Solution of the Dirichlet problem for the Laplace equation for a multiply connected region with point symmetry

I. F. Spivak-Lavrov

''Dunie'' Aktyubinsk Nongovernmental Higher Educational Institute, 463000 Aktyubinsk, Kazakhstan
(Submitted June 3, 1997; resubmitted February 9, 1998)
Zh. Tekh. Fiz. **69**, 1–9 (March 1999)

A method is proposed which uses an expansion of the potential in irreducible representations of the symmetry group of the field-defining elements of a system. A boundary-value problem is solved for multipole systems with planar plate electrodes for the C_{nv} symmetry group. A quadrature expression is obtained for the field potential of these systems. Constraints imposed on the electrode potentials, under which such a solution is possible, are determined. Results of calculations of the potential distribution are presented for various specific systems.
 [S1063-7842(99)00103-8]

1. INTRODUCTION

A method which can be used to find a solution of the Dirichlet problem for the Laplace equation for two-dimensional multiply connected regions having two mutually perpendicular symmetry axes was proposed in Ref. 1. The solution of the Dirichlet problem for this type of multiply connected region reduces to solving several Dirichlet–Neumann boundary-value problems for each of the four potential components, but in a singly connected region bounded by the symmetry axes and by the portions of the electrodes of the initial system lying directly in the respective quadrant. In the particular case of two-dimensional or conical fields, this method can yield closed analytic formulas for the electrostatic potential of various systems with plate electrodes with allowance for the sizes of the gaps between them.

It was shown in Ref. 2 that the method proposed in Ref. 1 can be generalized to multiply connected regions possessing arbitrary geometric symmetry. In the present paper, a detailed analysis is made of the case where the symmetry of the field-defining elements of an ion-optical system is described by the C_{nv} group (see Ref. 3, for instance). This is the symmetry group of a body with n symmetry planes passing through an nth-order symmetry axis. This symmetry may be exhibited by the electrode system of a multipole deflector, where $n=2,3,4,\ldots$ correspond to quadrupole, sextupole, octupole, ... systems. Note that the symmetry of the systems analyzed in Ref. 1 corresponds to the C_{2v} group.

2. EXPANSION OF THE POTENTIAL

If the electrodes of an ion-optical system have a specific geometric symmetry, the potential φ of the field of this system can be given by[2,3]

$$\varphi=\sum_{\alpha}\ \varphi_{\alpha}, \tag{1}$$

where φ_{α} are the potential components transforming according to different irreducible representations of the symmetry group of the field-defining surfaces of the system:

$$\varphi_{\alpha}=g^{-1}f_{\alpha}\sum_{G}\ \chi_{\alpha}(G)^{*}\hat{G}\varphi. \tag{2}$$

Here g is the order of the group, f_{α} is the dimension of the irreducible representation, $\chi_{\alpha}(G)$ is the character of the irreducible representation assigned to the G element of the group. The symmetry properties can be used to find values of the potentials at the electrodes for which the boundary conditions for the various components φ_{α} can be determined in a singly connected region formed by the field-defining surfaces and symmetry surfaces, which is the goal of the method.

We shall henceforth analyze a two-dimensional problem, assuming that the electrode system belongs to the C_{nv} symmetry group. Let us assume that the potential $\varphi\equiv\varphi(x,y)$ depends only on the two Cartesian coordinates x and y. Using tables of the characters $\chi_{\alpha}(G)$ of the irreducible representations of the C_{nv} group,[3] we can use formula (2) to write expressions for the components $\varphi_{\alpha}(x,y)$. For the C_{2v} symmetry group we obtain the expressions given in Refs. 1 and 2

$$\varphi_1(x,y)=\frac{1}{4}[\varphi(x,y)+\varphi(-x,-y)+\varphi(x,-y)+\varphi(-x,y)], \tag{3}$$

$$\varphi_2(x,y)=\frac{1}{4}[\varphi(x,y)+\varphi(-x,-y)-\varphi(x,-y)-\varphi(-x,y)], \tag{4}$$

$$\varphi_3(x,y)=\frac{1}{4}[\varphi(x,y)-\varphi(-x,-y)+\varphi(x,-y)-\varphi(-x,y)], \tag{5}$$

$$\varphi_4(x,y)=\frac{1}{4}[\varphi(x,y)-\varphi(-x,-y)-\varphi(x,-y)+\varphi(-x,y)]. \tag{6}$$

For even $n=2p$ the number of irreducible representations of the C_{nv} group is $p+3$, whereas for odd $n=2p+1$ it is $p+2$. Using the familiar formula for the dimensions of irreducible representations

$$\sum_{k=1}^{r} f_k^2 = g, \tag{7}$$

where r is the number of different irreducible representations of the group, we can show for the C_{nv} group ($g=2n$) that the number of one-dimensional irreducible representations is 4 for even n and 2 for odd n. The remaining irreducible representations are two-dimensional, the number of these being $p-1$ for even $n=2p$ and p for odd $n=2p+1$.

The components φ_α can be conveniently defined in the polar coordinates ρ, ψ, matched with the Cartesian coordinates chosen, since in this case, the four components corresponding to the one-dimensional representations of the C_{nv} group can be given for any even n in the form:

$$\varphi_1(\rho,\psi)=\frac{1}{2n}\sum_{j=0}^{n-1}\left[\varphi\left(\rho,\frac{2\pi}{n}j+\psi\right)+\varphi\left(\rho,\frac{2\pi}{n}j-\psi\right)\right], \tag{8}$$

$$\varphi_2(\rho,\psi)=\frac{1}{2n}\sum_{j=0}^{n-1}\left[\varphi\left(\rho,\frac{2\pi}{n}j+\psi\right)-\varphi\left(\rho,\frac{2\pi}{n}j-\psi\right)\right], \tag{9}$$

$$\varphi_3(\rho,\psi)=\frac{1}{2n}\sum_{j=0}^{n-1}(-1)^j\left[\varphi\left(\rho,\frac{2\pi}{n}j+\psi\right)+\varphi\left(\rho,\frac{2\pi}{n}j-\psi\right)\right], \tag{10}$$

$$\varphi_4(\rho,\psi)=\frac{1}{2n}\sum_{j=0}^{n-1}(-1)^j\times\left[\varphi\left(\rho,\frac{2\pi}{n}j+\psi\right)-\varphi\left(\rho,\frac{2\pi}{n}j-\psi\right)\right]. \tag{11}$$

For the case $n=2$ these expressions, which are equivalent to (3)–(6), exhaust all the irreducible representations of the group. For any odd n the two components φ_1 and φ_2 corresponding to the one-dimensional representations are also given by formulas (8) and (9).

Expressions for the potential components corresponding to the two-dimensional representations of the C_{nv} group cannot be given in a general form for any n. However, the sum of all these components for even $n=2p$ is

$$\varphi-\sum_{\alpha=1}^{4}\varphi_\alpha=\frac{1}{p}\left[(p-1)\varphi(\rho,\psi)-\sum_{j=1}^{p-1}\varphi\left(\rho,\frac{2\pi}{p}j+\psi\right)\right], \tag{12}$$

and for odd $n=2p+1$

$$\varphi-\sum_{\alpha=1}^{2}\varphi_\alpha=\frac{1}{2p+1}\left[2p\varphi(\rho,\psi)-\sum_{j=1}^{2p}\varphi\left(\rho,\frac{2\pi}{2p+1}j+\psi\right)\right]. \tag{13}$$

Using expressions (12) and (13) and the orthogonality conditions for the characters of the irreducible representations, we can determine the characters $\chi_\alpha(G)$ for each specific n and then, using formula (2), the components corresponding to the two-dimensional representations. For example, for the C_{4v} group ($n=4$) we obtain

$$\varphi_5(\rho,\psi)=\frac{1}{2}[\varphi(\rho,\psi)-\varphi(\rho,\pi+\psi)], \tag{14}$$

for the C_{6v} group ($n=6$)

$$\varphi_5(\rho,\psi)=\frac{1}{6}\left[2\varphi(\rho,\psi)+2\varphi(\rho,\pi+\psi)-\varphi\left(\rho,\frac{2\pi}{3}+\psi\right)-\varphi\left(\rho,\frac{4\pi}{3}+\psi\right)-\varphi\left(\rho,\frac{\pi}{3}+\psi\right)-\varphi\left(\rho,\frac{5\pi}{3}+\psi\right)\right], \tag{15}$$

$$\varphi_6(\rho,\psi)=\frac{1}{6}\left[2\varphi(\rho,\psi)-2\varphi(\rho,\pi+\psi)-\varphi\left(\rho,\frac{2\pi}{3}+\psi\right)-\varphi\left(\rho,\frac{4\pi}{3}+\psi\right)+\varphi\left(\rho,\frac{\pi}{3}+\psi\right)+\varphi\left(\rho,\frac{5\pi}{3}+\psi\right)\right], \tag{16}$$

and for the C_{3v} group ($n=3$)

$$\varphi_3(\rho,\psi)=\frac{1}{3}\left[2\varphi(\rho,\psi)-\varphi\left(\rho,\frac{2\pi}{3}+\psi\right)-\varphi\left(\rho,\frac{4\pi}{3}+\psi\right)\right]. \tag{17}$$

The components $\varphi_\alpha(\rho,\psi)$ ($\alpha=1,2,3,4$) corresponding to the one-dimensional representations of the group play a special role in the solution of the boundary-value problem. It is easy to show that on the symmetry planes where $\psi=\pi(j-1)/n$ ($j=1,2,3,\ldots,2n$) the components given by the expressions (8)–(11) satisfy the following boundary conditions:

$$\varphi_{1\psi}\left(\rho,\frac{\pi}{n}(j-1)\right)=0 \quad (j=1,2,3,\ldots,2n), \tag{18}$$

$$\varphi_2\left(\rho,\frac{\pi}{n}(j-1)\right)=0 \quad (j=1,2,3,\ldots,2n), \tag{19}$$

$$\varphi_{3\psi}\left(\rho,\frac{\pi}{n}(j-1)\right)=0 \quad (j=1,3,\ldots,2n-1),$$

$$\varphi_3\left(\rho,\frac{\pi}{n}(j-1)\right)=0 \quad (j=2,4,\ldots,2n), \tag{20}$$

$$\varphi_{4\psi}\left(\rho,\frac{\pi}{n}(j-1)\right)=0 \quad (j=2,4,\ldots,2n),$$

$$\varphi_4\left(\rho,\frac{\pi}{n}(j-1)\right)=0 \quad (j=1,3,\ldots,2n-1). \tag{21}$$

In these formulas the subscript ψ denotes the partial derivative with respect to the coordinate ψ. Thus, on all the symmetry planes either the potential components corresponding to the one-dimensional representations of the sym-

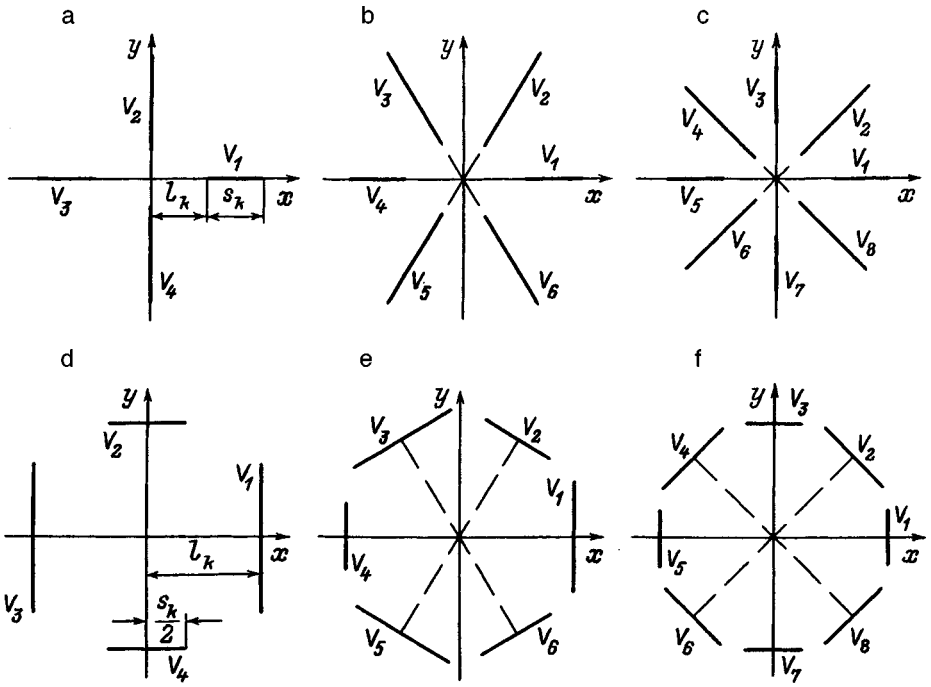

FIG. 1. Multipole systems with plate electrodes having C_{nv} symmetry: a, d — quadrupole ($n=2$), b, e — sextupole ($n=3$), and c, f — octupole ($n=4$).

metry group or their normal derivative vanish, making it possible to demarcate $2n$ singly connected regions with assigned Dirichlet–Neumann boundary conditions on them for φ_1, φ_3, and φ_4 and with simply Dirichlet boundary conditions for φ_2.

In order to calculate the potential distribution $\varphi(x,y)$, the boundary conditions (18)–(21) should be supplemented by the boundary conditions at the system electrodes. If the potentials V_k ($k=1,2,\ldots,N$) are known at N electrodes, formulas (8)–(11) can be used to find the potentials $V_{\alpha k}$ at the electrodes for the components φ_α ($\alpha=1,2,3,4$). The elements $V_{\alpha k}$ of the potential matrix V are such that

$$\sum_{\alpha=1}^{4} V_{\alpha k}=V_k. \tag{22}$$

The boundary conditions for the potential components corresponding to the two-dimensional representations [see formulas (14)–(17)] cannot be defined on any of the symmetry planes, but the problem can be solved for those values of the potentials V_k for which these components vanish identically.

3. CONFORMAL MAPPING AND BOUNDARY CONDITIONS

The proposed method can be illustrated by considering multipolar ion-optical systems with plate electrodes exhibiting C_{nv} ($n=2,3,4$) symmetry, as shown in Fig. 1. The electrode potentials are denoted by V_k ($k=1,2,\ldots,8$). The parameter l_k assigns the distance between the symmetry axis and the electrode, and s_k is the size of an electrode with the potential V_k. The dashed lines define the positions of the symmetry planes when the latter do not coincide with the coordinate planes.

For each of the systems shown in Fig. 1 we can demarcate singly connected regions bounded by the rays $\psi=\pi(j-1)/n$ and $\psi=\pi j/n$ ($j=1,2,3,\ldots,2n$) and by the parts of the electrodes located in the relevant sector.

Each of the singly connected regions demarcated in the complex plane $z=x+iy=\rho\exp(i\psi)$ can be mapped onto the upper half-plane $w=u+iv$ by the analytic functions $w(z^{(j)})$ ($j=1,2,3,\ldots,2n$), where the superscript j is introduced for the one-sheeted mapping and indicates that the values of z belong to the corresponding singly connected region with the number j. The potential distribution in the upper half-plane can also be given by the $2n$ functions $\varphi^{(j)}(u,v)$, which describe the potential distribution in the corresponding singly connected regions.

For the systems shown in Figs. 1a–1c, each sector with the index j, bounded by the rays $\psi=\pi(j-1)/n$ and $\psi=\pi j/n$ ($j=1,2,3,\ldots,2n$) is mapped onto the upper half-plane of the w plane by the analytic function

$$w(z^{(j)})=\rho^n e^{i(n\psi-\pi(j-1))}. \tag{23}$$

For the systems shown in Fig. 1d–1f, the singly connected regions demarcated are octagons, one of which corresponds to $j=1$ and is shown in Fig. 2a. The mapping of the upper half-plane of the w plane onto these octagons in the z plane is given by the conformal transformation[2]

$$z^{(j)}(w)=C\exp\left(i\pi\frac{j-1}{n}\right)\int_0^w f(w)\,dw$$

$$(j=1,2,\ldots,2n), \tag{24}$$

where C is a real constant, and the integrand is

$$f(w)=\frac{w^{\frac{1-n}{n}}(w-a_e)(w+a_4)}{\sqrt{(w-a_1)(w-a_2)(w+a_3)(w+a_5)}}.$$

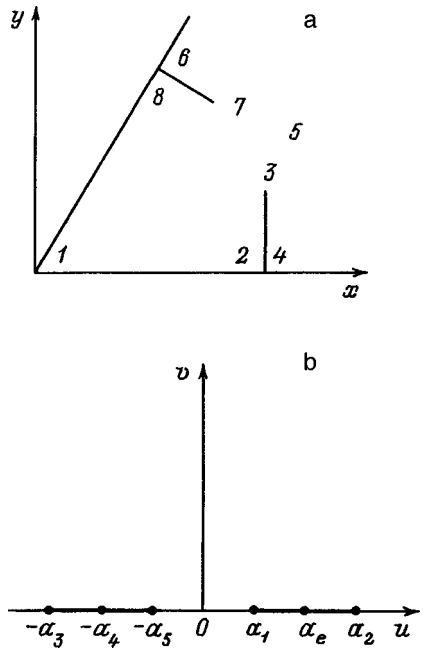

FIG. 2. Demarcation and mapping of a singly connected region: a — boundary-value problem in the $z=x+iy$ plane; *1, 2, . . . , 8* — vertices of an octagon; b — boundary-value problem in the $w=u+iv$ plane.

The vertices 1,2, . . . ,8 of the octagon in the mapping (24) correspond to the points 0, a_1, a_e, a_2, ∞, $-a_3$, $-a_4$, and $-a_5$ on the real u axis, respectively, and the electrode segments correspond to the segments $[a_1,a_2]$ and $[-a_3, -a_5]$. The mapped region in the w plane is shown in Fig. 2b.

By virtue of the boundary conditions (18)–(21) for the components $\varphi_\alpha^{(j)}$ ($\alpha=1,2,3,4$) on the segments $u<-a_3$, $-a_5<u<a_1$ and $u>a_2$ of the real u axis, either the components $\varphi_\alpha^{(j)}(u,0)$ or their normal derivative $\varphi_{\alpha v}^{(j)}(u,0)$ are zero. Here the subscript v denotes the partial derivative with respect to the coordinate v. On the other segments of the real axis the values of the components $\varphi_\alpha^{(j)}(u,0)$ are determined by the potential matrix V with the elements $V_{\alpha k}$, where

$$\varphi_\alpha^{(j)}(u,0)=\begin{cases} V_{\alpha j}, & \text{if} \quad a_1\leqslant u\leqslant a_2, \\ V_{\alpha j+1}, & \text{if} \quad -a_3\leqslant u\leqslant -a_5. \end{cases} \tag{25}$$

This last formula is valid for all the multipole systems being studied if allowance is made for the cyclicity condition $V_{\alpha j+2n}\equiv V_{\alpha j}$.

