the *N* layers and the angle of incidence varied between 0 and 90°. The current–voltage characteristics, and the angular and spectral dependences of the relative quantum efficiency η were measured using GaAs : N/n-GaAs and GaP : N/n-GaP heterojunctions with average photosensitive surface areas of 5×5 mm.

EXPERIMENTAL RESULTS AND DISCUSSION

1. Typical steady-state current–voltage characteristics for these heterojunctions are plotted in Figs. 1 and 2 (curves *1*) and some of the parameters are indicated in Table I. It can be seen that the heterojunctions exhibit well-defined rectification, and under forward biases corresponding to a negative external voltage on the substrates and exceeding the cutoff voltage $U > U_0$ the current obeys the relation

$$I = (U - U_0)/R_0, \tag{1}$$

where R_0 is the residual resistance.

The cutoff voltage, like the resistance R_0, in these heterojunctions varies fairly widely (see Table I) and reflects the influence of the conditions of formation of the *N* layers on the electrical properties of the heterojunctions. In the range of weak forward currents, the current–voltage characteristics obey the well-known exponential law for diode structures.

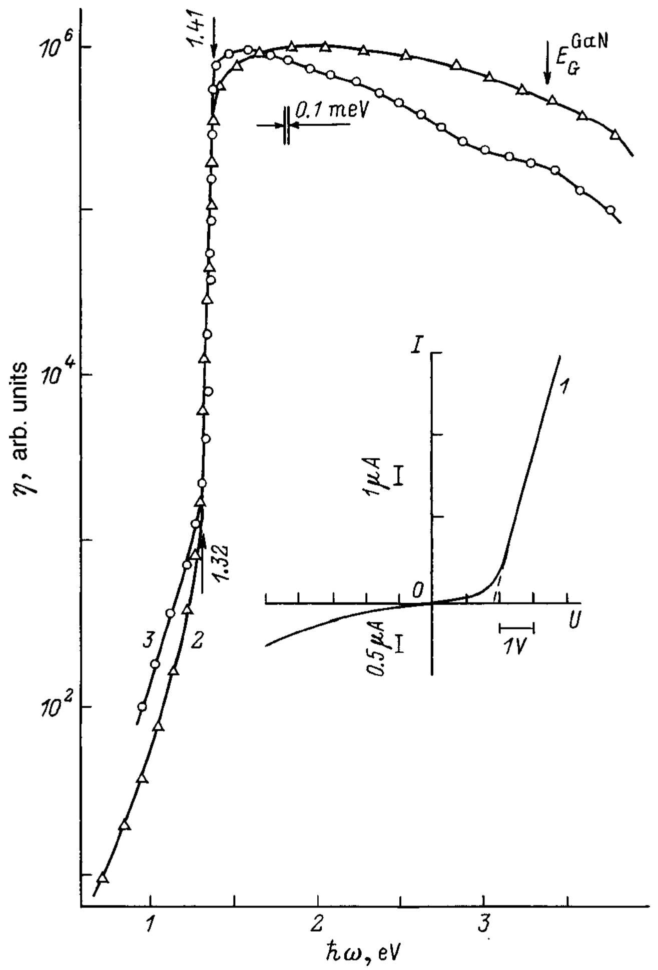


FIG. 1. Steady-state current–voltage characteristic (*1*) and spectral dependences of the relative photoconversion quantum efficiency (*2, 3*) of GaAs : N/n-GaAs heterojunctions at $T = 300$ K in natural radiation: *1, 3* — sample No. 1, *2* — sample No. 6; illuminated on side of wide-gap layer.