For a quadrupole system with C_{2v} symmetry and arbitrary values of the electrode potentials V_k ($k=1,2,3,4$) the potential matrix has the form

$$V=\begin{pmatrix} \frac{V_1+V_3}{2} & \frac{V_2+V_4}{2} & \frac{V_1+V_3}{2} & \frac{V_2+V_4}{2} \\ 0 & 0 & 0 & 0 \\ \frac{V_1-V_3}{2} & 0 & \frac{V_3-V_1}{2} & 0 \\ 0 & \frac{V_2-V_4}{2} & 0 & \frac{V_4-V_2}{2} \end{pmatrix}, \tag{26}$$

from which it follows, with allowance for Eq. (19), that $\varphi_2\equiv 0$.

In cases of C_{3v} symmetry, as in the previous case, $\varphi_2\equiv 0$, and the potential component φ_3 and its normal derivative cannot be determined at any boundaries of the singly connected regions for arbitrary values of the electrode potentials. The problem can be solved when the electrode potentials are such that formula (17) gives $\varphi_3\equiv 0$. This constraint is satisfied if

$$2V_1-V_3-V_5=0, \quad 2V_2-V_4-V_6=0,$$

$$2V_3-V_5-V_1=0, \quad 2V_4-V_6-V_2=0,$$

$$2V_5-V_1-V_3=0, \quad 2V_6-V_2-V_4=0.$$

Hence it follows that $V_1=V_3=V_5$, $V_2=V_4=V_6$, and only the first line of the potential matrix is nonzero and that $V_{11}=V_{13}=V_{15}=V_1$ and $V_{12}=V_{14}=V_{16}=V_2$. In this case, we can also use the equivalent potential matrix for which

$$V_{11}=V_{13}=V_{15}=U, \quad V_{12}=V_{14}=V_{16}=-U. \tag{27}$$

Here $U=(V_1-V_2)/2$, the potential at infinity obtained using the conditions (27) is zero, and the initial potential distribution can be obtained from it by adding the constant $(V_1+V_2)/2$. Thus, for sextupole systems a solution can be obtained only when the potential distribution also has C_{3v} symmetry.

As in the previous systems, for C_{4v} symmetry $\varphi_2\equiv 0$. However, $\varphi_5\equiv 0$ is also obtained if the electrode potentials are such that $V_5=V_1$, $V_6=V_2$, $V_7=V_3$, and $V_8=V_4$. Under these conditions, the potential matrix for an octupole system is in fact similar to the potential matrix (26) for a quadrupole system and the fifth column of this matrix is the same as the first column of the matrix (26), the sixth column is the same as the second, the seventh is the same as the third, and the eighth is the same as the fourth.

Thus, the field of multipole systems can be calculated for arbitrary values of the electrode potentials in the case of quadrupole systems only when $n=2$, which corresponds to the systems considered in Ref. 1. In cases where n is even, a solution can only be obtained for relations between the electrode potentials which actually reduce the problem to calculating the field of a quadrupole system. For example, for the systems shown in Figs. 1a and 1c the potential distribution in the multipole system can be obtained from that in the corresponding quadrupole system by using the simple mapping z^p, where $p=n/2$, and the symmetry principle. Note also that with this transformation, as would be expected, the expressions (8)–(11) become the corresponding formulas for the components of a quadrupole system with $n=2$, and the expression (12) vanishes. For odd n, a solution can be obtained only for boundary conditions of the type (27) under which the potential distribution $\varphi(x,y)$ also has C_{nv} symmetry.

4. EXPRESSIONS FOR THE POTENTIALS

Applying the Keldysh–Sedov method (see Ref. 4, for example) as in Refs. 1 and 2, we introduce the complex components of the potential

$$\Omega_\alpha^{(j)}(w)=\phi_\alpha^{(j)}(u,v)+i\varphi_\alpha^{(j)}(u,v) \quad (\alpha=1,3,4), \tag{28}$$

where the derivative

$$\frac{d\Omega_{\alpha}^{(j)}}{dw}=\frac{\partial\varphi_{\alpha}^{(j)}}{\partial v}+i\frac{\partial\varphi_{\alpha}^{(j)}}{\partial u}. \tag{29}$$

For a quadrupole system, taking into account the boundary conditions (25), we can write the following expressions for the derivatives of the complex components of the potential

$$\frac{d\Omega_{1}^{(j)}}{dw}=\frac{\gamma_{j1}}{\sqrt{(w-a_1)(a_2-w)(w+a_3)(w+a_5)}}=\gamma_{j1}f_1(w), \tag{30}$$

$$\frac{d\Omega_{3}^{(j)}}{dw}=\frac{\gamma_{j3}}{\sqrt{w(w-a_1)(a_2-w)}}=\gamma_{j3}f_3(w), \tag{31}$$

$$\frac{d\Omega_{4}^{(j)}}{dw}=\frac{\gamma_{j4}}{\sqrt{w(w+a_3)(w+a_5)}}=\gamma_{j4}f_4(w), \tag{32}$$

where γ_{j1}, γ_{j3}, and γ_{j4} are real constants.

These constants are obtained from the correspondence to the boundary conditions (25) and we can write the following expressions for them:

$$\gamma_{j\alpha}=\frac{V_{\alpha j+1}-V_{\alpha j}}{J_{\alpha}} \quad (\alpha=1,3,4). \tag{33}$$

Here J_{α} denote the following integrals:

$$J_{\alpha}=\mathrm{Im}\int_{a_e}^{-a_4}f_{\alpha}(w)\,dw, \tag{34}$$

whose value is determined by the geometry of the system. Formulas (30)–(34) are valid for an octupole provided $V_5=V_1$, $V_6=V_2$, $V_7=V_3$, and $V_8=V_4$ and for a sextupole system provided $V_1=V_3=V_5$, $V_2=V_4=V_6$, and the only nonzero potential component is $\varphi_1^{(j)}$, which can be determined using the boundary conditions (27).

Substituting the integrated expressions (30)–(32) into formula (1), we finally obtain the following expression for the potential distribution $\varphi^{(j)}(u,v)$ in the jth sector:

$$\varphi^{(j)}(u,v)=\mathrm{Im}\int_{a_e}^{w}[\gamma_{j1}f_1(w)+\gamma_{j3}f_3(w)+\gamma_{j4}f_4(w)]\,dw+V_j, \tag{35}$$

which, combined with the conformal transformation (24) or (23), determines the potential sought $\varphi(x,y)$ in the respective sector. Note that in those cases where the inverse transformation $w^{(j)}(z)$ is not obtained explicitly, as in the case of the transformation (24), for example, it is better to use the distribution $\varphi^{(j)}(u,v)$ rather than the potential distribution $\varphi(x,y)$ to calculate the trajectories of charged particles.[5] Moreover, the same potential distribution in the w plane can describe the properties of different multipole systems, for which the form of the conformal transformation $z(w)$ differs.

5. RESULTS OF NUMERICAL CALCULATIONS

We shall henceforth confine our analysis to the case of greatest practical interest, in which the dimensions s_k and l_k of all the electrodes are the same and are equal to s and l, respectively. On account of the symmetry, we have $a_3=a_2$, $a_4=a_e$, and $a_5=a_1$, and the functions in the integrands in Eqs. (24) and (35) have the form

$$f(w)=\frac{w^{\frac{1-n}{n}}(w^2-a_e^2)}{\sqrt{(w^2-a_1^2)(w^2-a_2^2)}}, \tag{36}$$

$$f_1(w)=\frac{1}{\sqrt{(w^2-a_1^2)(a_2^2-w^2)}}, \tag{37}$$

$$f_3(w)=\frac{1}{\sqrt{w(w-a_1)(a_2-w)}}, \tag{38}$$

$$f_4(w)=f_3(-w). \tag{39}$$

It is easy to show that $J_4=J_3$ for these systems.

When the dimensions and configuration of all the electrodes are the same in the systems shown in Figs. 1a–1c, we have

$$a_1=l^n, \quad a_2=(l+s)^n. \tag{40}$$

For the systems shown in Figs. 1d–1f with plate electrodes of the same size we calculated the parameters of the conformal transformation (24) a_1 and a_2 and the integrals J_1 and J_3 as a function of the system geometry. We assumed that $a_4=a_e=1$. For each value of a_1 we selected the value of a_2 which satisfies the equality

$$\int_{a_1}^{a_e}|f(u)|\,du=\int_{a_e}^{a_2}|f(u)|\,du=\frac{s}{2C}, \tag{41}$$

where the function $f(u)$ is given by (36). For these values of a_1 and a_2 we calculated the integral

$$\int_{0}^{a_1}|f(u)|\,du=\frac{l}{C} \tag{42}$$

and obtained the parameter $a=s/2l$, which determines the geometry of the respective system. The results of the calculations are given in Table I for the three values $n=2$, 3, and 4 and the following values of a: 0.1,0.2, . . . ,0.9 for $n=2$, 0.1,0.2, . . . ,0.5 for $n=3$, and 0.1, 0.2, 0.3, 0.4 for $n=4$.

We calculated the potential distribution $\varphi(x,y)$ for the quadrupole system shown in Fig. 1a with $V_1=0$, $V_2=1.0$, $V_3=2.0$, $V_4=3.0$; $l=0.6$, $s=0.6$. We calculated the integrals J_{α} using formula (34) for $a_e=a_4=1$ with allowance for (37), (38), and (40), and obtained the following values: $J_1=2.21700$, $J_3=J_4=2.80958$. The coefficients $\gamma_{j\alpha}$ were obtained from formula (33) using the potential matrix (26). These values of the coefficients form the following matrix:

$$\gamma=\begin{pmatrix} 0.451059 & 0 & 0.355925 & -0.355925 \\ -0.451059 & 0 & 0.355925 & 0.355925 \\ 0.451059 & 0 & -0.355925 & 0.355925 \\ -0.451059 & 0 & -0.355925 & -0.355925 \end{pmatrix}. \tag{43}$$

The values of the coefficients $\gamma_{j\alpha}$ were used in formula (35) to calculate the potential distributions $\varphi^{(j)}(u,v)$ ($j=1,2,3,4$) in the upper half-plane of the w plane. Using the mapping (23), from each of the four branches of the potential

TABLE I. Values of the conformal mapping parameters a_1, a_2, C, and the integrals J_1, J_3 as functions of the system geometry.

$a=s/(2l)$		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$n=2$	a_1	0.80982	0.63907	0.48800	0.35713	0.24681	0.15715	0.08795	0.03892	0.00970
	a_2	1.21036	1.44382	1.70722	2.01403	2.38961	2.88314	3.60191	4.83540	7.81772
	C/l	0.49874	0.49482	0.48781	0.47700	0.46135	0.43931	0.40842	0.36402	0.29371
	J_1	2.99432	2.29609	1.87960	1.57233	1.31822	1.09045	0.87233	0.64972	0.40186
	J_3	3.69425	3.00984	2.61089	2.323301	2.08806	1.87621	1.66557	1.43156	1.12394
$n=3$	a_1	0.72691	0.49877	0.30886	0.15610	0.04492				
	a_2	1.33467	1.76940	2.40018	3.52374	6.73568				
	C/l	0.33096	0.32295	0.30796	0.28228	0.23401				
	J_1	2.56531	1.81245	1.31437	0.89199	0.46642				
	J_3	3.27862	2.56121	2.09824	1.69260	1.21251				
$n=4$	a_1	0.65108	0.37690	0.16092	0.01280					
	a_2	1.47477	2.22170	3.89622	18.2539					
	C/l	0.24763	0.23713	0.21468	0.14969					
	J_1	2.24722	1.42075	0.80666	0.17211					
	J_3	2.97667	2.20372	1.60841	0.73544					

$\varphi^{(j)}(u,v)$ we can find the corresponding potential distribution $\varphi(x,y)$ in one of the quadrants of the z plane, whose number is specified by the value of j.

Using formulas (23) and (24), from the potential distributions $\varphi^{(j)}(u,v)$ obtained for $a_1=a_5=0.36$ and $a_2=a_3=1.44$ we find the field distribution in corresponding multipole systems of both the star (Figs. 1a and 1c) and polygon (Figs. 1d and f) types. Using the mapping inverse to (23) with $n=2$, we obtain the potential distribution in the initial quadrupole system (Fig. 1a) with $V_1=0$, $V_2=1.0$, $V_3=2.0$,

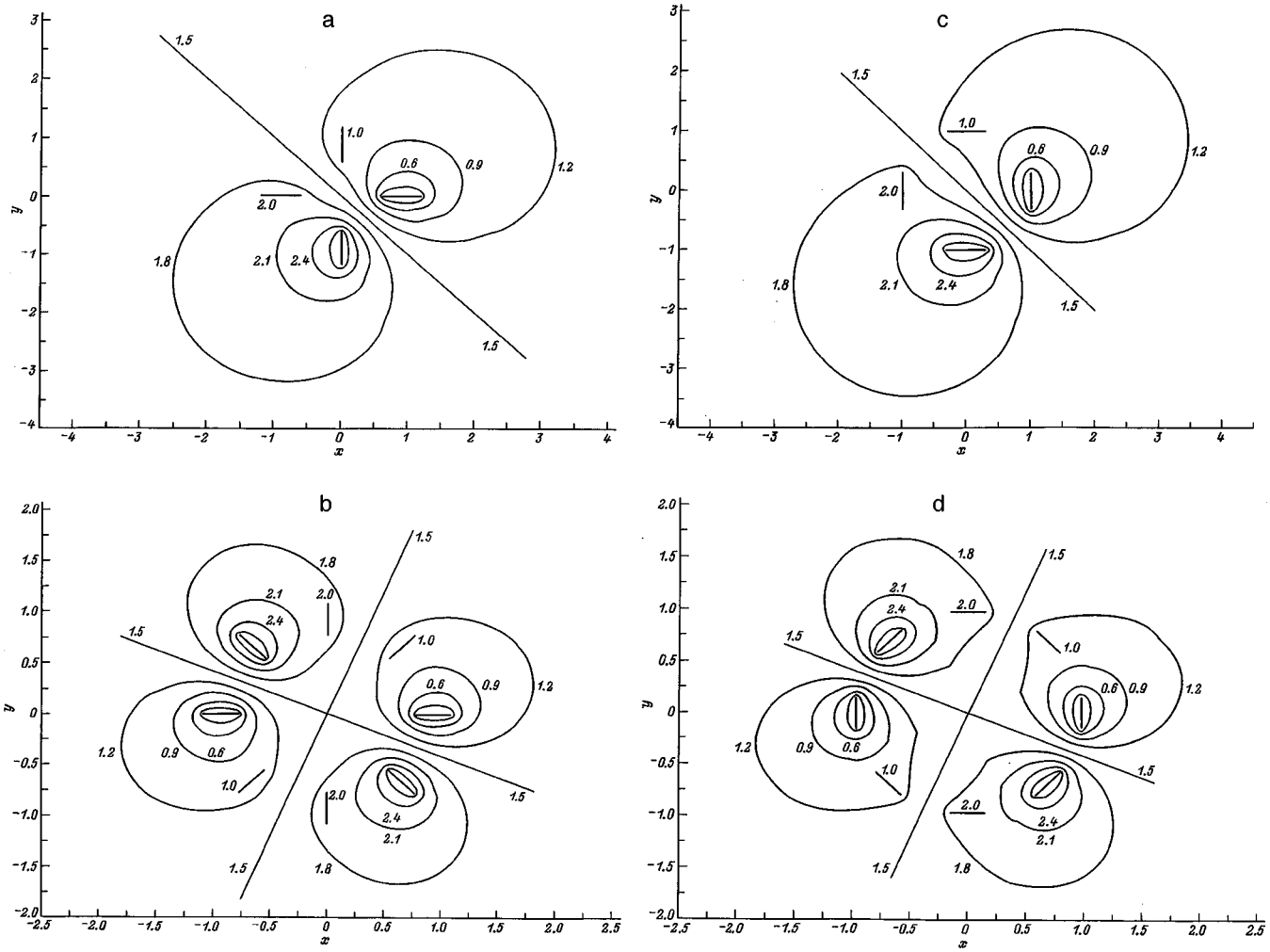

FIG. 3. Potential distribution in multipole systems for $V_1=0$, $V_2=1.0$, $V_3=2.0$, and $V_4=3.0$: a, b — star systems, c, d — polygon systems.

$V_4=3.0$, $l=0.6$, and $s=0.6$, and for the case of $n=4$ we obtain the distribution in an octupole system (Fig. 1c) with $V_1=V_5=0$, $V_2=V_6=1.0$, $V_3=V_7=2.0$, $V_4=V_8=3.0$, $l=0.774597$, and $s=0.320848$. Using the mapping (24) with integrand defined by (36), for $n=2$, $a_1=0.36$, $a_2=1.44$, $a_e=a_4=0.646295$, and $C=0.541029$ we find the potential distribution in a polygon-type quadrupole system (Fig. 1d) with $V_1=0$, $V_2=1.0$, $V_3=2.0$, $V_4=3.0$, and $l=1.0$, $s=0.625980$, and for $n=4$, $a_1=0.36$, $a_2=1.44$, $a_e=a_4=0.579398$, and $C=0.259266$ we obtain the distribution in an octupole system (Fig. 1f) with $V_1=V_5=0$, $V_2=V_6=1.0$, $V_3=V_7=2.0$, $V_4=V_8=3.0l=1.0$, and $s=0.299851$.

All the calculations were made with an accuracy which ensures that the potential has at least five significant digits. The equipotentials of these distributions corresponding to 0.3,0.6, . . . ,2.7 are plotted in Figs. 3a and b for star systems and in Figs. 3c and d for polygon systems.

6. CONCLUSIONS

This method for calculating the potential of the field of multipole systems can be used to obtain exact quadrature expressions for solving the Laplace equation for symmetric multiply connected boundary regions. Group-theoretical methods can be applied to find the relations between the electrode potentials for which such a solution is possible. The proposed method is particularly valuable when the electrodes have points, sharp unrounded edges, and other features which make it difficult to obtain solutions by known numerical methods. This method may also prove useful for determining the accuracy of calculations made by other numerical methods such as the finite element method.[6,7] It can be used not only for calculations of two-dimensional fields but also for calculations of fields which can be reduced to two-dimensional analogs[8] and also to solve essentially non-two-dimensional problems.

In many cases, the geometric symmetry of the electrodes in multipole systems is not clearly manifested in the field distribution since it may be impaired by the asymmetry of the boundary potential values. However, this latent symmetry can be identified by group-theoretical methods. Group methods can also be successfully used in calculations of electric and magnetic fields made by alternative numerical methods.

[1] L. G. Glikman, V. V. Radchenko, I. F. Spivak-Lavrov, and A. K. Shektybaev, Zh. Tekh. Fiz. **62**(8), 23 (1992) [Sov. Phys. Tech. Phys. **37**, 844 (1992)].
[2] I. F. Spivak-Lavrov, in *Proceedings of the SPIE Conference on Electron-Beam Sources and Charged-Particle Optics*, San Diego, Proc. SPIE **2522**, 149 (1995).
[3] L. D. Landau and E. M. Lifshits, *Quantum Mechanics: Non-Relativistic Theory*, 3rd. ed, (Pergamon Press, Oxford, 1977) [Russ. original, 3rd ed., Nauka, Moscow 1974, 702 pp.].
[4] M. A. Lavrent'ev and B. V. Shabat, *Theory of Functions of a Complex Variable* [in Russian], Nauka, Moscow (1972), 716 pp.
[5] I. F. Spivak-Lavrov, Nucl. Instrum. Meth. Phys. Res. Sect. A **363**, 491 (1995).
[6] B. Lencova, Nucl. Instrum. Meth. Phys. Res. Sect. A **363**, 190 (1995).
[7] X. Zhu, H. Liu, and E. Munro, in *Proceedings of the SPIE Conference on Electron-Beam Sources and Charged-Particle Optics*, San Diego, Proc. SPIE **2522**, 66 (1995).
[8] G. A. Doskeev and I. F. Spivak-Lavrov, Zh. Tekh. Fiz. **59**(1), 144 (1989) [Sov. Phys. Tech. Phys. **34**, 84 (1989)].

Translated by R. M. Durham

Superspherical cumulation. Converging shock waves with amplitudes increasing faster than in spherical cumulation

P. A. Voĭnovich

Center for Promising Technologies and Developments, 194156 St. Petersburg, Russia

M. O. Mdivnishvili and M. I. Taktakishvili

Institute of Physics, Georgian Academy of Sciences, 380077 Tbilisi, Georgia

I. V. Sokolov

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia
(Submitted October 14, 1996)
Zh. Tekh. Fiz. **69**, 10–18 (March 1999)

Experimental, numerical, and theoretical investigations are made of a gas flow generated by a pulsed high-current discharge in an axisymmetric cavity bounded by a spherical lens adjacent to a flat plate. It is shown that the shock wave forming in the discharge and converging toward the axis is accelerated and amplified as it converges. The amplitude of the shock wave increases faster than does that of a spherical converging shock wave. © *1999 American Institute of Physics.* [S1063-7842(99)00203-2]

1. INTRODUCTION

Nonsteady-state flows with converging shock waves have attracted the attention of researchers for many years. When the area of the front of a converging shock wave decreases with time and vanishes, the hydrodynamic energy density increases sharply in the flow behind the front, i.e., hydrodynamic cumulation[1] which is of considerable scientific and practical interest, takes place. Guderley, Landau, and Stanyukovich showed[1–3] that the flow behind the front of a converging spherical (or cylindrical) strong shock wave is self-similar and that the pressure behind the front P increases as the front radius r decreases, following a power law

$$P \sim r^{-2\delta}, \tag{1}$$

where $\delta = \text{const}$.

The exponent δ for a spherical wave is approximately twice that for a cylindrical wave[4] so that in both these cases the dependence of the pressure on the front area S can be considered to be the same with a high degree of accuracy.

Chester suggested an approximate theory (described in Ref. 4) for the amplification of a nonsteady-state shock wave propagating in a tapering channel. The theory is based on the assumption that the dependence $P(S)$ extracted from the self-similar solutions for cylindrical and spherical converging shock waves is also be applicable to the description of the amplification of a nonsteady-state shock wave in a channel of tapering cross section if the channel cross section is substituted as S. This theory, developed by Chisnell and Whitham, led to the so-called geometric dynamics of shock waves[4] (the CCW theory), which will be described briefly and applied below. In Ref. 5, we reviewed the known theoretical and experimental studies of converging shock waves in gases, as of the end of the nineteen-eighties. Note that in laboratory experiments on cylindrical converging shock waves, researchers usually do not attempt to create an extended shock wave along the symmetry axis, but investigate a fragment of a converging wave bounded by two plane walls (Fig. 1), which serve as optical windows or mirrors. Perry and Kantrowitz[6] first used this method to obtain a shadowgram of a converging cylindrical shock wave, and Takayama *et al.* obtained remarkable photographs of a wave front and its instabilities as ''flow sequences'' in Ref. 7 and their later studies.

A natural generalization of this approach is to produce converging shock waves in axisymmetric channels as shown in Figs. 2 and 3, i.e., channels bounded by the surfaces

$$z = \text{const}\, r^{m-1}, \tag{2}$$

where r, z, and φ are cylindrical coordinates, and $m = 1$, 2, and 3 for the channels shown in Figs. 1–3, respectively.

Bearing in mind the known advantages of the configuration shown in Fig. 1, we shall briefly discuss the scope for the experimental investigation of converging shock waves in the channels shown in Figs. 2 and 3. A converging shock wave which is similar in form to the fragment of a converging spherical wave can be generated in a cavity bounded by two conical surfaces. In principle, one of the channel walls may be a transparent conical refractor (refractive index n)

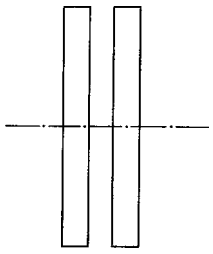

FIG. 1. Channel geometry for recording cylindrical cumulation.

with the expansion half-angle θ_1, and the other may be a conical reflector with the expansion half-angle

$$\Theta_2 = \arccos(\cos\Theta_1(n\sin\Theta_1 - \sqrt{1-n^2\cos^2\Theta_1})), \quad (3)$$

which is selected so that if the incoming optical beam is parallel to the optical axis, the outgoing beam will also be parallel, as shown in Fig. 2. It is then possible to use various optical methods. However, fairly serious experimental constraints are imposed by the astigmatism observed in the optical image of point sources in conical optical elements.

Conversely, the cavity shown in Fig. 3 is formed by adjacent walls having the most natural shape for optics, i.e., a spherical lens and a plane (or spherical) wall, which can serve as an optical window or plane mirror. The converging shock waves in this system are not only a convenient object for experimental investigation but are also interesting from the theoretical point of view, since an even stronger cumulation effect is observed for these than for a spherical shock wave (for brevity, we shall henceforth call it superspherical cumulation). This type of flow is analyzed in the present study. A brief report of this work containing some of the results presented here was published in Ref. 8.

2. INVESTIGATION OF A CONVERGING SHOCK WAVE BY OPTICAL METHODS

The dynamics of a converging and reflected wave in a cavity (Fig. 3) were studied using time-resolved shadowgraphy with slit scanning. A shock wave was created in atmospheric-pressure air using the annular gas-discharge source described in Ref. 9. This source can initiate a high-current filamentary gas discharge, and the discharge filament in this case was a thin toroid of major radius $R_0 = 50$ mm. The discharge time was ~ 20 μs, and the total energy input ~ 300 J. The annular shock wave source was clamped between a flat transparent Perspex plate and a thin spherical lens made of optical glass. The gap between the lens and the plate was 10 mm at the discharge point (which generates the shock wave) and <1 mm at the center of the system. The flow was illuminated using a flashlamp positioned on the side of the flat plate (on the left-hand side in Fig. 3). A slit which forms a shadow image on a photographic film was located on the optical axis of the system. An SFR-7 device was used for the time scanning. The photographic film recorded the position of the front of the converging (and then diverging) shock wave, i.e., an *XT* diagram.

For comparison we also recorded *XT* diagrams for a cylindrical shock wave. For this purpose the spherical lens was

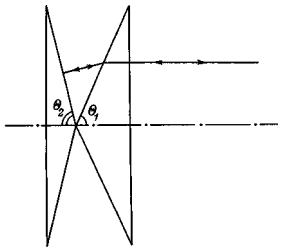

FIG. 2. Channel geometry for recording a fragment of spherical cumulation.

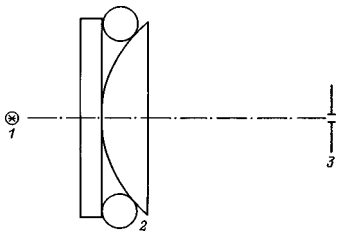

FIG. 3. Channel geometry and system for recording superspherical cumulation: *1* — flashlamp, *2* — annular shock wave source, and *3* — gap.

replaced by a second flat plate, so that the shock wave source was clamped between two disks. Reference measurements showed that the relative change in the optical magnification ratio caused by this substitution did not exceed 2%, so that the differences on the shadowgrams for cylindrical and supercylindrical cumulation are attributable to the different shock wave dynamics and not to changes in the properties of the optical system.

Shadowgrams for cylindrical and supercylindrical cumulation are shown in Figs. 4a and 4b, respectively. To facilitate comparison, Fig. 4c shows a photograph which is the result of combining Figs. 4a and 4b (partially retouched to compensate for the difference in the contrast of the photographs). The spatial coordinate is plotted along the vertical axis, and the distance along a vertical between the lines in Fig. 4 (trajectories of the fronts) corresponds to the distance between the converging and diverging fronts. The time is plotted along the horizontal axis. The slope of the front trajectory relative to the horizontal axis is proportional to the front velocity. In Fig. 4 the spatial scale is shown on the vertical axis, and the time scale on the horizontal axis.

The trajectories of shock-wave fronts from the left-hand edge of the figure to the point of intersection of the fronts (the geometric center of the system) correspond to a shock wave converging toward the center, while those running from the point of intersection to the right-hand edge of the figure correspond to a diverging shock wave. Figures 4a and 4b clearly show an increase in the angle between the trajectory of the shock-wave front on the diagram and the symmetry axis, which indicates that the converging front is accelerated. Since the velocity of the front of a strong shock wave is directly related to the hydrodynamic energy density behind the front, the acceleration of the front demonstrates an increase in the energy density, i.e., cumulation.

A comparison between the front trajectories for superspherical and cylindrical cumulation (Fig. 4c) leads to two conclusions. First, up to distances $\sim 0.4R_0$, i.e., 20 mm, from the center, the dynamics of the converging shock wave are almost the same in both cases since the trajectories of the fronts closely coincide. This implies that the differences in the experimental conditions of shock-wave generation caused by the different configurations of the cavity walls may be considered to be negligible. Conversely, at distances less than $0.4R_0$, as can be seen from Fig. 4c, the acceleration of the front for superspherical cumulation significantly exceeds that for cylindrical cumulation. The difference in acceleration is manifested as a steeper increase in the angle

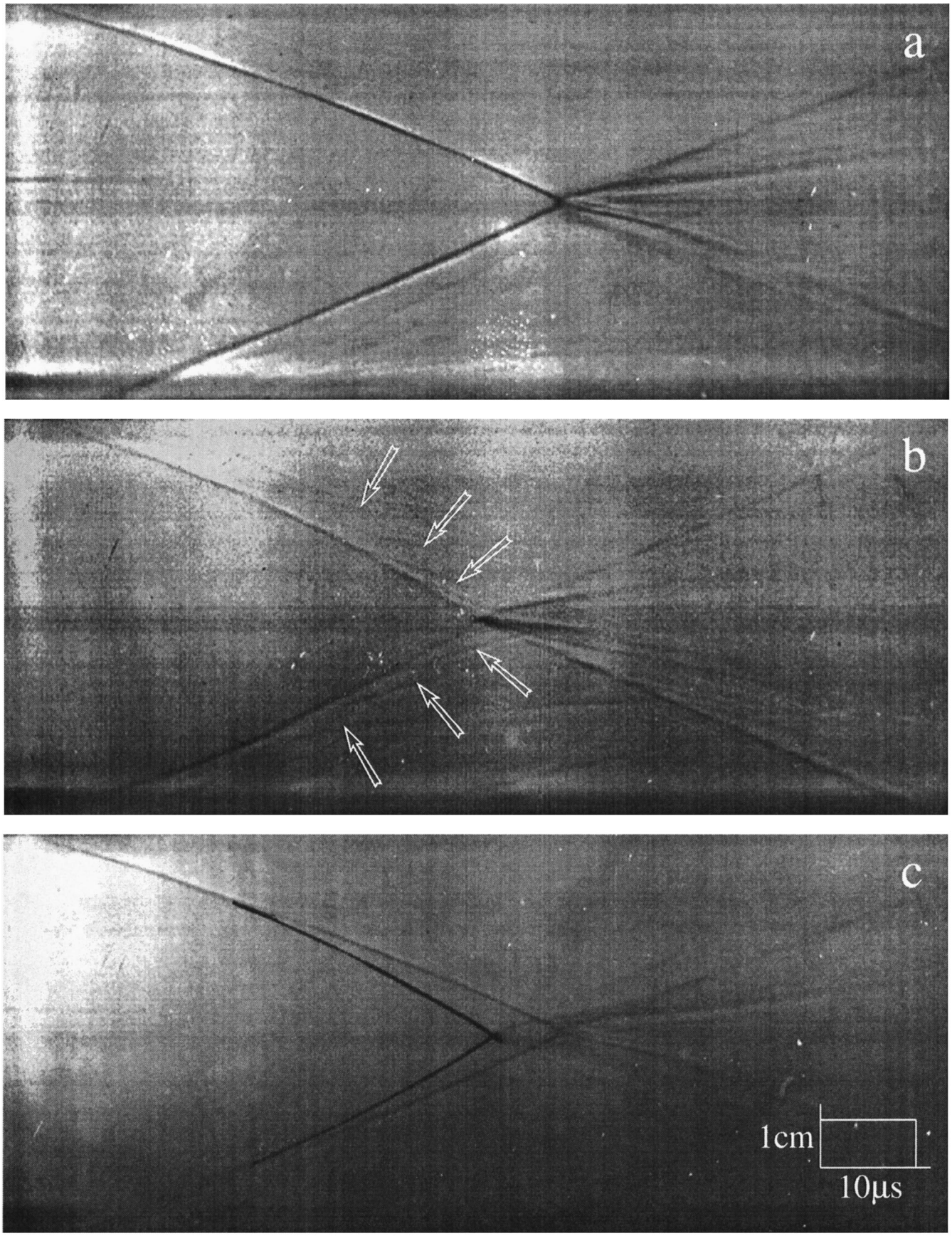

FIG. 4. Experimental time-resolved shadowgrams of cylindrical cumulation (a) and superspherical cumulation (b) and their superposition (c).

between the trajectory of the shock-wave front and the symmetry axis, and thus the times of convergence of the front toward the axis differ by 5 μs$\pm$50%.

The difference in the shock wave velocity measured from the slope of the front trajectory to the horizontal axis at $0.1R_0=5$ mm is 1.3×10^5 cm/s$\pm$20% for superspherical cumulation and $0.9\times10^5\pm15\%$ for cylindrical cumulation. The velocity measurements are unreliable closer to the center.

We regard these results as experimental evidence that superspherical cumulation leads to a steeper increase in energy density compared with cylindrical cumulation. This conclusion is fully consistent with the results of numerical calculations and an analytical theory presented below, which show that in terms of the degree of energy concentrating, even spherical cumulation cannot compete with the superspherical effect.

It is also interesting to note that the shadowgram in Fig. 4b clearly shows traces of hydrodynamic disturbances (indicated by the arrows) behind the converging shock-wave front, which have the following properties: the trajectories of the disturbances begin on the front trajectory (i.e., they are generated at the front of the converging shock wave); the disturbances propagate toward the center at a velocity slightly lower than the shock-wave velocity (and most likely propagate toward the center at the velocity of the gas flow behind the shock-wave front); finally, disturbances are observed in the shadowgram in Fig. 4b over almost the entire field of view but do not appear in Fig. 4a, i.e., their presence is typical of superspherical cumulation. It is natural to assume that these disturbances are contact discontinuities formed as a result of the successive irregular interaction between the converging shock wave and the cavity walls. This interaction between the shock wave and the walls and the associated stepped increase in pressure are important characteristic features of superspherical cumulation, as will be shown in the next section.

3. NUMERICAL SIMULATION OF SUPERSPHERICAL CUMULATION

A numerical simulation was made of a converging shock wave in the cavity shown in Fig. 3. The dimensions of the cavity and the experimental parameters differed from those used experimentally. The spherical wall had a radius of 250 mm and was in contact with a plane wall. In the radial direction the cavity was bounded by a cylindrical wall of radius $R_{\max}=50$ mm. The gas filling the cavity had an initial pressure of 1 atm, and the adiabatic exponent was $\gamma=7/5$.

The energy source for hydrodynamic motion was an initially assigned excess pressure $P_0=100$ atm in the gas volume bounded by the walls and the cylindrical surface $R_1=47$ mm (i.e., $R_1<r<R_{\max}$). A strong shock wave formed on the boundary of this volume at $r=R_1$ and then converged toward the symmetry axis.

The resulting flow was calculated using a unstructured flow-adapting mesh. The calculation scheme, criteria, and methods for restructuring of the mesh were described in Ref. 10. The calculations were continued until the radius of the converging shock wave was 5 mm.

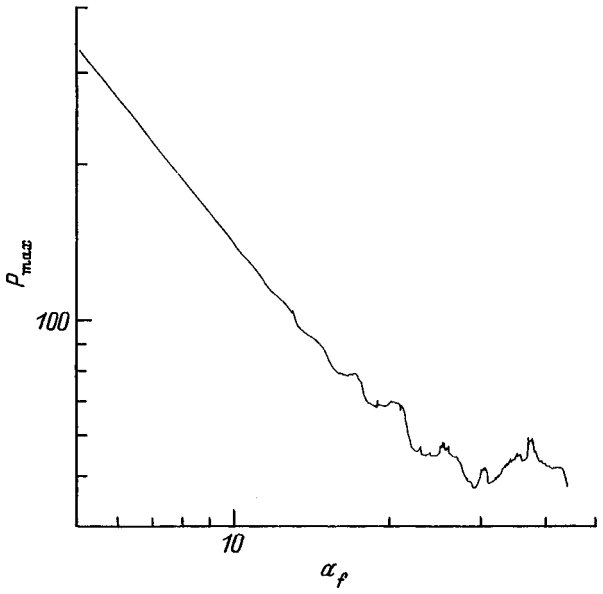

FIG. 5. Calculated pressure behind a converging shock-wave front for superspherical cumulation as a function of its radius (a.u.).

Initially, the converging shock wave underwent successive irregular interactions alternately with the plane and spherical walls. As a result, the shock wave amplification process took place discontinuously and was accompanied by the appearance of gasdynamic disturbances propagating transversely to the front of the converging shock wave, from wall to wall.

Figure 5 gives a logarithmic plot of the maximum pressure $P_{\max}$ along the flow at a given time as a function of the front radius a_f of the converging wave at that time. Figure 5 shows that as the shock wave converges, the pressure jumps become relatively weaker. At small radii the $P_{\max}(a_f)$ curve follows a power law (a straight line on the logarithmic scale), which specifically indicates that superspherical cumulation is self-similar at short distances from the axis.

The slope of the line in Fig. 5 corresponds to the self-similar dependence $P_{\max}\approx a_f^{-2\delta}$, where $\delta\approx0.62$. This value shows good agreement with the theory presented below, which gives $\delta\approx0.59$. The pressure-rise law obtained in the numerical calculations is steeper than that for a spherical shock wave in a gas with the same adiabatic exponent.

Thus, the numerical simulation shows that superspherical cumulation is accompanied by a rapid pressure rise, points out its self-similar behavior, and possibly explains the appearance of disturbances behind the converging shock-wave front which can be seen on the experimental shadowgrams.