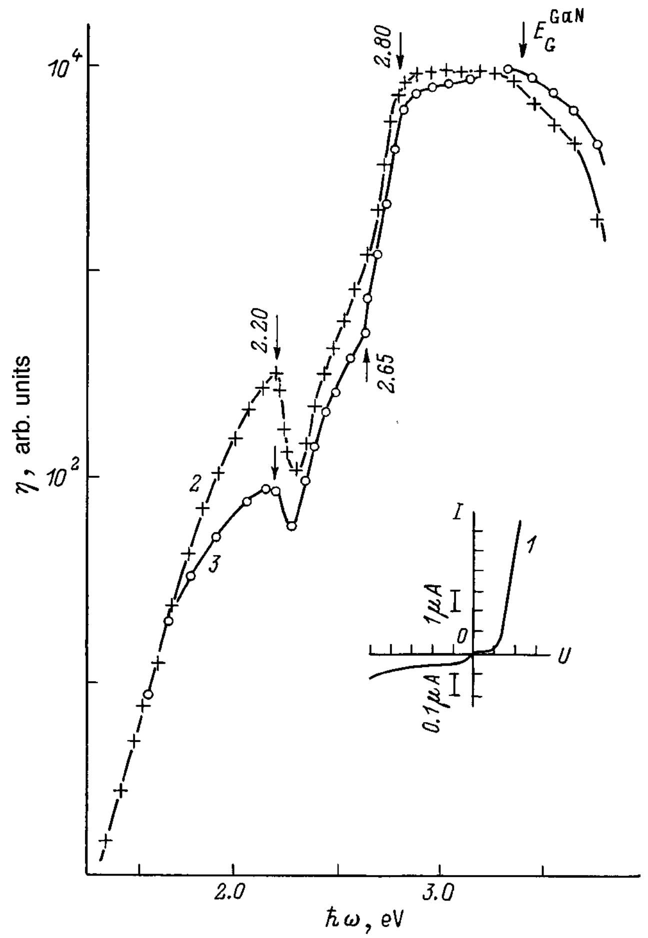


FIG. 2. Steady-state current–voltage characteristic (*1*) and spectral dependences of the relative photoconversion quantum efficiency (*2, 3*) of GaP : N/n-GaP heterojunctions at $T = 300$ K in natural radiation: *1, 2* — sample No. 4, *3* — sample No. 1; illuminated on side of wide-gap layer

The results of estimating the ideality factor β suggest that in N/n-GaAs structures the forward current is of the recombination type typical of GaAs homojunctions.[2] The high values of β in N/n-GaP heterojunctions reflect the influence of the series resistance on the current–voltage characteristic and thus cannot be used to assess the nature of the current transport. In reverse-biased heterojunctions the current in most cases obeys the power law $I \sim U^\gamma$ where the exponent γ lies in the range of 1.2–2, which indicates some imperfection at the edge of the heterojunction. The range of variation of the reverse currents I_{tr} is shown in Table I and indicates that the best of these heterojunctions are of fairly good quality.

2. When both types of heterojunctions were illuminated, the *N*-layers were always positively charged relative to the substrates. The photovoltaic effect predominates when the junctions are illuminated on the side of the *N*-layers and the range of values of the voltage sensitivity S_U is indicated in Table I. The polarity of the open-circuit photovoltage in all the heterojunctions corresponds to the direction of transmission and does not depend on the point of incidence of the optical probe on the surface nor on the energy of the incident photons.

Spectral dependences of the relative photoconversion quantum efficiency $\eta(\hbar w)$ for several heterojunctions are plotted in Figs. 1 and 2 (curves *2* and *3*). These spectra give some idea of the possibilities of this method of obtaining in-plane control of the heterojunction photosensitivity. The principal features of these spectral dependences of η are as follows. For both types of structure we observe broad-band photosensitivity typical of ideal heterojunctions. It can be seen from Figs. 1 and 2 that the maximum photosensitivity is observed in the range between the band gaps of the substrate and GaN crystals.[3] The long-wavelength component of the

TABLE I. Photoelectric properties of GaAs : N/n-GaAs and GaP : N/n-GaP heterojunctions at $T = 300$ K.

Heterojunction parameters	GaAs : N/n-GaAs	GaP : N/n-GaP
d_s, mm	0.3	0.3
d_L, μm	$\cong 0.5$	$\cong 0.8$
R_0, Ω	$4 \cdot 10^3 - 10^5$	$10^5 - 10^6$
U_0, V	$0.25 - 1.8$	1–1.2
β	1.8–2.0	7–10
I_{tr}, μA	0.2–1 ($U_{tr} = 1$ V)	0.1–1 ($U_{tr} = 5$ V)
S_U, V/W	$10^2 - 10^4$	$10 - 10^3$
s, eV^{-1}	80–115	14–18
$\delta_{1/2}$, eV	1.1–1.8	0.8–1.0

photosensitivity spectra ($\hbar w < 1.8\,\mathrm{eV}$ for N/n-GaAs and $\hbar w < 3$ eV for N/n-GaP) is caused by photoactive absorption in the GaAs and GaP substrates. The differences observed in the long-wavelength photosensitivity (Figs. 1 and 2) can be attributed to fluctuations in the perfection of the initial substrates and also to the influence of the growth conditions of the *N*-layers on their properties. The slope of the long-wavelength increase in photosensitivity $s = \delta(\ln I)/\delta(\hbar w)$ is given in Table I and is typical of binary GaAs and GaP semiconductors. The full-width of the spectral dependences $\eta(\hbar w)$ at half-height $\delta_{1/2}$ describes their broad-band property and the observed changes in this parameter reflect the influence of the layer formation conditions (see Table I). The short-wavelength decay of the photosensitivity for these heterojunctions is localized near E_G^{GaN} and this may be taken as the basis for assuming that the atomic composition in the surface region of the *N*-layers approaches GaN stoichiometry.