4. ONE-DIMENSIONAL APPROXIMATION AND SELF-SIMILAR SOLUTION OF GASDYNAMIC EQUATIONS DESCRIBING SUPERSPHERICAL CUMULATION

In order to show that the theoretical description of superspherical cumulation is a general case of spherical and cylindrical cumulation, we shall investigate the transition to

a one-dimensional approximation (where the motion depends only on the ''radial'' coordinate) for all three cases simultaneously, taking $n=1,2,3$ in Eq. (1) as the parameter. We introduce the family of coordinate surfaces

$$\beta = z/r^{m-1} \tag{4}$$

and the coordinate α, which is conjugate to β,

$$\alpha^2 = (n-1)\times z^2 + r^2, \tag{5}$$

so that $\operatorname{grad}\alpha \cdot \operatorname{grad}\beta = 0$. The motion of the shock wave toward the center corresponds to $\alpha \to 0$, and the boundaries of the cavity are the coordinate lines $\beta=\beta_1$ and $\beta=\beta_2$.

The square of the length $dl^2 = dr^2 + dz^2 + r^2 d\varphi^2$ is expressed in the orthogonal coordinates α, β, and φ in terms of metric coefficients:

$$dl^2 = g_{\alpha\alpha} d\alpha^2 + g_{\beta\beta} d\beta^2 + g_{\varphi\varphi} d\varphi^2. \tag{6}$$

Here

$$g_{\alpha\alpha} = \alpha^2/[\alpha^2 + (n-1)(n-2)\beta^2 r^{2m-2}], \tag{7}$$

$$g_{\beta\beta} = g_{\alpha\alpha} r^{2m}/\alpha^2, \tag{8}$$

$$g_{\varphi\varphi} = r^2, \tag{9}$$

where for each given value of m, r should be expressed in terms of α and β using the equation

$$r^2 = \alpha^2 - (m-1)\beta^2 r^{2m-2}. \tag{10}$$

The equations of continuity and the Euler equations were given in Ref. 11 for an arbitrary orthogonal coordinate system. Assuming that no motion takes place along φ and that the flow does not depend on φ, we have

$$\frac{\partial \rho}{\partial t} + \frac{u_\alpha}{\sqrt{g_{\alpha\alpha}}}\frac{\partial \rho}{\partial \alpha} + \frac{u_\beta}{\sqrt{g_{\beta\beta}}}\frac{\partial \rho}{\partial \beta} + \frac{\rho}{\sqrt{g_{\alpha\alpha} g_{\beta\beta} g_{\varphi\varphi}}} \times \left[\frac{\partial}{\partial \alpha}(\sqrt{g_{\beta\beta} g_{\varphi\varphi}}\cdot u_\alpha) + \frac{\partial}{\partial \beta}(u_\beta \sqrt{g_{\alpha\alpha} g_{\varphi\varphi}})\right] = 0, \tag{11}$$

$$\frac{\partial u_\alpha}{\partial t} + \frac{u_\alpha}{\sqrt{g_{\alpha\alpha}}}\frac{\partial u_\alpha}{\partial \alpha} + \frac{u_\beta}{\sqrt{g_{\beta\beta}}}\frac{\partial u_\alpha}{\partial \beta} + \frac{1}{2} \times \left(\frac{u_\alpha u_\beta}{g_{\alpha\alpha}\sqrt{g_{\beta\beta}}}\frac{\partial g_{\alpha\alpha}}{\partial \beta} - \frac{u_\beta^2}{g_{\beta\beta}\sqrt{g_{\alpha\alpha}}}\frac{\partial g_{\beta\beta}}{\partial \alpha}\right) = -\frac{1}{\rho\sqrt{g_{\alpha\alpha}}}\frac{\partial P}{\partial \alpha}, \tag{12}$$

$$\frac{\partial u_\beta}{\partial t} + \frac{u_\alpha}{\sqrt{g_{\alpha\alpha}}}\frac{\partial u_\beta}{\partial \alpha} + \frac{u_\beta}{\sqrt{g_{\beta\beta}}}\frac{u_\beta}{\partial \beta} + \frac{1}{2} \times \left(\frac{u_\alpha u_\beta}{g_{\beta\beta}\sqrt{g_{\alpha\alpha}}}\frac{\partial g_{\beta\beta}}{\partial \alpha} - \frac{u_\alpha^2}{g_{\alpha\alpha}\sqrt{g_{\beta\beta}}}\frac{\partial g_{\alpha\alpha}}{\partial \beta}\right) = -\frac{1}{\rho\sqrt{g_{\beta\beta}}}\frac{\partial P}{\partial \beta}, \tag{13}$$

where u_α and u_β are the physical components of the velocity, and P and ρ are the pressure and density.

Equations (11)–(13) should be supplemented by the condition for adiabatic flow behind the shock-wave front and the boundary conditions

$$u_\beta\big|_{\beta=\beta_1,\beta_2} = 0. \tag{14}$$

We shall examine the possibility of separating the radial motion along α, i.e., the possible existence of motion which depends only on α and has only the velocity component u_α. This motion is described by

$$\rho_t + u_\alpha \frac{\partial \rho}{\partial \alpha} + \frac{\rho}{\alpha^m}\frac{\partial}{\partial \alpha}(\alpha^m u_\alpha) = 0, \tag{15}$$

$$\frac{\partial u_\alpha}{\partial t} + u_\alpha \frac{\partial u_\alpha}{\partial \alpha} = -\frac{1}{\rho}\frac{\partial P}{\partial \alpha}, \tag{16}$$

and condition (14) is automatically satisfied. We can go over from Eqs. (11)–(13) to Eqs. (15) and (16) provided the following conditions are satisfied

$$g_{\alpha\alpha} = 1 \quad \text{and} \quad \frac{\partial}{\partial \alpha}\left(\frac{r}{\alpha}\right) = 0.$$

For $m=1,2$ the following relations are satisfied identically

$$g_{\alpha\alpha} = 1 \quad \text{and} \quad \frac{\partial}{\partial \alpha}\left(\frac{r}{\alpha}\right) = 0,$$

and Eqs. (15) and (16) are valid in all space and describe cylindrically and spherically symmetric flows, respectively. The case of $m=3$ is more complex since the expressions for

$$g_{\alpha\alpha} - 1 \quad \text{and} \quad \frac{\partial}{\partial \alpha}\left(\frac{r}{\alpha}\right)$$

contain terms which depend on α and β. However, it is easy to see that these additional terms tend to zero when $\alpha \to 0$ (i.e., when the shock wave converges infinitely close to the center). Thus, whereas in the case of $m=1,2$ the radial motion can be separated in all space, for $m=3$ only local separation can be achieved as $\alpha \to 0$. Note that the local nature of the one-dimensional approximation (15) and (16) in the case of $m=3$ does not too severely restrict the validity of our self-similar solutions of Eqs. (15) and (16) since these solutions are only valid for $\alpha \to 0$.

We now turn to Eq. (13) to estimate the error of the one-dimensional approximation (15) and (16) for $\alpha \to 0$ and also to explain the physical meaning of the additional terms in $g_{\alpha\alpha}$. In addition to the pressure gradient along β, this equation also contains a generalized force, whose density is given by

$$f_\beta = \frac{1}{2}\frac{\rho u_\alpha^2}{\sqrt{g_{\beta\beta}}}\frac{\partial \ln g_{\alpha\alpha}}{\partial \beta}. \tag{17}$$

The quantity $(\partial/\partial\beta)\ln g_{\alpha\alpha}$ is directly related to the curvature of the coordinate line $\beta=\text{const}$, along which the radial motion takes place. Physically it is clear that motion along a curved line leads to the appearance of inertial forces transverse to the trajectory, such as the force (17).

Since for $m=3$ the distance between the walls decreases more rapidly ($\propto\alpha^2$) as $\alpha\to 0$ than does the distance from the center ($\propto\alpha$), as the shock wave converges, the motion bounded by the walls along β is rapidly damped behind the shock-wave front. Then in Eq. (13) all the terms containing u_β vanish, and the generalized force (17) is compensated by the pressure gradient $\partial P/\partial\beta$. This gives the β-dependent pressure variation

$$\delta P=\frac{1}{2}\rho u_\alpha^2\ln g_{\alpha\alpha}\approx-\rho u_\alpha^2\alpha^2\beta^2. \tag{18}$$

In particular, the pressure for the same value of α is slightly higher at a plane wall (where $\beta=\beta_2\neq 0$) than that at a spherical wall (where $\beta=\beta_1=0$). For $\alpha\to 0$ and bounded β, or, in dimensional quantities, at short distances from the center compared with the radius of the spherical wall, the pressure variation (18) is relatively small: $\delta P\ll P\sim\rho u_\alpha^2$. This inequality justifies going over to the one-dimensional approximation and permits estimation the error of the approximation as $0[\alpha^2\beta^2]$.

We shall now search for self-similar solutions of Eqs. (15) and (16) for $m=3$ which describe superspherical cumulation. It is interesting to note that these equations would describe a converging shock wave which is symmetric with respect to three polar angles in four-dimensional space for $m=3$. As a result of the more abrupt decrease in the area of the front as the wave converges ($S\propto\alpha^m$), we can naturally expect more abrupt cumulation compared with the cylindrical case. This conclusion was confirmed, and, moreover, the self-similarity exponent δ [see Eq. (1)] was three times (with exceptionally high accuracy) that for cylindrical cumulation and 1.5 times that for spherical cumulation, i.e., $\delta\propto m$.

The procedure for searching for the self-similarity exponent is described in Ref. 2 for a spherical shock wave. Thus, without changing the notation used in Ref. 2, we can merely present the equation for the functions represented by z and v which is valid for any m ($m=2$ in Ref. 2)

$$\frac{dz}{dv}=\frac{z}{1-v}\times\left[\frac{(((m+1)\gamma+1-m)v-2\delta-2)[z-(1-v)^2]}{v(1-v)(1+\delta-v)-z\left((m+1)v-\frac{2\delta}{\gamma}\right)}+(\gamma-1)\right], \tag{19}$$

as well as establish the relationship between the self-similarity index $\alpha_{[2]}$ introduced in Ref. 2 and the exponent δ appearing in Eq. (1): $1+\delta=1/\alpha_{[2]}$. Then, all the arguments in Ref. 2 hold for $m=3$.

A numerical integration of Eq. (19) can be used to find δ. For a polytropic index $\gamma=7/5$ this gives $\delta\approx 0.59$, and the pressure behind the shock-wave front increases as

$$P\propto\alpha^{-1.17}, \tag{20}$$

in good agreement with the results of the numerical calculations ($P\propto\alpha^{-1.24}$). For a spherical wave we would have $P\propto\alpha^{-0.78}$, and for a cylindrical wave[1)] we would have $P\propto\alpha^{-0.39}$.

The self-similarity exponent is proportional to m or, which amounts to the same thing, the function $P(S)$ is the same for all m to a high degree of accuracy. To some extent this detracts from the significance of the results obtained in this section: the universal nature of the dependence $P(S)$ in fact implies that the CCW theory can be applied to describe superspherical cumulation but in this theory the transition to the one-dimensional approximation is made more easily (and even more rigorously) than in pure hydrodynamics, after which the calculation of δ reduces to multiplying the known quantities by m.

5. APPLICATION OF THE CCW THEORY TO CALCULATE THE STABILITY OF SUPERSPHERICAL CUMULATION

The CCW theory[4] uses the function $\Phi(x,y,z)$ (where x, y, and z are the coordinates) to describe the dynamics of shock-wave fronts such that at each time t the surface of the front is described by the equation

$$\Phi(x,y,z)+V_{s0}\,t=0, \tag{21}$$

where V_{s0} is the velocity of sound in the undisturbed gas through which the shock wave propagates.

The Mach number M is then given by

$$M=\frac{1}{|\operatorname{grad}\Phi|}. \tag{22}$$

We introduce a system of coordinate lines or rays orthogonal to the surfaces (21) and assume that as each section of the surface of a shock-wave front moves along a ray tube the variation of M is determined only by the variation of the area of the front and, for a strong shock wave ($M\gg 1$), is given by

$$M\propto A^{-\delta_1}, \tag{23}$$

$$\delta_1=1\Big/\left(1+\frac{2}{\gamma}+\sqrt{\frac{2\gamma}{\gamma-1}}\right). \tag{24}$$

The equations of differential geometry give a relation between A and gradΦ and yield the equation for Φ

$$\operatorname{div}(M^{1+1/\delta_1}\operatorname{grad}\Phi)=0. \tag{25}$$

In the curvilinear coordinates α, β, and φ with the metric (7)–(9), Eq. (25) has the form

$$\frac{\partial}{\partial\alpha}\left[\alpha^m\left(\frac{r}{\alpha}\right)^{m+1}M^{1+1/\delta_1}\frac{\partial\Phi}{\partial\alpha}\right]+\alpha^{2-m}\frac{\partial}{\partial\beta}\times\left[\left(\frac{r}{\alpha}\right)^{1-m}M^{1+1/\delta_1}\frac{\partial\Phi}{\partial\beta}\right]+g_{\alpha\alpha}\left(\frac{r}{\alpha}\right)^{m-1}\times\alpha^{m-2}\frac{\partial}{\partial\varphi}M^{1+1/\delta_1}\frac{\partial\Phi}{\partial\varphi}=0. \tag{26}$$

Equation (26) for the radial motion is

$$\frac{\partial}{\partial\alpha}\left(\alpha^m\frac{\partial\Phi^{-1/\delta_1}}{\partial\alpha}\right)=0, \tag{27}$$

where for $m=1,2$ the variables (and the motion) are strictly separated, but for $m=3$ they are locally separated when $\alpha\to 0$, since in this last case we have $r/\alpha\to 1$ and $g_{\alpha\alpha}\to 1$. The solution of Eq. (27) is obviously

$$\Phi_0=cr^{1+m\delta_1},\quad c=\text{const}. \tag{28}$$

From Eqs. (22) and (28) we have

$$M\propto r^{-m\delta_1}, \tag{29}$$

and the self-similarity exponent δ introduced earlier, which is defined such that $P\approx M^2\approx r^{-2\delta}$, is

$$\delta=m\delta_1. \tag{30}$$

Thus, in the CCW theory the result $\delta\propto m$ obtained in the previous section (which, of course, is approximate, although highly accurate) is obtained as the exact relation (30). Since δ_1, which is defined by formula (24) in the CCW theory, is almost the same as the self-similarity exponent of a cylindrical converging shock wave calculated directly using the hydrodynamic equations, the CCW theory predicts a simple and very accurate expression for the self-similarity exponent $\delta=3\delta_1$ for superspherical cumulation.

We shall now linearize Eq. (26) relative to the small correction Φ_1 to Φ_0. For $m=3$ we also take into account small corrections to the metric coefficients for $\alpha\to 0$. We then have

$$\begin{aligned}&\frac{1}{\delta_1}\left(\frac{r}{\alpha}\right)^{m+1}\frac{\partial}{\partial\alpha}\left[\alpha^{-m\delta_1}\frac{\partial\Phi_1}{\partial\alpha}\right]-\alpha^{2-2m-m\delta_1}\frac{\partial}{\partial\beta}\\&\quad\times\left[\left(\frac{r}{\alpha}\right)^{1-m}\frac{\partial\Phi}{\partial\beta}\right]-\alpha^{-m\delta_1-2}\left(\frac{r}{\alpha}\right)^{m-1}\frac{\partial^2\Phi}{\partial\varphi^2}\\&\quad=\frac{\partial\Phi_0}{\partial\alpha}\alpha^{-m\delta_1}\frac{\partial}{\partial\alpha}\left(\frac{r}{\alpha}\right)^{m+1}.\end{aligned} \tag{31}$$

The boundary condition at the cavity walls is

$$\frac{\partial\Phi_1}{\partial\beta}\Big|_{\beta=\beta_1,\beta_2}=0. \tag{32}$$

For $m=3$ the right-hand side of Eq. (31) (which vanishes for $m=1,2$) equals $-8\alpha^{1-m\delta_2}\beta^2(\partial\Phi_0/\partial\alpha)$, and within Eq. (31) two essentially different formulations of the problem are possible. First, we can seek the particular solution Φ_{10} of Eq. (31) with a nonzero right-hand side, having imposed the constraint that for this particular solution $\Phi_{10}=0[\Phi_0]$ for $\alpha\to 0$. Such a solution does in fact exist, but here we shall not present the elementary procedure for constructing it as an expansion in powers of α. The corresponding small correction to Φ_0 is a refinement of the solution (28) and ensures that small (near the center) deviations of r/α and $g_{\alpha\alpha}$ from unity are taken into account. The existence of a solution Φ_{10} with these properties implies that the one-dimensional approximation of Φ_0 is accurate in first-order perturbation theory with respect to α.

The construction of a general solution of Eq. (31) with a zero right-hand side has a completely different meaning. The search for these corrections and the study of their behavior as $\alpha\to 0$ permits assessment of the stability of the one-dimensional solution (28) toward non-one-dimensional perturbations of the initial data.

Calculations of the behavior of the general solution of Eq. (31) with a zero right-hand side show that for $m=1-3$ all the non-one-dimensional disturbance modes are unstable since they decay more slowly than Φ_0 as $\alpha\to 0$. This implies an increase in the relative deviations of the shock wave profile from the symmetric one.

The growth rate of the various modes depends on m. As a cylindrical shock wave converges, sausage modes (i.e., disturbances which depend on z and do not depend on φ) grow preferentially. In a spherical shock wave all the disturbances which depend on the polar angles grow at the same rate. Here we present some calculations for the stability of superspherical cumulation ($m=3$).

For disturbances having a dependence on φ and β given by

$$\Phi=A(\alpha)e^{il\varphi}\cos\left(\pi k\frac{\beta-\beta_1}{\beta_2-\beta_1}\right), \tag{33}$$

where l and k are integers, $i^2=-1$, the equation for the amplitude has the form

$$\frac{1}{\delta_1}\frac{d^2A}{d\alpha^2}-\frac{3}{\alpha}\frac{dA}{d\alpha}+l^2\frac{A}{\alpha^2}+\frac{\pi^2k^2}{(\beta_2-\beta_1)^2}\frac{A}{\alpha^4}=0. \tag{34}$$

For $k\neq 0$ substituting $u=1/\alpha$ reduces Eq. (34) to

$$\begin{aligned}&\frac{1}{\delta_1}\left(u^4\frac{d^2A}{du^2}+2u^3\frac{dA}{du}\right)+3u^3\frac{dA}{du}+u^2l^2A\\&\quad+\frac{\pi^2k^2}{(\beta_2-\beta_1)}u^4A=0.\end{aligned} \tag{35}$$

The solution (35) is expressed in terms of Bessel functions with an imaginary index. The asymptotic behavior for $u\to\infty$ ($\alpha\to 0$) is determined by the factor

$$\begin{aligned}A&\propto u^{-(1+3\delta_{1/2})}\exp\left(i\frac{\pi k\sqrt{\delta_1}}{\beta_2-\beta_1}u\right)\\&=\alpha^{1+3\delta_{1/2}}\exp\left[\frac{i\pi k\sqrt{\delta_1}}{\alpha(\beta_2-\beta_1)}\right],\end{aligned} \tag{36}$$

i.e., Φ_1 oscillates with an amplitude which tends to zero as $|\Phi_1|\propto\alpha^{1+3\delta_1/2}$. The relative magnitude of the disturbance $|\Phi_1/\Phi_0|$, which characterizes the change in the shape of the front, increases as

$$\left|\frac{\Phi_1}{\Phi_0}\right|\propto\alpha^{-3\delta_{1/2}}. \tag{37}$$

Thus, disturbances which depend on β lead to a distortion of the wave-front profile which increases as $\alpha^{-0.3}$. However, purely azimuthal disturbances with $k=0$, i.e., which do not depend on β and depend only on the polar angle φ, grow considerably faster. Substituting $k=0$ into Eq. (34), we find that $A=\text{const}\,\alpha^b$, where

$$b=\frac{3\delta_1+1\pm\sqrt{(3\delta_1+1)^2-4\delta_1 l^2}}{2}. \tag{38}$$

Modes with $l\geqslant 2$ have the disturbance growth rate

$$\left|\frac{\Phi_1}{\Phi_0}\right|\propto\alpha^{-3\delta_{1/2}-1/2}\approx\alpha^{-0.8}. \tag{39}$$

Finally, a disturbance with $l=1$ splits into two modes, one of which gives the growth rate

$$\left|\frac{\Phi_1}{\Phi_0}\right|\propto\alpha^{-\frac{3\delta_1}{2}-\frac{1}{2}-\sqrt{\left(\frac{3}{2}\delta_1+\frac{1}{2}\right)^2-\delta_1}}\approx\alpha^{-1.5}. \tag{40}$$

It is easily understood that the extremely fast growth of this mode corresponds to accelerated displacement of the center of symmetry of the shock-wave front relative to the center of the cavity in a certain direction, which is determined by the initial form of the disturbance, as the shock wave converges. The action of this instability makes the experimental investigations of superspherical cumulation considerably more complex.