3. When all these heterojunctions were illuminated by linearly polarized radiation along the normal to the plane of the *N*-layers ($\Theta = 0$), the photocurrent did not depend on the polarization. This implies that under these conditions, the photoactive absorption is isotropic and the coefficient of natural photopleochroism is $P_N = 0$ over the entire range of photosensitivity of the heterojunction.[4–6] As soon as the angle of incidence of the linearly polarized radiation becomes nonzero, all the heterojunctions exhibit differences between the photocurrents i^P when the electric vector of the light wave **E** lies in the plane of incidence (PI), i.e., $\mathbf{E}||\mathrm{PI}$ and i^S ($\mathbf{E}\perp\mathrm{PI}$). Under the conditions $\Theta > 0$, over the entire range of photosensitivity the photocurrent is a periodic function of the azimuthal angle φ between **E** and the plane of incidence

$$i_\varphi = i^P \cos^2\varphi + i^S \sin^2\varphi. \tag{2}$$

For heterojunctions with different atomic compositions and differently oriented substrates the inequality $i^P > i^S$ is satisfied for $\Theta > 0$. Figures 3 and 4 show typical dependences $i^P(\Theta)$ and $i^S(\Theta)$ for these heterojunctions. It can be seen that for both types of heterojunction the photocurrent i^P initially increases with increasing angle of incidence, reaching a maximum $i^P_{\max}$, and then begins to decrease. This behavior is consistent with the results of analyzing the propagation of a light wave across the interface between two media using Fresnel relationships.[7] The increase in the photocurrent is caused by the reflection losses being eliminated, these being minimal near the Brewster angle. The experimental values of the ratio $i^P_{\max}/i^0_0$ for both types of heterojunction confirm that the increase in the photocurrent is related to the elimination of the reflection losses. At the same time, the experimentally observed characteristic $i^P(\Theta)$ taking into account the results of Refs. 4, 8, and 9 suggests that these *N*-layers on GaAs and GaP substrates are of fairly high optical quality.

A second important results of these polarization studies of the photosensitivity is that contrary to the monotonic decrease in the photocurrent i^S with increasing Θ deduced from the Fresnel formulas,[4,9] for most of our heterojunctions the photocurrent i^S behaves as i^P with increasing angle of incidence (Figs. 3 and 4, curves *2*). As a result, the curves of $i^P(\Theta)$ and $i^S(\Theta)$ converge, with this convergence being more defined for GaP:N/n-GaP structures (Fig. 4, curve *2*). This behavior was observed previously and was attributed to interference of the linearly polarized radiation in the thin layer.[10,11] Thus, there is reason to assume that the anomalous behavior of $i^S(\Theta)$ observed for these heterojunctions may be attributed to interference effects of the linearly polarized radiation in the *N*-layers. This interference also indicates that these layers are of high optical quality.