6. CONCLUSIONS

The results indicate that symmetric shock waves which are accelerated and amplified as they converge can be created in specially shaped axisymmetric cavities. The amplitude of these shock waves increases faster than does that of a spherical converging shock wave.

This work was financed by the Russian Fund for Fundamental Research, Grant No. 94-02-06691.

The authors are grateful to É. M. Barkhudarov, I. A. Kossyĭ, and E. V. Timofeev for supporting this work and E. L. Satunina and V. E. Terekhin for their technical assistance.

[1)] It can be said that superspherical cumulation is as strong as the spherical and cylindrical effects taken together ($\alpha^{-1.17}=\alpha^{-0.78}\alpha^{-0.39}$).

[1] E. I. Zababakhin and I. E. Zababakhin, *Unbounded Cumulation Phenomena* [in Russian], Nauka, Moscow 1988.
[2] L. D. Landau and E. M. Lifshits, *Fluid Mechanics*, 2nd ed. (Pergamon Press, Oxford, 1976) [Russ. original, 3rd ed., Nauka, Moscow 1985].
[3] K. P. Staniukovich, *Unsteady Motion of Continuous Media* (Pergamon Press, New York, 1960) [Russ. original, Gostekhizdat, Moscow (1955)].
[4] G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974) [Russ. trans., Mir, Moscow, 1977].
[5] I. V. Sokolov, Usp. Fiz. Nauk. **160**(11), 143 (1989) [Sov. Phys. Usp. **33**, 960 (1989)].
[6] R. W. Perry and A. Kantrowitz, J. Appl. Phys. **51**, 3126 (1954).
[7] K. Takayama, H. Kleine, and H. Groenig, Exp. Fluids **5**, 315 (1987).
[8] P. A. Voĭnovich, and I. V. Sokolov, Pis'ma Zh. Tekh. Fiz. **22**(13), 68 (1996) [Tech. Phys. Lett. **22**(7), 549 (1996)].
[9] I. V. Sokolov and V. E. Terekhin, in *Proceedings of the International Conference on Plasma Physics*, Delhi, India, 1989, edited by A. Sen and P. K. Kaw, p. 293.
[10] A. A. Fursenko, N. P. Mende, K. Oshima *et al.*, J. Comp. Fluid Dyn. **2**(1), 1 (1992).
[11] L. I. Sedov, *A Course in Continuum Mechanics, Vol. 1: Basic Equations and Analytical Techniques* (Wolters-Noordhoff, Groningen, 1971) [Russ. original, 3rd. ed., Nauka, Moscow, 1976, Vol. 1, pp. 178–179].

Translated by R. M. Durham

Numerical calculations of the radiation-induced strain rate of interstitial solid solutions. II Allowance for the main channels for the influence of impurities on radiation-induced creep

Yu. S. Pyatiletov and A. D. Lopuga

Institute of Atomic Energy, Scientific Nuclear Center of the Republic of Kazakhstan, 480082 Almaty, Kazakhstan

(Submitted July 8, 1996; resubmitted October 23, 1997)
Zh. Tekh. Fiz. **69**, 19–23 (March 1999)

A previously developed numerical method for calculating the radiation-induced creep rate [Yu. S. Pyatiletov and A. D. Lopuga, Tech. Phys. **45** (1999)] is used to study the influence of impurity atmospheres around dislocations and pores, impurity traps, and mobile impurity-vacancy and impurity-interstitial complexes on the radiation-induced strain rate of interstitial alloys. Quantitative data are obtained on the creep rate as a function of impurity concentration, and a physical interpretation allowing for the recombination of interstitial atoms and vacancies directly with one another, on impurity traps, and on mobile complexes is put forward. © *1999 American Institute of Physics.* [S1063-7842(99)00303-7]

1. INTRODUCTION

Self-consistent numerical calculations of the strain rate $\dot{\varepsilon}$ of interstitial alloys using the dislocation model of radiation-induced creep described in Ref. 1 require knowledge of the values of the parameters characterizing the material, its defect structure, and test conditions and then implementation of the calculation procedure using the method developed in Ref. 1. This we have done, and we present the results of calculating $\dot{\varepsilon}$ with allowance for the main channels for the influence of impurities on the strain rate of interstitial solid solutions. These include the formation of impurity atmospheres around dislocations and vacancy pores, the formation of fixed impurity–intrinsic-point-defect (IPD) complexes, and the formation of mobile impurity-IPD complexes. For these cases we calculated the concentration profiles of IPD's and impurity-IPD complexes near dislocations and pores, we determined the fluxes of vacancies, interstitials and impurity-interstitial complexes to sinks, and we obtained dependences of the radiation-induced creep rate on impurity concentration for various binding energies and migration energies of the impurity-IPD complexes. The calculations took into account the recombination of interstitial atoms and vacancies directly as a result of encounters between unlike IPD's, on impurity traps, and on mobile impurity-IPD complexes.

We adopt the notation introduced in Ref. 1 without any explanation. References to formulas from Ref. 1 are prefixed by the Roman numeral I.

2. ROLE OF IMPURITY ATMOSPHERES

The initial system of equations (I-12)–(I-15) can be used to calculate steady-state concentrations of vacancies, interstitials, and impurity-vacancy and impurity-interstitial complexes near dislocations and to perform further calculations of the radiation-induced creep rate with simultaneous allowance for all these channels for the influence of impurities on $\dot{\varepsilon}$. In order to identify the role of the impurity atmospheres alone, we shall assume that the impurities do not form complexes with IPD's. Then, if we set $D_{i\beta}=0$, $C_{i\beta}=0$ and $\chi_{i\beta}=0$, the expression (I-10) for the creep rate is simplified to

$$\dot{\varepsilon}=\frac{L(\sigma)}{6\Lambda}\rho_d(Z_I D_I C_I^0 - Z_V D_V C_V^0), \tag{1}$$

and instead of Eqs. (I-12)–(I-15) we obtain the two equations

$$D_V\nabla^2 C_V+\frac{D_V}{k_b T}(\nabla C_V \nabla E_V^{(d)}+C_V\nabla^2 E_V^{(d)})+G - D_V k_V^{(h)2} C_V-\alpha C_V C_I=0, \tag{2}$$

$$D_I\nabla^2 C_I+\frac{D_I}{k_b T}(\nabla C_I \nabla E_I^{(d)}+C_I\nabla^2 E_I^{(d)}) + G-D_I k_I^{(h)2} C_I-\alpha C_V C_I=0. \tag{3}$$

The expression for $E_\beta^{(d)}(\mathbf{r})$, which is the interaction energy of an IPD with a dislocation surrounded by an impurity atmosphere, has the form[2]

$$E_\beta^{(d)}(\mathbf{r})=\frac{\mu b(1+\nu)\Delta V_\beta}{3\pi(1-\nu)}\frac{\sin\varphi}{r} -\sum_{\mathbf{r}'}\frac{\Omega_1(1+\nu)^2\mu\Delta V_\beta}{18\pi(1-\nu)(1-2\nu)|\mathbf{r}-\mathbf{r}'|^3} \times\left(1-\frac{3(\mathbf{n},\mathbf{r}-\mathbf{r}')^2}{|\mathbf{r}-\mathbf{r}'|^2}\right)C_i^{(d)}(\mathbf{r}'). \tag{4}$$

Here ΔV_β is the relaxation volume of a β point defect, ν is the Poisson ratio, μ is the shear modulus, Ω_1 is the deviator part of the tensor which characterizes an elastic dipole simulating an interstitial impurity atom; r, φ, and z are the cylin-

drical coordinates associated with the dislocation line, and $C_i^{(d)}(\mathbf{r})$ is the equilibrium concentration of impurities in the atmosphere around a dislocation, which can be determined from the transcendental equation[2]

$$C_i^{(d)}(\mathbf{r})=\left\{1+\frac{1-C_i}{C_i}\exp\left[\frac{w^{id}(r,\varphi)}{k_bT}+\frac{1}{k_bT}\times\sum_{\mathbf{r}'}C_i^{(d)}(\mathbf{r}')\cdot(w^{ii}(\mathbf{r}-\mathbf{r}')-w^{ii}(\mathbf{r}_c^d-\mathbf{r}'))\right]\right\}^{-1}, \tag{5}$$

where $w^{id}(\mathbf{r})$ and $w^{ii}(\mathbf{r})$ are the interaction energies of an impurity atom with a dislocation and with another impurity atom, respectively, and $\mathbf{r}_c^d$ is the distance at which interaction between impurities and dislocations can be neglected.

The boundary conditions for the system of equations (1) and (2) are now assigned by formulas (I-23) and (I-31a), where the IPD concentration at the outer boundary of the calculation cell is determined from the system of nonlinear algebraic equations (I-27) and (I-28), where we set $C_{i\beta}=0$ and $\chi_\beta=0$. Introducing these changes into the calculation scheme proposed in Ref. 1, we calculate the radiation-induced creep rates using the material parameters of carbonaceous martensite: $b=2.48\times10^{-10}$ m, $\mu=8\times10^4$ MPa, $\nu=0.3$, $\Omega=1.2\times10^{-29}$ m^3, $\Delta V_I=1.1\ \Omega$, $\Delta V_V=-0.5\ \Omega$, $r_{dV}^0=3.40\times10^{-10}$ m, $r_{dI}^0=11.85\times10^{-10}$ m, $D_I^0=10^{-7}$ m^2/s, $D_V^0=0.58\times10^{-4}$ m^2/s, $E_I^m=0.2$ eV, $E_V^m=1.3$ eV (D_β^0 are the preexponential factors of the diffusion coefficients D_β, and E_β^m are the migration energies of β point defects). We also set $G=10^{-6}$ displacements per atom per second, $N_h=3.16\times10^{21}$ m^{-3}, $\rho_d=3.16\times10^{14}$ m^{-2}, $r_h=5\times10^{-9}$ m, and $T=660$ K.

The dashed curve in Fig. 1a gives the radiation-induced creep rate relative to $\dot\varepsilon_0$ (where $\dot\varepsilon_0$ is the creep rate of an impurity-free material having the same parameters) as a function of impurity concentration without allowance for the formation of complexes. It can be seen that $\dot\varepsilon/\dot\varepsilon_0$ decreases with increasing impurity concentration, the effect beginning to manifest itself at $C_i\approx10^{-4}$ and reaching 20% at $C_i=5\times10^{-3}$. This behavior is caused by the shielding of the elastic field of the dislocations by the impurity atmospheres, which reduces the capacity of the dislocations to trap interstitial atoms and vacancies. The most important factor was that the efficiency of the absorption of interstitial atoms Z_I by dislocations decreases more rapidly than the absorption efficiency for vacancies Z_V. A direct consequence is the decrease in the preference factor $B=(Z_I-Z_V)/Z_I$ in Fig. 1b (dashed curve), which leads to decreases in the dislocation climb rate and the creep rate as the impurity concentration increases.

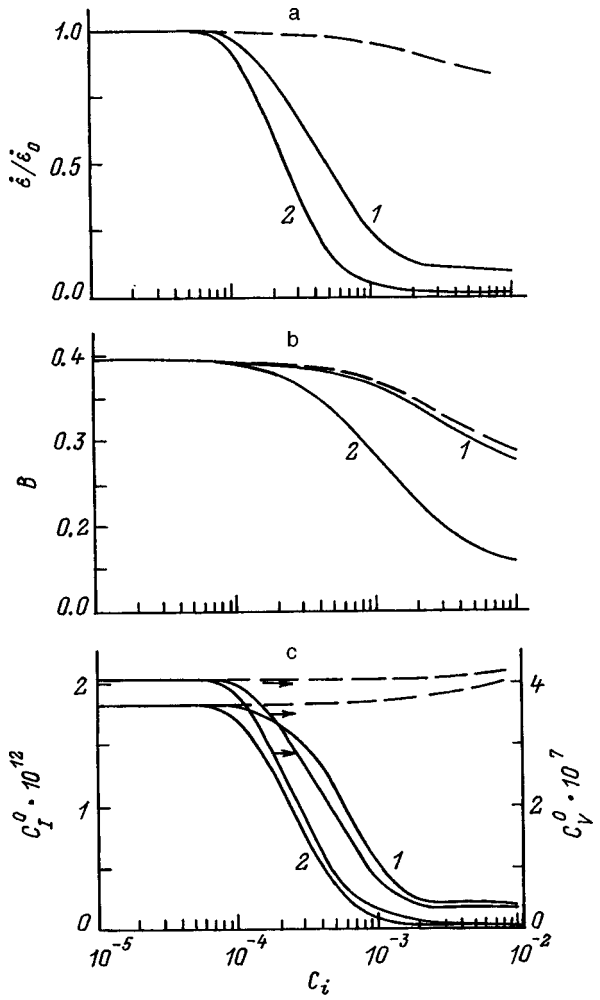

FIG. 1. Creep rate (a), preference factor (b), and steady-state IPD concentrations (c) as functions of the relative impurity concentration calculated for two values of the impurity-vacancy binding energy: $E_{iV}^b=0.3$ (*1*) and 0.5 eV (*2*).

3. ROLE OF IMPURITY TRAPS

Another channel for the influence of impurities on the strain rate of interstitial alloys under irradiation is the impurity trapping of interstitial atoms and vacancies and the formation of low-mobility complexes with a positive binding energy.[3,4] An IPD entering such an impurity trap becomes bound, becomes unable to migrate toward sinks, and, after a certain time has elapsed, annihilates with an incoming mobile IPD of opposite sign. Thus, impurity traps serve as IPD recombination centers and thereby influence the steady-state distribution of interstitial atoms and vacancies in the sample and, consequently, the values of Z_β and C_β^0 which determine the radiation-induced creep rate. The effect of the impurity traps should be enhanced when allowance is made for the formation of impurity atmospheres near sinks since the IPD fluxes to sinks are determined by the concentration profile of the IPD's not bound into complexes directly at the sink surfaces.

Since impurity-IPD complexes are assumed to be fixed, formula (I-10) for $\dot\varepsilon$ takes the form (1). In order to calculate the steady-state concentrations of vacancies and interstitial atoms near a dislocation with allowance for these complexes, we need to set $D_{i\beta}=0$ and $\alpha'=0$ in the system of equations (I-12)–(I-15), which is then rewritten as follows:

$$D_V\nabla^2 C_V+\frac{D_V}{k_bT}(\nabla C_V\nabla E_V^{(d)}+C_V\nabla^2 E_V^{(d)})+G$$
$$-D_V k_V^{(h)2}C_V-\alpha C_V C_I+\gamma_{iV}C_{iV}-\mu_I C_V C_{iI}$$
$$-\chi_V C_V(C_i^{(d)}-C_{iI}-C_{iV})=0, \quad (6)$$

$$D_I\nabla^2 C_I+\frac{D_I}{k_bT}(\nabla C_I\nabla E_I^{(d)}+C_I\nabla^2 E_I^{(d)})+G-D_I k_I^{(h)2}C_I$$
$$-\alpha C_V C_I+\gamma_{iI}C_{iI}-\mu_V C_I C_{iV}-\chi_I C_I$$
$$\times(C_i^{(d)}-C_{iI}-C_{iV})=0, \quad (7)$$

$$\chi_V C_V(C_i^{(d)}-C_{iI}-C_{iV})-\mu_V C_I C_{iV}-\gamma_{iV}C_{iV}=0, \quad (8)$$

$$\chi_I C_I(C_i^{(d)}-C_{iI}-C_{iV})-\mu_I C_V C_{iI}-\gamma_{iI}C_{iI}=0. \quad (9)$$

The boundary conditions for the vacancies and interstitial atoms are given by formulas (I-23) and (I-23a), and their concentration at the outer boundary of the calculation cell is determined from the system of equations (I-27)–(I-30) with $D_{i\beta}=0$ and $\alpha'=0$.

We shall first ascertain how the recombination of interstitial atoms and vacancies on impurity traps influences their concentration profiles around dislocations and pores for various impurity concentrations C_i and various binding energies of impurity-IPD complexes. For the calculations we shall use the material parameters given in the previous section. We established that effects associated with recombination of IPD's on impurity traps begin to play a significant role when the binding energy of the impurity-vacancy complexes E_{iV}^b exceeds 0.2 eV and that of the impurity-interstitial complexes E_{iI}^b exceeds 1.3 eV. Since the values of E_{iV}^b for carbon impurity atoms in α-Fe given by various authors[5,6] exceed 0.4 eV and the value of E_{iI}^b from Ref. 7 is 0.5 eV, in our case the impurity-interstitial complexes can be neglected. Thus, the calculations only take account of impurity-vacancy complexes, whose binding energy will be varied between 0.3 and 0.5 eV. The relative impurity concentration C_i will be varied between zero and 10^{-2}.

As a result of these calculations, we established that allowance for IPD recombination on impurity traps leads to two effects: first, the concentration of free vacancies in the bulk of the sample decreases and thus the concentration of impurity-vacancy complexes becomes nonzero; second, the slope of the $C_\beta(r)$ curves at the sink surfaces decreases, i.e., the fluxes of interstitial atoms and vacancies to sinks decrease. Both effects are enhanced as E_{iV}^b increases. Figure 1c gives the dependence obtained using these data of the IPD concentrations (C_I^0 and C_V^0), which appear directly in the formula (1) for the creep rate, on impurity concentration. Figure 1c shows that $C_I^0(C_i)$ and $C_V^0(C_i)$ behave similarly, falling sharply from 10^{-4} to 10^{-3} as C_i increases. The effect is enhanced as E_{iV}^b increases. Figure 1b gives the preference factor B, which specifies how much $\dot\varepsilon$ is influenced by the difference between the efficiencies of absorption of interstitial atoms and vacancies by dislocations, as a function of C_i. Figure 1 shows that the preference factor B also decreases monotonically with increasing C_i but not as strongly as C_β^0. Finally, Fig. 1a (solid curves) shows the relative creep $\dot\varepsilon/\dot\varepsilon_0$ as a function of the impurity concentration C_i for $E_{iV}^b=0.3$ and 0.5 eV. Also plotted is the dependence of $\dot\varepsilon/\dot\varepsilon_0$ on C_i obtained without consideration of impurity traps (dashed curve). We can see from a comparison of the solid and dashed curves that when impurity traps are taken into account, the creep rate decreases considerably faster with increasing C_i and that the effect is enhanced as E_{iV}^b increases.