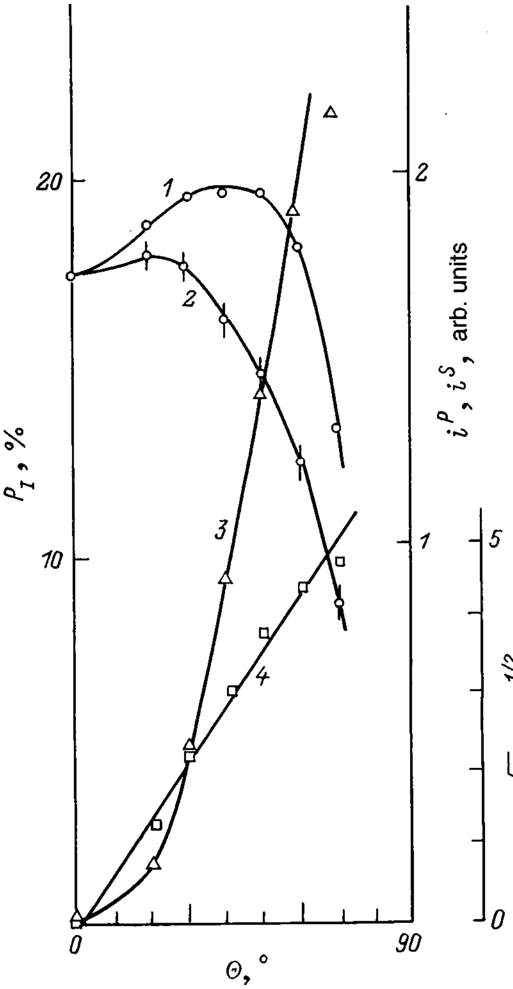


FIG. 3. Short-circuit photocurrent (*1* — i^P, *2* — i^S) and coefficient of induced photopleochroism (*3, 4*) as a function of the angle of incidence of linearly polarized radiation on the receiving plane of a GaAs:N/n-GaAs heterojunction at $T = 300$ K. Illumination is on the side of the wide-gap layer, $\lambda = 0.60\,\mu$m, sample No. 6.

The coefficient of induced photopleochroism[5,6] calculated from the curves $i^P(\Theta)$ and $i^S(\Theta)$

$$P_I = (i^P - i^S)/(i^P + i^S) \tag{3}$$

for these heterojunctions becomes nonzero for $\Theta > 0$ and increases quadratically with increasing angle of incidence $P_I \sim \Theta^2$ (Figs. 3 and 4, curves *3*). These dependences are linearized in the coordinates $P_I^{1/2}$–Θ (Fig. 3, curve *4*) in agreement with the results from Ref. 12.

Figure 5 gives spectral dependences of the coefficient of induced photopleochroism for $\Theta = \mathrm{const}$ for several hetero-

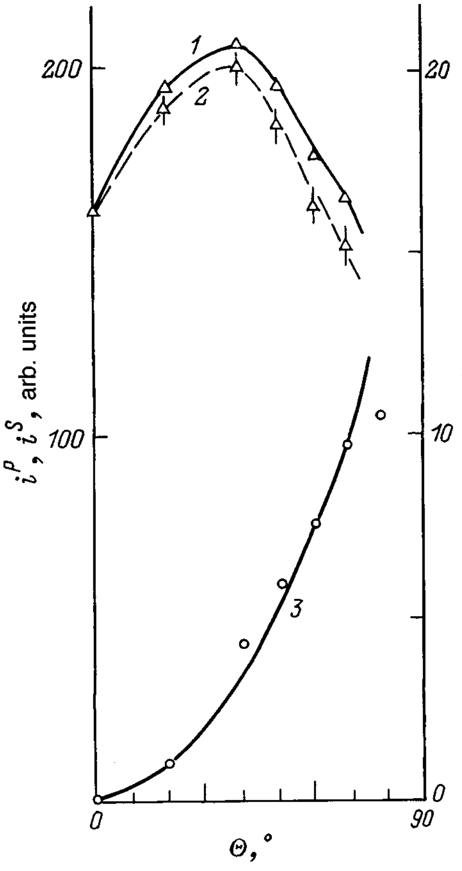


FIG. 4. Short-circuit photocurrent (*1* — i^P, *2* — i^S) and coefficient of induced photopleochroism (*3*) as a function of the angle of incidence of linearly polarized radiation on the receiving plane of a GaP : N/n-GaP heterojunction at $T = 300$ K. Illumination is on the side of the wide-gap layer, $\lambda = 0.63$ μm, sample No. 1.

junctions. It can be seen that the coefficient of induced photopleochroism for these heterojunctions has a clearly defined dependence on the incident photon energy, at variance with the analysis.[12] Medvedkin *et al.*[12] showed that the coefficient of induced photopleochroism is proportional to the refractive index n and this dependence has been used for some time to determine n from polarization measurements of the photosensitivity.[4,6] It was recently established that under the conditions of interference of linearly polarized radiation, this unique relationship $P_I \sim n$ ceases to hold and some dispersion of the induced photopleochroism occurs.[10,11] Taking account of Refs. 10 and 11, the spectral dependence of the coefficient of induced photopleochroism observed here can also be attributed to the antireflecting properties of the N-layers. The curves $P_I(\hbar w)$ suggest that a bleaching effect for which the condition is $P_I \rightarrow 0$ (Refs. 10 and 12), shows up most clearly for N/n-GaP structures (Fig. 5, curve *3*) for which the experimental values are $P_I \rightarrow 0$ over a fairly wide range between 2.5 and 3.5 eV, i.e., in the range where the photosensitivity of these heterojunctions has a maximum (Fig. 2). From the point of view of the bleaching criterion $P_I \rightarrow 0$ (Refs. 10 and 11), we can conclude that this effect is much weaker in N/n-GaAs heterojunctions, although it should be stressed that the minimum of P_I for both types of heterojunctions is localized in the same spectral range. Figure 5 shows spectra of $P_I(\hbar w)$ for two different N/n-GaAs heterojunctions (curves *1* and *2*). The similarity between the values of P_I and its spectral profile for the two different heterojunctions indicates that the quality of the N-layers has fairly good technological reproducibility.