A comparison of Figs. 1a and 1b reveals that a decrease in the concentration of free point defects as a result of the capture of vacancies by impurity traps and the recombination of vacancies with interstitial atoms on these traps has a stronger influence on the creep rate than do changes in the concentration profiles of IPD's near sinks, which determine the values of the parameters Z_β, Y_β, and B. The contribution made by changes in C_β^0 to the reduction in $\dot\varepsilon$ is between four and five times greater than that caused by changes in B.

4. ROLE OF MOBILE IMPURITY–INTRINSIC-POINT-DEFECT COMPLEXES

Let us now consider the situation when the impurity-vacancy and impurity-interstitial complexes are mobile. Quite clearly, these complexes are not only recombination centers for interstitial atoms and vacancies but, like free IPD's, diffuse toward sinks. Having reached a sink, the complexes dissociate: the IPD's are absorbed by the sinks, and the freed impurity atoms diffuse back into the bulk of the sample. Thus, some of the vacancies and interstitial atoms migrate toward sinks in the free form and some migrate in a bound form as complexes. In this case, the fluxes of all these mobile components to dislocations contribute to the creep rate calculated using the general formula (I-10) and the magnitudes of these fluxes are determined by the concentration profiles of the IPD's and mobile impurity-IPD complexes at the sink surfaces.

We shall make numerical calculations of these concentration profiles by solving the system of equations (I-12)–(I-15) with the boundary conditions (I-23), (I-24), and (I-31) using the same calculation parameters as in the previous sections. We shall also set $E_{iI}^b=0.6$ eV, $E_{iV}=0.5$ eV, $D_{iV}^0=D_V^0$, and $D_{iI}^0=D_I^0$. The dilatation volume of the complexes $\Delta V_{i\beta}$, which determines the interaction energy of the complex $i\beta$ with dislocations and pores, is taken as the sum of the dilatation volumes of an IPD ΔV_β and a carbon impurity atom ΔV_i. Since no reliable data are available on the migration energies $E_{i\beta}^m$ of the complexes, we shall vary these in a range where the lower limit corresponds to the migration energy of the most mobile component and the upper limit corresponds to the value at which the complex is virtually fixed. For instance, for the impurity-vacancy complexes we shall vary E_{iV}^m between 0.8 (which corresponds to the migration energy of a carbon atom in α-Fe) and 2.2 eV, and for the impurity-interstitial complexes we shall vary E_{iI}^m between 0.3 (which corresponds to the migration energy of an interstitial atom) and 1 eV. The range of impurity concentrations studied is between 10^{-5} and 5×10^{-3}.

The calculations show that allowance for the mobility of the impurity-IPD complexes leads to two principal effects: first, the concentration of free IPD's in the bulk of the

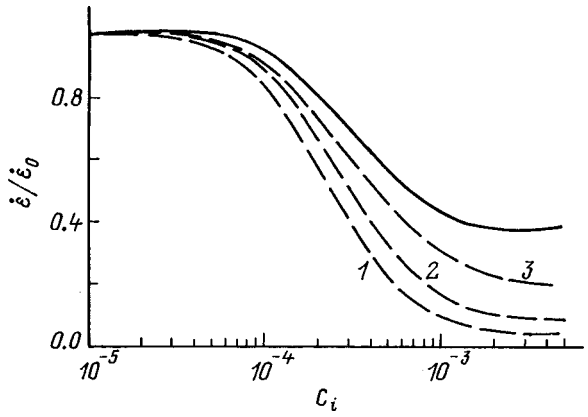

FIG. 2. Radiation-induced creep rate versus relative impurity concentration for cases where both types of complexes are mobile (solid curve), both types of complexes are fixed (*1*), only impurity-interstitial complexes are mobile (*2*), and only impurity-vacancy complexes are mobile (*3*).

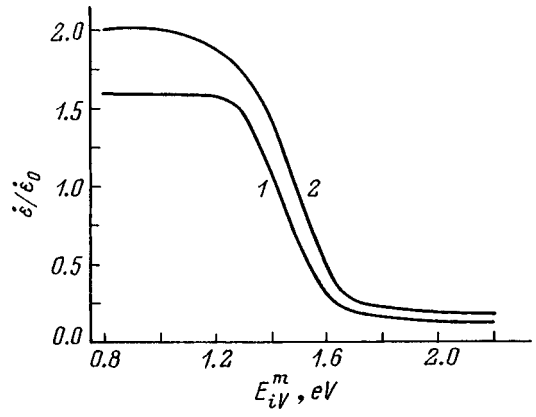

FIG. 3. Radiation-induced creep rate versus migration energy of impurity-vacancy complexes when the impurity-interstitial complexes are fixed (*1*) and when they are mobile ($E^m_{iI}=0.3$ eV) (*2*).

sample increases in comparison with fixed complexes; second, fluxes of mobile complexes to sinks appear, and the slope of the $C_\beta(r)$ curves increases at the sink surfaces, i.e., the fluxes of free IPD's to sinks increase. These effects are enhanced as the complexes become more mobile.

In order to identify the direct influence of the mobile complexes on the creep rate, we calculated the fluxes of all the mobile components, i.e., interstitial atoms (I), vacancies (V), and complexes (iI and iV), to dislocations. The results of calculations of these fluxes as a function of impurity concentration for $E^m_{iV}=1.6$ and 0.3 eV showed that as C_i increases, the IPD fluxes $I^{(d)}_I$ and $I^{(d)}_V$ to dislocations decrease, whereas the fluxes of complexes $I^{(d)}_{iI}$ and $I^{(d)}_{iV}$ increase. The creep rate is determined by the flux differences $I^{(d)}_I - I^{(d)}_V$ and $I^{(d)}_{iI} - I^{(d)}_{iV}$. Thus, at low impurity concentrations the main contribution to $\dot{\varepsilon}$ comes from fluxes of free IPD's, whereas as C_i increases, the fluxes of IPD's in complexes play a more significant role. Hence, we can postulate that the creep rate will be higher when the mobility of the complexes is taken into account than in the case of fixed complexes. This is confirmed by calculations of the creep rate as a function of impurity concentration, whose results are plotted in Fig. 2 (solid curve). For comparison, this figure also shows the dependence of $\dot{\varepsilon}/\dot{\varepsilon}_0$ on C_i for fixed complexes (curve *1*).

In order to establish how the creep rate is influenced by each type of mobile complex separately (iI or iV), we calculated $\dot{\varepsilon}/\dot{\varepsilon}_0$ for two cases. In the first case, we assumed that only the iI complexes are mobile (curve *2* in Fig. 2) and that the migration energy of the complexes equals the migration energy of an interstitial atom, i.e., the complexes exhibited maximum mobility. In the second case, only the iV complexes were mobile (curve *3* in Fig. 2), but the migration energy of the complexes was taken to be fairly high ($E^m_{iV}=1.6$ eV), i.e., the iV complexes had a comparatively low mobility. Nevertheless, in the second case, the plot of $\dot{\varepsilon}/\dot{\varepsilon}_0$ as a function of C_i deviates more substantially from curve *1* for fixed iI and iV complexes than in the first case and is closer to curve *3*. This behavior suggests that the mobile impurity-vacancy complexes have a much stronger influence on the creep rate than do the impurity-interstitial complexes.

We calculated $\dot{\varepsilon}/\dot{\varepsilon}_0$ as a function of the migration energy of the impurity-vacancy complexes E^m_{iV} to obtain quantitative confirmation of this conclusion (Fig. 3). We considered two situations: when the iI complexes are fixed (curve *1*) and when they have maximum mobility ($E^m_{iI}=0.3$ eV, curve *2*). Thus, curves *1* and *2* demarcate the range of all possible values of the migration energy for the iI and iV complexes. Figure 3 shows that at high values of E^m_{iV} the creep rate is close to that for a material with fixed complexes. As the mobility of the iV complexes increases and E^m_{iV} approaches the migration energy of free vacancies E^m_V, the creep rate becomes higher than that for an impurity-free material. This can be explained by the fact that some of the IPD's are transferred to dislocations and pores as part of complexes whose interaction energy with sinks is higher than that of free IPD's, and thus the dislocation preference is higher.

5. DISCUSSION

In the preceding sections we have examined three main channels for the influence of impurities on the rate of radiation-induced creep $\dot{\varepsilon}$ in interstitial alloys. The simplest scenario for interpretation purposes is that impurity-IPD complexes do not form and the influence of impurities on $\dot{\varepsilon}$ is mediated only by the impurity atmospheres around dislocations and pores. The impurity atmospheres have their own elastic field, which partly shields the elastic field of the dislocations and suppresses its action on the IPD's. As a result, the dislocations have a lower capacity for trapping IPD's, which reduces their absorption efficiencies for interstitial atoms (Z_I) and vacancies (Z_V) in comparison to the absorption observed in an impurity-free metal, with Z_I decreasing more substantially than Z_V. Consequently, an increase in the impurity concentration in the material reduces the preference factor B and, therefore, the creep rate.

The situation is more difficult to analyze when impurity-IPD complexes with some degree of mobility form in the irradiated material. The formation of these complexes essentially leads to the appearance of new types of IPD recombination and new IPD fluxes to sinks as part of the mobility of

the complexes. Thus, the interstitial atoms and vacancies formed continuously in an alloy as a result of irradiation are inactivated in two ways: either they escape to sinks, diffusing toward them in free form or as part of complexes, or they are lost as a result of recombination. Several types of IPD recombination can take place in the presence of defects. First, ordinary mutual recombination occurs when free interstitial atoms and vacancies migrating in the material meet; second, IPD's recombine when they encounter complexes containing IPD's of opposite type; and, third, recombination takes place when two complexes containing IPD's of opposite type meet.

Before the processes taking place in irradiated interstitial alloys undergoing deformation can be fully described, we need to know the quantitative relations between the fractions of IPD's lost as a result of these types of recombination. An analysis of the results of the numerical calculations showed that when the material contains fixed complexes, the recombination of IPD's upon encounters between interstitial atoms and impurity-vacancy complexes predominates. When the $i\beta$ complexes are mobile, IPD's are mainly lost when unlike complexes encounter one another. The fraction of defects lost as a result of all three types of recombination increases with decreasing temperature.

[1] Yu. S. Pyatiletov and A. D. Lopuga, Zh. Tekh. Fiz. **69**(1), 64 (1999) [Tech. Phys. **44**, 59 (1999)].
[2] Yu. S. Pyatiletov, S. B. Kislitsin, and N. I. Edemskyi, Radiat. Eff. Defects Solids **25**, 223 (1993).
[3] A. N. Orlov and Yu. V. Trushin, in *Radiation Defects in Metallic Crystals*, edited by Sh. Sh. Ibragimov [in Russian], Nauka, Alma-Ata (1978), pp. 30–40.
[4] L. K. Mansur and M. H. Yoo, J. Nucl. Mater. **74**(2), 228 (1978).
[5] R. A. Johnson, Phys. Rev. **134**, A1329 (1964).
[6] K. Masuda, Phys. Status Solidi B **166**, 9 (1983).
[7] A. Vehanen, P. Hautojanvi *et al.*, Phys. Rev. B **25**, 762 (1979).

Translated by R. M. Durham

Control of the degree of long-range compositional order in ceramics of the compound perovskites $Pb(B'_{1/2}B''_{1/2})O_3$ using recrystallization

V. Yu. Shonov and I. P. Raevskiĭ

Rostov State University, Physics Research Institute, 344090 Rostov-on-Don, Russia
(Submitted January 22, 1997; resubmitted April 7, 1998)
Zh. Tekh. Fiz. **69**, 24–30 (March 1999)

It is established that an ordered state forms in $Pb(Yb_{1/2}Nb_{1/2})O_3$ at fairly low temperatures (650–800 °C). Subsequent high-temperature heat treatment of the ceramic without additives (sintering or additional annealing) does not produce any significant change in the degree of long-range compositional order *s* because of the low diffusion rate of the Yb and Nb ions. The addition of Li_2CO_3, which forms a liquid phase, creates conditions for the dissolution of grains with a high value of *s*, the nucleation of new crystallization centers, and the growth of grains with a new equilibrium value of *s* at the sintering (annealing) temperature. © *1999 American Institute of Physics.* [S1063-7842(99)00403-1]

1. INTRODUCTION

The phenomenon of compositional ordering, i.e., a variation of the long-range order *s* in the placement of ions of different types in identical crystallographic positions in compound oxides of the perovskite family, has been observed in $Pb(B^{III}_{1/2}B^{V}_{1/2})O_3$ ferroelectrics, where B^{III} = Sc or In and B^{V} = Nb or Ta (Refs. 1–5). Recently ordering effects were observed in perovskite compound oxides, where B^{III} = Lu, Yb, Tm, or Er (Refs. 6–8). For all these compounds, changes in *s* are possible as a result of an order–disorder phase transition at values of T_{OD} exceeding 1000 °C. Changes in *s* have a very strong influence on the ferroelectric properties, especially the Curie temperature T_C. For instance, in the ordered state the perovskite oxides $Pb(B^{III}_{1/2}B^{V}_{1/2})O_3$ (B^{III}=Lu, Yb, Tm, or Er; B^{V}=Nb or Ta) are antiferroelectrics with a high Curie point, whereas in the disordered state they are ferroelectrics with a Curie point 150–220 °C lower. The value of *s* can be determined quantitatively by means of x-ray structural analysis.

The usual method for achieving different values of *s* is based on obtaining oxides at different temperatures in the vicinity of T_{OD}, while additional high-temperature annealing is used to change *s* (Refs. 1–3). The process of obtaining different values of *s* as a result of sintering compound perovskite ceramics at different temperatures, in fact, combines the two processes, i.e., the formation of an ordered or a disordered state during synthesis and then the alteration of *s* as a result of subsequent annealing at the sintering temperature T_s. However, high-temperature annealing is frequently ineffective since a short holding time does not change *s* and an increase in the annealing time leads to a transformation of the perovskite phase into a nonferroelectric pyrochlore phase.[9,10]

These methods did not produce any significant changes in *s* for ceramics of the compound oxides $Pb(B^{III}_{1/2}B^{V}_{1/2})O_3$, where B^{III}=Lu, Yb, Tm, or Er, and B^{V}=Nb or Ta. Sintering these ceramics at various T_s causes only negligible deviations of *s* from the values for ordered states and slightly shifts T_C toward lower temperatures. However, a highly disordered state has been achieved[6–8] by sintering these perovskite compound oxide ceramics with added Li_2CO_3 in the same temperature range as without this additive (Fig. 1).

Here we take the example of $Pb(Yb_{1/2}Nb_{1/2})O_3$ (T_{OD} = 1125 °C) (Ref. 10) to investigate the mechanism for the formation of states with an equilibrium value of *s* during the heat treatment (sintering and subsequent annealing) of $Pb(B^{III}_{1/2}B^{V}_{1/2})O_3$ ceramics with a small addition of Li_2CO_3.

2. SAMPLE PREPARATION AND MEASUREMENT METHOD

Polycrystalline $Pb(Yb_{1/2}Nb_{1/2})O_3$ samples were prepared by a conventional ceramic technology using a two-stage technique (preliminary synthesis at 600–900 °C for 4–6 h followed by sintering at 850–1200 °C for between 20 min and 6 h) or a single-stage technique (sintering at 850–1200 °C for between 20 min and 6 h) both with and without an addition of 2–15 mol % Li_2CO_3. The addition of Li_2CO_3 was introduced before sintering. The initial materials were ultrapure PbO, Yb_2O_3, and Nb_2O_5 oxides. In some cases, to increase the content of the perovskite phase, the compound

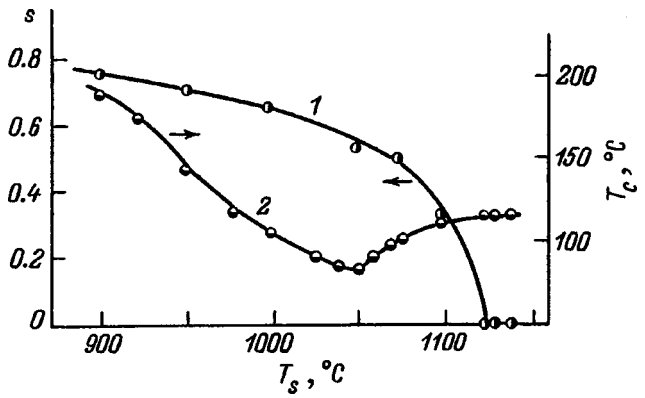

FIG. 1. Dependence of the degree of long-range order *s* (*1*) and the Curie point T_C (*2*) on the sintering temperature T_s for a PYN(L) ceramic.

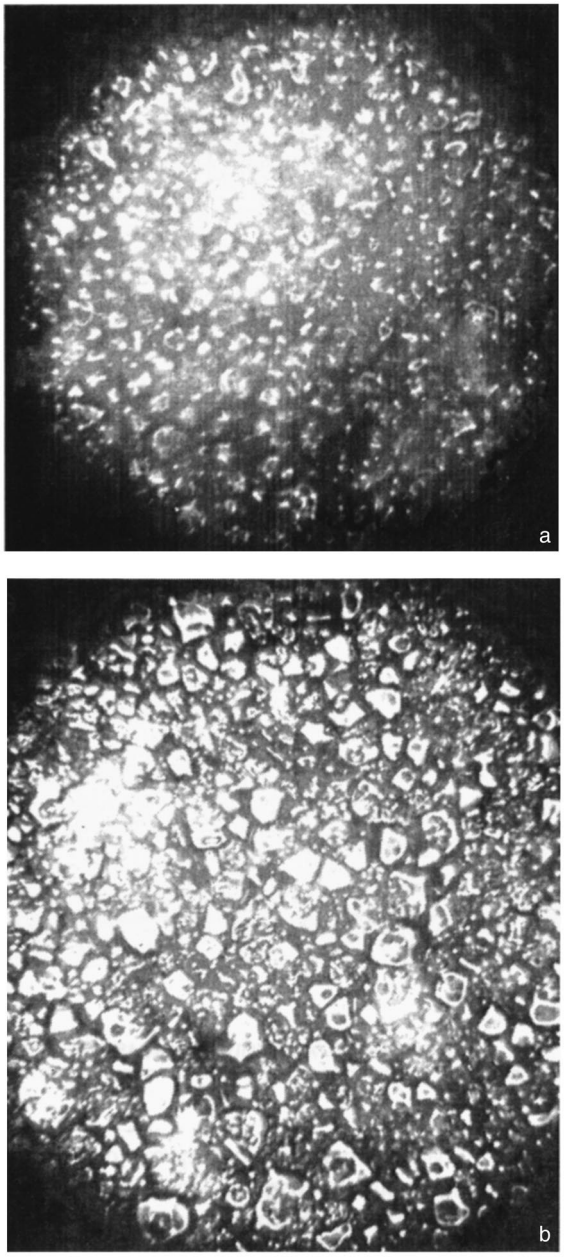

FIG. 2. Influence of annealing at 1100 °C for 15 min on the microstructure of PYN samples: a — synthesized at 800 °C for 3 h; b — sintered at 900 °C for 2 h. Magnification: 1000×.

$YbNbO_4$, which was preliminarily synthesized at 1000 °C for 4 h, was used instead of Yb_2O_3 and Nb_2O_5,[11] and excess PbO was also introduced to 20 wt. %, which evaporated completely during the anneal.

Samples pressed into disks 10 mm in diameter and 1–2 mm thick at a pressure of $3–4\times10^7$ Pa were sintered in a closed alundum crucible. A charge of powdered $Pb(Mg_{1/3}Nb_{2/3})O_3$ was used to produce a suitable atmosphere in the crucible. For the measurements the samples were polished and electrodes were deposited by firing silver paste.

The x-ray structural analysis of the samples was performed on a Dron-2.0 diffractometer using Cu K_α radiation. The value of s was determined using a method described by Setter and Cross[12] by comparing the experimentally determined intensity ratio of the superstructural and main reflections with the ratio calculated for a completely ordered structure

$$s^2=(I_{111}/I_{200})_{\text{observed}}(I_{111}/I_{200})^{-1}_{\text{calc}}.$$

The temperature dependences of the dielectric constant ε were investigated using a Tesla BM-484 capacitance bridge at 1.6 kHz at a sample heating (cooling) rate of 2 K/min.

The microstructure of the ceramics was investigated using a Carl Zeiss–Jena NU-2E optical microscope.

TABLE I. Average grain size of PYN and PYN(L) ceramics as a function of the sintering temperature T_s and the additional annealing temperature T_{ann}.

Ceramic	T_s, °C	T_s, h	T_{ann}, °C	T_{ann}, h	d, μm
PYN	900	2	⋯	⋯	1.6
	900	2	1100	1/4	2.9
PYN(L)	900	3	⋯	⋯	4.2
	900	3	1100	3/4	0.9
	1050	2.5	⋯	⋯	2.4
	1075	3	⋯	⋯	1.1
	1075	3	950	5	1.4
	1110	2	⋯	⋯	1.0

3. EXPERIMENTAL RESULTS

For the investigations we used $Pb(Yb_{1/2}Nb_{1/2})O_2$ samples with (PYN(L)) and without (PYN) added Li_2CO_3, which were sintered at various temperatures $T_s=850–1200$ °C. The samples all underwent additional annealing in the range 900–1200 °C at temperatures both above and below T_s. Before and after annealing we carried out x-ray structural investigations, determined s, and investigated the microstructure and the dependence $\varepsilon(T)$.

The samples of PYN sintered at $T_s=900–1200$ °C had high values of s (about 0.8) and T_C (between 260 and 290 °C). Additional high-temperature annealing near the order–disorder phase transition had no significant influence on the ordered structure formed during synthesis and annealing. An increase in the annealing time reduced the content of the perovskite phase while increasing the content of the pyrochlore phase. For the annealed PYN samples the maximum of $\varepsilon(T)$, which corresponds to T_C, did not shift, but its value decreased as the content of the pyrochlore phase increased. In all cases, x-ray structural analysis of the PYN samples revealed monoclinic distortion of the perovskite subcell and the presence of two types of superstructural reflections, which are caused by antiparallel displacement and compositional ordering of the Yb and Nb cations.

The PYN ceramic sintered at low temperatures was fine-grained. Increases in T_s (T_{ann}) and the holding time at high temperature increased the average grain diameter d (Fig. 2, see Table I).

Sintering PYN(L) ceramics in the range $T_s=850–1200$ °C produced samples with different values of s (Fig. 1), depending on the proximity of T_s to the order–disorder phase transition temperature. When the PYN(L)

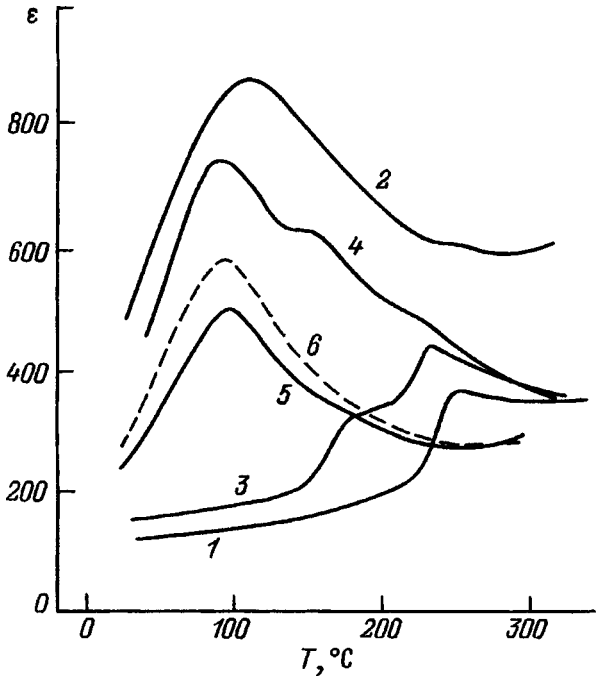

FIG. 3. Influence of nondestructive annealing on the course of $\varepsilon(T)$ for PYN(L) samples with an addition of 7.5 mol % Li_2CO_3: *1, 3, 5* — before annealing, *2, 4, 6* — after annealing; *1* — $T_s=900\,^\circ$C, 20 min; *2* — $T_{ann}=1050\,^\circ$C, 2 h; *3* — $T_{ann}=750\,^\circ$C, 3 h; $T_s=900\,^\circ$C, 3 h; *4* — $T_{ann}=1100\,^\circ$C, 1.5 h; *5* — $T_s=1050\,^\circ$C, 3 h; *6* — $T_{ann}=950\,^\circ$C, 5 h.

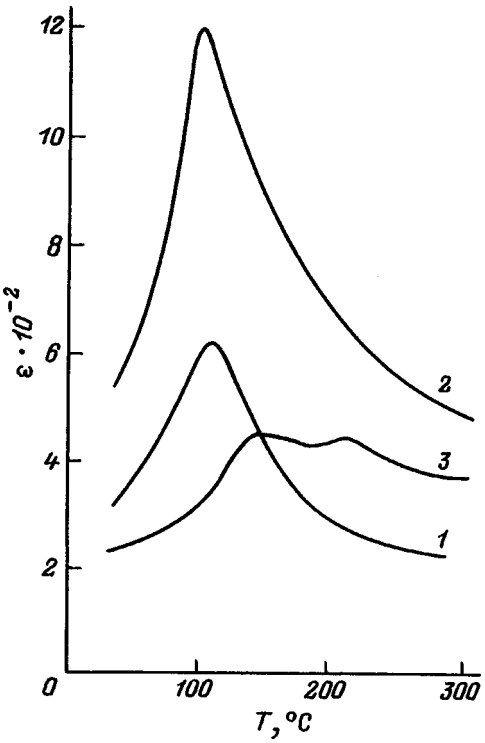

FIG. 4. Influence of annealing at $T_{ann}=900\,^\circ$C, $t=2.5$ h on the course of $\varepsilon(T)$ for PYN(L) samples sintered at $T_s=1100\,^\circ$C, $t=50$ min with an addition of 7.5 mol % Li_2CO_3: *1* — before annealing, *2* — after annealing with intermediate grinding, *3* — after annealing with intermediate grinding and an addition of Li_2CO_3.

samples were additionally annealed, the values of s and T_C only varied if T_s was less than $\sim 1020\,^\circ$C and the holding time under sintering was 5–6 h.

Figure 3 shows the temperature dependences of ε for PYN(L) ceramics sintered and annealed at various temperatures. The shift of the $\varepsilon(T)$ peak corresponds to the change in s. The changes in s and T_C under additional annealing took place within a comparatively short time (between a few tens of minutes and a few hours, depending on T_{ann}). In this case, there was no increase in the content of the pyrochlore phase, which did not exceed 15% for various samples. Attempts to change s for PYN(L) ceramics by means of second or further anneals were generally unsuccessful. The value of s for ceramics which had already been annealed, as well as for PYN(L) samples sintered at $T_s \geqslant 1020\,^\circ$C, was changed as a result of annealing only when an additional quantity (2–15 mol %) of Li_2CO_3 was added to the bulk of the sample. This was accomplished by grinding the ceramic, adding Li_2CO_3 to the powder, and pressing out a new sample which then underwent further annealing.

Figure 4 shows the plot of $\varepsilon(T)$ for a PYN(L) ceramic obtained at $T_s=1110\,^\circ$C and annealed at $T_{ann}=900\,^\circ$C after these operations. The annealing induces a transition from a disordered to a more ordered state, which is manifested by the appearance of a high-temperature maximum on the $\varepsilon(T)$ curve (curve *3*). After the first such anneal for 2.5 h, the $\varepsilon(T)$ curve begins to resemble the analogous curve for a PYN(L) ceramic obtained after sintering at $T_s=T_{ann}$. Annealing these samples with no additional Li_2CO_3 did not change the form of the $\varepsilon(T)$ curve if we disregard the increase in the height of the maximum of ε caused by an increase in the density of the ceramic (curves *1* and *2*).

Under high-temperature annealing, in addition to the main maximum of $\varepsilon(T)$, the PYN(L) ceramics sometimes exhibited additional anomalies, which diminished or disappeared completely when the holding time for the first anneal was increased or repeated anneals were carried out after adding Li_2CO_3.

Unlike the case of the PYN ceramics, an increase in T_s or the annealing temperature T_{ann} for PYN(L) samples reduced the average grain diameter d (see Table I and Figs. 5a–5e). For samples sintered at T_s above 1020 °C with no additional Li_2CO_3, no significant change in d was observed as a result of annealing for 2–5 h. A further increase in the annealing time enhanced the porosity and increased the content of the pyrochlore phase. The PYN(L) samples sintered at $T_s \leqslant 1020\,^\circ$C and annealed at $T_{ann} > T_s$ exhibit a correlation between the decrease in d and the appearance and growth of new anomalies on the $\varepsilon(T)$ curve as a function of the annealing time (Fig. 5f).

Additional high-temperature annealing ($900 < T_{ann} < 1120\,^\circ$C) of PYN ceramics to which Li_2CO_3 was added after annealing can give the same values of s and T_C as in the case of the sintering of a PYN(L) ceramic provided $T_{ann}=T_s$.

4. DISCUSSION OF RESULTS

According to our results, the presence of Li_2CO_3 in the bulk of a PYN ceramic plays a significant role in the formation of disordered states during sintering or in changing the

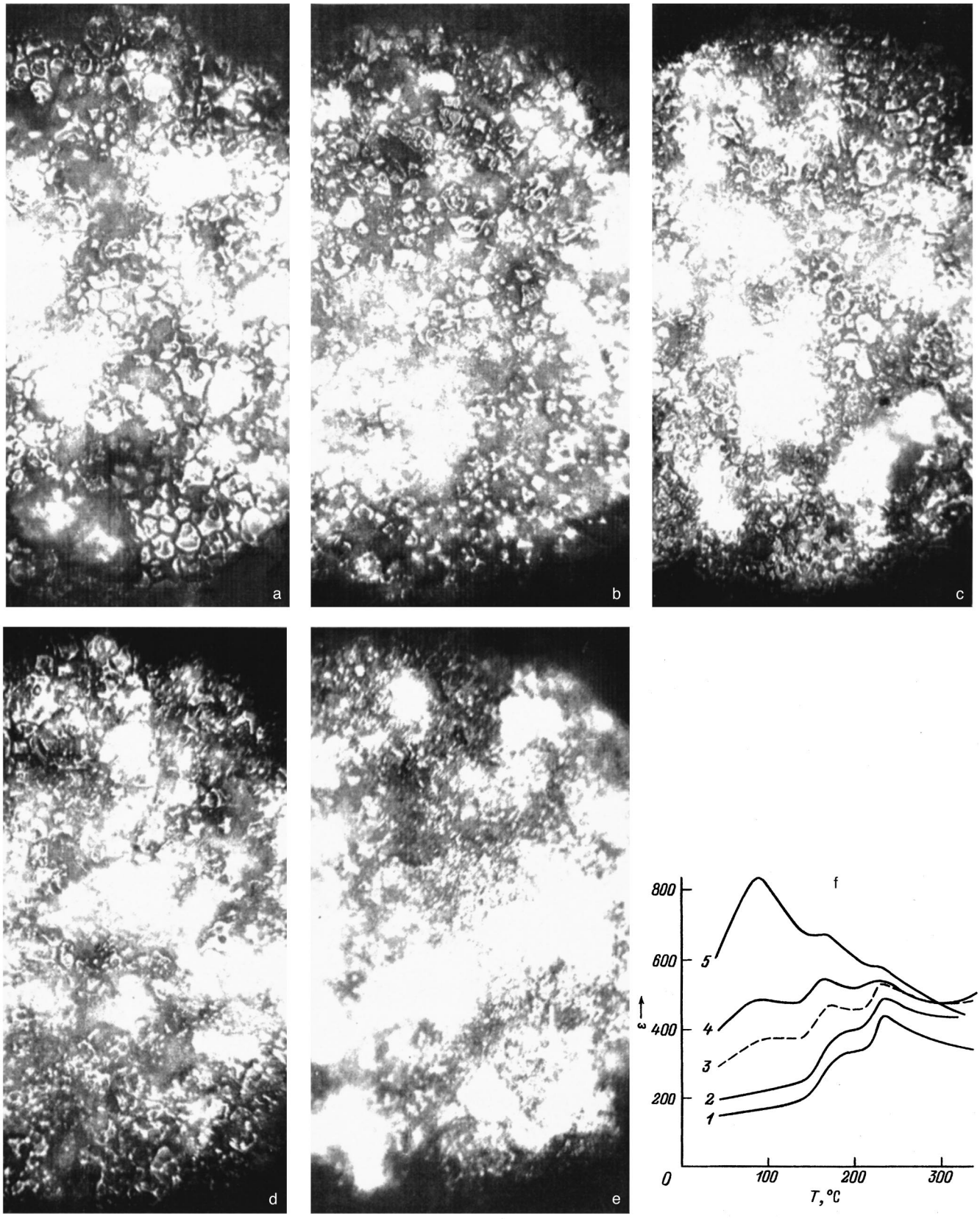

FIG. 5. Influence of the annealing time t at 1100 °C on the microstructure (a–e) and $\varepsilon(T)$ curve (f) for PYN(L) ceramics synthesized at 700 °C for 6 h and sintered at 900 °C for 3 h; T_{ann}, min: *1* — 0, *2* — 5, *3* — 10, *4* — 20, *5* — 40. Magnification: 1000×.

existing degree of compositional order as a result of additional high-temperature annealing near T_{OD}. In order to ascertain the specific mechanism for the action of this additive on the ordering process, let us examine what we consider to

be the two most probable mechanisms among the various possibilities.

1. The additive increases the rate of ordering (disordering) by increasing the diffusion coefficient of the Yb and Nb ions. Such an increase may be attributed to the appearance of additional vacancies in Yb and Nb sites caused by the dissociation of some solid solution, which forms in the initial stage of synthesis and contains lithium in Yb and Nb sites, as a result of the intensive evaporation of Li at high temperatures.

2. The additive provides for the crystallization of new ceramic grains with the equilibrium value of the compositional order s_e at $T_s(T_{\text{ann}})$ and the dissolution of grains with nonequilibrium s. Favorable conditions for the formation of states with the equilibrium value s_e at the sintering (or annealing) temperature are created by the formation of an additional liquid phase by the mineralizer, from which new grains (with $s=s_e$) can form without needing to overcome the energy barrier for the solid-phase processes.

It should be noted that the influence of lithium on the properties of PYN is clearly not associated with the formation and/or dissociation of solid solutions, as is the case of $BaTiO_3$, for example. We know[13,14] that the addition of lithium salts to $BaTiO_3$ can substantially reduce the annealing temperature and T_C because of the formation of a solid solution where Li replaces Ti. However, as a result of the evaporation of Li, the value of T_C for these solid solutions depends strongly on the amount of additive and the annealing time τ. As τ increases, T_C increases monotonically to values typical of $BaTiO_3$ without additives as a result of the evaporation of Li and dissociation of the solid solution. For PYN(L) an investigation of the influence of the amount of Li_2CO_3 and the holding time during annealing indicated that changes in T_C are not caused by the formation or dissociation of solid solutions involving lithium but are caused by a change in the value of s (Refs. 8 and 10). Thus, in our view the second mechanism is dominant.

As was shown in Refs. 8 and 10, the solid-phase synthesis of PYN is accompanied by the formation of a structure with a high value of s, which is the equilibrium value at fairly low temperatures of 700–800 °C, where the perovskite modification of PYN forms. During sintering or additional annealing at higher temperatures, the ordered structure remains in a nonequilibrium state since the rate of the transition to a more stable state with lower s is low. This low rate of conversion is the result of the high activation energy for the diffusion of B ions. It is quite safe to say that the sintering or additional annealing of a PYN ceramic does not result in the formation and growth of regions of a new disordered phase. Otherwise, the intensity of the superstructural reflections caused by compositional and antiferroelectric ordering would decrease and additional anomalies would appear on the $\varepsilon(T)$ curve, in contradiction to the experimental results. An increase in T_s, T_{ann}, or the holding time at high temperatures increases the average grain diameter d in PYN ceramics (see Table I and Fig. 2), which is normal for perovskite compound oxide ceramics.[15] A small number of large grains grow from the finely dispersed mass in polycrystalline samples as a result of secondary recrystallization.

Crystalline phases which are not in equilibrium under specific conditions of temperature, pressure, and composition occur fairly frequently in ceramic systems. An example may be provided by quartz,[16] whose presence in a porcelain mass during annealing at 1200–1400 °C does not result in the formation of stable tridymite or cristobalite forms. In this case, the rapid establishment of an equilibrium state is promoted by the introduction of a liquid-phase-forming mineralizer, i.e., calcium oxide. Quartz dissolves in the liquid phase and then crystallizes in the form of a new phase. The mineralizer permits the transfer of material by dissolution and thus overcomes the energy barrier which exists for direct phase conversion.[16]

The additive Li_2CO_3, which has a low melting point[17] (730 °C), probably also functions as a mineralizer, forming a liquid phase when a PYN ceramic is sintered. The formation of a PYN(L) ceramic evidently takes place as follows. As the temperature rises in the range ~600–750 °C, solid-phase synthesis accompanied by the growth of grains of a perovskite phase with a high degree of order s_0, which is the equilibrium value in this temperature range, takes place. Then, at higher temperatures a liquid phase forms, creating conditions for the nucleation of new crystallization centers and the growth of grains with the equilibrium degree of order s_e at the sintering temperature T_s. The driving force for the formation of new grains with s_e stems from the difference between the free energies of the grains with s_0 (a nonequilibrium value at T_s) and s_e (the equilibrium value at T_s). The grains with the nonequilibrium degree of order s_0 at T_s are gradually dissolved by the liquid phase. In this context, we note that not any additive which leads to the appearance of a liquid phase and does not form solid solutions with PYN is a mineralizer. For instance, using additions of GeO_2 and SiO_2 instead of Li_2CO_3 during the sintering or annealing of PYN near T_{OD} allowed us to obtain disordered states. The samples always had a high degree of long-range compositional order s and were antiferroelectrics with a Curie point $T_C=230$–285 °C.