The $P_I(\hbar w)$ curves indicate that in order to achieve high coefficients of induced photopleochroism and to use these heterojunctions as broad-band photoanalyzers for linearly polarized radiation, the technological process must ensure that this radiation does not undergo interference in the wide-gap layer of these structures.

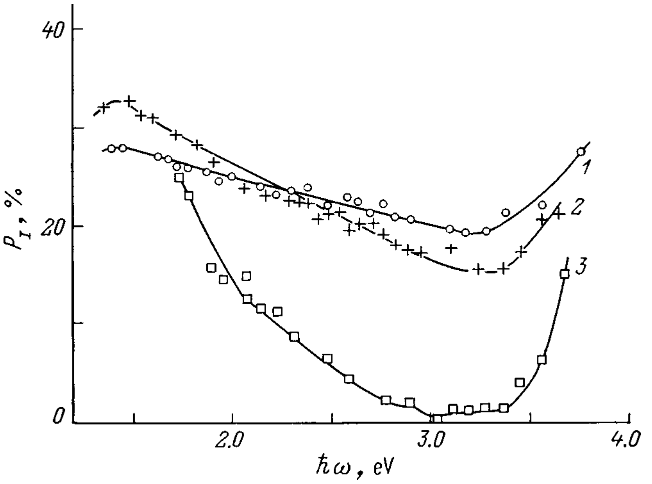


FIG. 5. Spectral dependences of the coefficient of induced photopleochroism of GaAs : N/n-GaAs heterojunctions (*1* — sample No. 1, *2* — sample No. 6) and GaP : N/n-GaP heterojunctions (*3* — sample No. 1) at $T = 300$ K and $\Theta = 70°$.

[1] V. F. Agekyan, V. I. Ivanov-Omskiĭ, Knyazevskiĭ *et al.*, Fiz. Tekh. Poluprovodn. **32**, 1203 (1998) [Semiconductors **38**, 1075 (1998)].

[2] *Gallium Arsenide*, edited by F. P. Kesamanla and D. N. Nasledova [in Russian], Nauka, Moscow (1975), 353 pp.

[3] *Handbook of Physicochemical Properties of Semiconductor Materials* [in Russian], Nauka, Moscow (1978), 339 pp.

[4] Yu. V. Rud', Izv. Vyssh. Uchebn. Zaved. Fiz. **29**(8), 68 (1986).

[5] V. Yu. Rud', Author's Abstract of Candidate's Dissertation [in Russian], St. Petersburg (1995), 17 pp.

[6] S. G. Konnikov, V. Yu. Rud', Yu. V. Rud' *et al.*, Jpn. J. Appl. Phys. **32**, 515 (1993).

[7] R. M. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (North-Holland, Amsterdam, 1977), 584 pp.

[8] Yu. V. Zhilyaev, N. Nazarov, V. Yu. Rud' *et al.*, Fiz. Tekh. Poluprovodn. **27**, 1610 (1993) [Semiconductors **27**, 890 (1993)].

[9] V. M. Botnaryuk, L. V. Gorchak, V. N. Pleshka *et al.*, Fiz. Tekh. Poluprovodn. **31**, 305 (1997) [Semiconductors **31**, 194 (1997)].

[10] V. Yu. Rud' and Yu. V. Rud', Fiz. Tekh. Poluprovodn. **31**, 309 (1997) [Semiconductors **31**, 197 (1997)].

[11] T. Walter, V. Yu. Rud', Yu. V. Rud', and G. V. Schock, Fiz. Tekh. Poluprovodn. **31**, 806 (1997) [Semiconductors **31**, 681 (1997)].

[12] G. A. Medvedkin and Yu. V. Rud', Phys. Status Solidi A **67**, 333 (1981).

Translated by R. M. Durham