During the annealing of a PYN(L) ceramic, the formation of a stable phase with s_e for T_{ann} depends on the content of lithium in the bulk of the sample, which rapidly evaporates at temperatures above 950 °C. In PYN(L) samples sintered at $T_s \geqslant 1020$ °C the amount of lithium left after sintering is probably insufficient to form a liquid phase. This explains why it is impossible to vary s by additional annealing (Figs. 3 and 4). Moreover, an increase in the annealing time does not cause any significant change in the average grain size (see Table I). The same degree of compositional order s can be obtained for PYN ceramics either directly by sintering the samples with added Li_2CO_3 or by annealing previously sintered samples. In the latter case several anneals with the introduction of an additional quantity of Li_2CO_3 into the sample before each one are sometimes required to achieve the equilibrium value s_e at T_{ann}. A PYN ceramic is considered to have the equilibrium value s_e during sintering, if an increase in the holding time does not lead to changes in s and T_s, and during additional annealing, if subsequent anneals also do not influence s and T_s. In our view, the existence of additional anomalies on the temperature depen-

dences of ε for PYN ceramics in addition to the maximum can be attributed to grains with different s and different structures being present in the bulk of the sample. Figures 5a–e clearly show the kinetics of the dissolution of some grains (for example, those with a high degree of order s) and the formation of others. Large grains ($d>4\mu$m) are obtained for PYN(L) samples synthesized at low values of T_s $\sim 900\,^\circ$C. For PYN(L) samples sintered at $T_s=900\,^\circ$C high-temperature annealing at $T_{\text{ann}}=1100\,^\circ$C ensures the gradual dissolution of large grains together with the formation and growth of grains of a new phase not exceeding $d<1\mu$m. Spatial inhomogeneities of the degree of order are thus manifested by the simultaneous existence of grains with different s, which gives rise to several anomalies on the temperature dependences of ε. This is consistent with the results of the x-ray structural investigations, according to which PYN(L) samples with a high degree of order belong to the orthorhombic system, whereas those with a low degree of order belong to the rhombohedral system, and those in the intermediate state ($0.35<s<0.65$) are mixtures of both these two phases. An increase in the total volume of some phases and a reduction in others alter the form of the $\varepsilon(T)$ curve (Fig. 5f). The presence of a large quantity of the liquid phase during sintering at low temperatures $\sim 900\,^\circ$C accelerates secondary recrystallization and causes accelerated grain growth. This explains the large size of the grains in the PYN(L) samples compared with PYN (Fig. 2 and Figs. 5a–5e; see Table I). In contrast, the influence of the liquid phase on grain growth at high $T_s(T_{\text{ann}})$ is negligible because of the rapid reduction in the amount of this phase caused by the rapid evaporation of Li from the sample. Another factor delaying the growth of equilibrium grains at $T_s(T_{\text{ann}})$ may be the presence of a pyrochlore phase, whose content increases with increasing T_s in PYN(L) samples, as was shown in Ref. 10. These factors cause the PYN(L) ceramics obtained or annealed at high temperature to be fine-grained.

As was noted in the Introduction, appreciable changes in the degree of compositional order as a result of sintering with added Li_2CO_3 are observed for the entire series of compounds $Pb(B^{III}_{1/2}B^{V}_{1/2})O_3$ (B^{III}=Lu, Yb, Tm, or Er; B^{V}=Nb or Ta). The mechanism described above for achieving the equilibrium value of s at the sintering or annealing temperature as a result of recrystallization is evidently realized in these compounds. The essential feature of the new technique for controlling the degree of compositional order in perovskite compound oxides is the initiation of a recrystallization process, which can be induced not only by using a mineralizer, but also by other methods.

5. CONCLUSIONS

In $Pb(Yb_{1/2}Nb_{1/2})O_3$ an ordered state forms at relatively low temperatures (650–800 °C). Subsequent high-temperature heat treatment of the ceramic without additives (sintering or additional annealing) does not cause any substantial change in s because of the low diffusion rate of the Yb and Nb ions. The addition of Li_2CO_3, which forms a liquid phase, creates conditions for the dissolution of grains with a high nonequilibrium value of s, the nucleation of new crystallization centers, and the growth of grains with a new equilibrium value of s at the sintering (annealing) temperature. Thus, the initiation of a recrystallization process during which grains are formed with the equilibrium value of s at the sintering (annealing) temperature is the essential feature of this new method for controlling s in complex compositionally ordered compounds.

[1] C. G. F. Stenger, F. L. Sholten, and A. J. Burggraaf, Solid State Commun. **32**, 989 (1979).
[2] C. G. F. Stenger and A. J. Burggraaf, Phys. Status Solidi A **61**, 257, 653 (1980).
[3] A. A. Bokov, I. P. Raevskiĭ, and V. G. Smotrakov, Fiz. Tverd. Tela. (Leningrad) **26**, 2824 (1984) [Sov. Phys. Solid State **26**, 1708 (1984)].
[4] A. Kania, Ferroelectric Lett. **11**, 107 (1990).
[5] N. Yasuda and S. Imamura, in *Abstracts of the 7th European Meeting on Ferroelectricity*, Dijon, France, 1991, p. 354.
[6] V. V. Shonov, A. A. Bokov, I. P. Raevski *et al.*, in *Proceedings of the International Conference on Electronics and Ceramics — Production and Properties*, Riga, 1990, pp. 116–118.
[7] A. A. Bokov, V. V. Shonov, I. P. Raevski *et al.*, in *Abstracts of the 7th European Meeting on Ferroelectricity*, Dijon, France, 1991, p. 128.
[8] V. Yu. Shonov, Dissertation for Candidate's Degree [in Russian], Rostov-on-Don (1992), 155 pp.
[9] N. Yaohiko, H. Inagaki, and S. Inamura, Jpn. J. Appl. Phys. **31**, 574 (1992).
[10] A. A. Bokov, V. V. Shonov, I. P. Raevsky *et al.*, J. Phys.: Condens. Matter **5**, 5491 (1993).
[11] T. R. Shrout, S. L. Swartz, and M. J. Haun, Bull. Am. Ceram. Soc. **63**, 808 (1984).
[12] N. Setter and L. E. Cross, J. Mater. Sci. **15**, 2478 (1980).
[13] J. M. Haussone, G. Desgardin, Ph. Bajolet, and B. Raveau, J. Am. Ceram. Soc. **66**, 801 (1983).
[14] J. M. Haussone, J. Lostec, O. Regreny *et al.*, Silic. Ind., No. 11–12, 231 (1984).
[15] K. Okazaki, *Ceramic Engineering for Dielectrics*, (Tokyo, 1969) [Russ. trans., Energiya, Moscow, 1976, 336 pp.].
[16] W. D. Kingery, *Introduction to Ceramics* (Wiley, New York–London–Sydney, 1960) [Russ. trans., Stroĭizdat, Moscow, 1964, 534 pp.].
[17] V. A. Rabinovich and Z. Ya. Khavin, *Concise Chemical Handbook* [in Russian], Khimiya, Leningrad (1977), 376 pp.

Translated by R. M. Durham

Theory of the oscillatory dependence of the conductivity of two-component dielectric composites at room temperatures

S. O. Gladkov

N. N. Semenov Institute of Chemical Physics, Russian Academy of Sciences, 117977 Moscow, Russia
(Submitted May 20, 1997)
Zh. Tekh. Fiz. **69**, 31–34 (March 1999)

It is predicted that at room temperatures a hopping mechanism of charge transfer plays a very important role and leads to temperature oscillations of the conductivity $\sigma(T)$ of a dielectric composite. The dependence of the conductivity $\sigma(\omega)$ on the frequency of an alternating electric field is calculated. The relation obtained can be used to determine, first, the electron relaxation times and, second, and more importantly, the frequency of electron tunneling through the dielectric matrix from measurements of the conductivity in various frequency ranges.
 [S1063-7842(99)00503-6]

Intensive studies of the physical properties of complex composite structures[1–4] have revealed some very interesting anomalies. For example, Bresler *et al.*[5] and Piccard and Derby[6] measured the resistance of polypropylene with conducting graphite inclusions as a function of the graphite concentration p, the temperature T, and the frequency ω of an applied alternating electric field $\mathbf{E}(t)=\mathbf{E}_0 e^{i\omega t}$, where $\mathbf{E}_0$ is the amplitude. In this context it is interesting to note that measurements of the resistance as a function of the concentration of graphite inclusions revealed a threshold concentration of graphite $p_{\rm cr}$, above which the composite becomes conducting. Curiously, $p_{\rm cr}$ can have different values, depending on the structure of the underlying composite matrix. For example, if a dielectric (say polypropylene) is simply filled with graphite, we find $p_{\rm cr}$ = 40–50% of the bulk composition of the matrix. However, if nonconducting spherical inclusions, whose radius R is considerably greater than the dimension of the graphite fillers b, are initially added to the polypropylene and graphite is then added to this structure, occupying only the region between the spheres (the technology is available to do this), $p_{\rm cr}$ =4–5% for such a structure. Qualitatively, the behavior of the resistance $\rho(p)$ is the same for both these composites (Fig. 1), although the bulk fractions of the conductor in them differ substantially.

The aim of the present paper is to describe the conductivity of a dielectric with conducting inclusions using a theoretical model and to determine its dependence on temperature, the frequency of an alternating field, and concentration. In order to assess how this composite behaves in an external alternating electric field, we envisage a single electron in a conducting particle, and we ''follow'' its path to the opposite end. Clearly, the electron will ''move'' along some elementary filament, in which conducting and dielectric regions alternate. If we stretch this elementary filament and take into account that the composite contains many such filaments, it becomes clear that its resistance is the sum of series-connected dielectric and metal components. Thus, the resistivity of this structure is

$$\rho(p)=p\rho_1+(1-p)\rho_2, \tag{1}$$

where $p=b/(b+d)$, d is the linear dimension of the dielectric region, ρ_2 is the resistivity of the metal, and ρ_1 is the resistivity of the dielectric.

Hence it follows that the conductivity is

$$\sigma(\omega,T,p)=\sigma_1(\omega,T)\sigma_2(\omega,T)/[p\sigma_2(\omega,T)+(1-p)\sigma_1(\omega,T)].$$

Allowance for the threshold concentration[7] slightly modifies this formula:

$$\sigma(\omega,T,g)=\sigma_1(\omega,T)\sigma_2(\omega,T)/[g\sigma_2(\omega,T)+(1-g)\sigma_1(\omega,T)], \tag{2}$$

where $g=p/p_{\rm cr}$ and $\sigma_{1,2}=1/\rho_{1,2}$.

Before calculating the dependences $\sigma_1(\omega,T)$ and $\sigma_2(\omega,T)$, we first need to select the charge-transfer mechanism. For a metal the situation is simple: in this case $\sigma_1(\omega,T)$ can be described using a formula that is valid for a gas consisting of almost free electrons.[8] We have

$$\sigma_1(\omega,T)=\mathrm{Re}\{e^2 n\tau_e m^*(1-i\omega\tau_e)\}=e^2 n\tau_e/m^*(1+\omega^2\tau_e^2), \tag{3}$$

where e is the charge of an electron, m^* is its effective mass, n is the electron concentration per unit volume, and τ_e is the electron relaxation time, which can be found either using the kinetic equation or using a diagram technique based on the temperature Green's functions (see Refs. 9–12, for example). It should be noted that if the field frequency is high ($\omega\tau_e \gg 1$), the time τ_e itself begins to depend on the field frequency (for a detailed description of this see the review presented in Ref. 11) and includes three main time-dependent components. That is to say.

$$\tau_e^{-1}=\tau_{ei}^{-1}+\tau_{ee}^{-1}+\tau_{eph}^{-1}, \tag{4}$$

where τ_{ei} is the electron-impurity elastic collision time, τ_{ee} is the electron-electron collision time, which governs the establishment of the temperature T_e in the electron subsystem, and τ_{eph} is the electron-phonon relaxation time, which gov-

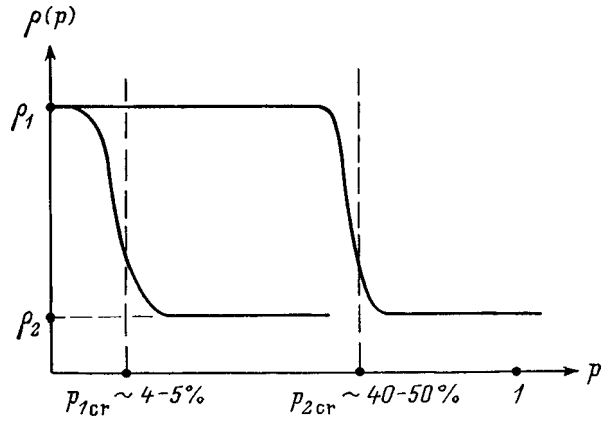

FIG. 1. Resistivity of a composite as a function of the concentration of metal additives for two different types of technologies for preparing the internal composition.

erns the establishment of the equilibrium thermostat temperature T in the electron gas with phonons acting as the thermostat.

The relationships between these three real physical parameters will be used to construct the temperature and frequency dependences of the conductivity of the composite. For the dielectric region we can take it as proven *a priori* that in most dielectrics charge transfer is mediated by a hopping mechanism. This viewpoint allows us to describe the conductivity in the dielectric region using the density-matrix formalism.[11,13]

Let the subscript i denote an impurity state in the band gap. As a result, charge is transferred to real states. The ''dumping'' of an electron into the ε_i level corresponds to an electron transition from the band gap to the ''tunneling'' band. Such a transition can occur as a result of the electron absorbing some particle or quasiparticle. For example, if the ε_i level is near the bottom of the band gap, in order to reach it, it is sufficient for the electron to absorb a phonon. If the level is high, the situation is more complex, and in order to reach this level, the electron must either absorb n_0 phonons (at low temperatures the probability of this process is reduced substantially, but increases at high temperatures, such as room temperature) or it can absorb any high-energy particle, say a photon. Without going into the details of the calculations (which are described in detail in Ref. 14), we can show that in this case the conductivity is determined by analogy with the calculation of the conductivity in a quantizing magnetic field.[9] With allowance for the alternating electric field we have

$$\sigma_{2\alpha\beta}(\omega,T)=\mathrm{Re}\Bigg\{\hbar e^2\sum_i\sum_j v_{\alpha ij}v^*_{\beta ij}[(\langle f_i\rangle-\langle f_j\rangle)/(\varepsilon_i-\varepsilon_j)][(\varepsilon_i-\varepsilon_j+i\hbar\tau_i^{-1}+\hbar\omega)^{-1}+(\varepsilon_i-\varepsilon_j+i\hbar\tau_i^{-1}-\hbar\omega)^{-1}]\Bigg\}. \quad (5)$$

Here $\langle f_{i,j}\rangle=e^{-\varepsilon_i/k_BT}$ is the equilibrium density matrix, $v_{\alpha ij}$ are the matrix elements of the velocity operator, and $1/\tau_i$ is the reciprocal electron lifetime in level i. This lifetime is related to the tunneling parameters by

$$1/\tau_i=\nu_0\begin{cases}e^{-(V_0-\varepsilon_i)/k_BT} & \text{at} \quad T>T^*,\\ e^{-S} & \text{at} \quad T<T^*,\end{cases} \quad (6)$$

where ν_0 is the tunneling frequency, which can be estimated experimentally (as will be discussed briefly below), V_0 is the barrier height, S is the quasiclassical action, and the temperature T^* is determined from

$$e^{-S}=e^{-(V_0-\varepsilon_i)/k_BT}$$

or

$$T^*=(V_0-\varepsilon_i)/k_B. \quad (7)$$

Formula (5) is simplified slightly if we introduce the random ''spread'' of impurity levels Δ, i.e., if we set

$$\varepsilon_j=\varepsilon_i+\Delta. \quad (8)$$

Then, after separating the real part, formula (5) gives

$$\sigma_{2xx}(\omega,T)=\sigma_2(\omega,T)=e^2\sum_i\sum_\Delta v_i(\varepsilon_i+\Delta)v_i^*(\varepsilon_i+\Delta)\times\exp(-\varepsilon_i/k_BT)[1-\exp(-\Delta/k_BT)]\tau_i/\Delta[((\Delta/\hbar)-\omega)^2\tau_i^2+1]^2. \quad (9)$$

Provided the spread of the impurity levels is small and has a maximum of Δ_0, formula (9) should be averaged over Δ. This operation corresponds to replacing the sum by the integral of (9), i.e.,

$$\sigma_2(\omega,T)=\int_{-\infty}^{+\infty}\{\text{formula (9)}\}\,d\Delta/2\Delta_0. \quad (10)$$

After simple integration, we obtain

$$\sigma_2(\omega,T)=(e^2/2\Delta_0)\sum_i[v_{ii}]^2e^{-\varepsilon_i/k_BT}\times\{1-\exp(-\hbar\omega/k_BT)[\tau_i\omega\sin(\hbar/\tau_iT)+\cos(\hbar/\tau_iT)]\}\tau_i/(1+\omega^2\tau_i^2). \quad (11)$$

Finally, assuming that the level ε_i created by each random impurity is the same on the average, i.e., $\varepsilon_i=\varepsilon_0$, and introducing the density of states of these levels $n(\varepsilon_i)=dN(\varepsilon_i)/d\varepsilon_i$, where $N(\varepsilon)$ is the number of i levels, we find

$$\sigma_2(\omega,T)=\mathfrak{M}e^2\langle v\rangle^2\nu e^{-\varepsilon_0/k_BT}\{1-(e^{-\hbar\omega/k_BT})\times[(\omega/\nu)\sin(\hbar\nu/k_BT)+\cos(\hbar\omega/k_BT)]\}, \quad (12)$$

where $\nu=1/\tau_0$, $\langle v\rangle^2=\mathfrak{M}^{-1}\int_0^\infty[v_{ii}]^2n(\varepsilon_i)\,d\varepsilon_i$ and $\mathfrak{M}$ is the number of electrons per unit volume of the dielectric.

Substituting $\sigma_1(\omega,T)$ and $\sigma_2(\omega,T)$ into (2) gives a formula for the total conductivity of the composite $\sigma(\omega,T)$. However, assuming that p-type conductivity also occurs in the general case, i.e., making the formal substitution

$$\sigma_1(\omega,T)\to\sigma_e(\omega,T)+\sigma_f(\omega,T), \quad (13)$$

where

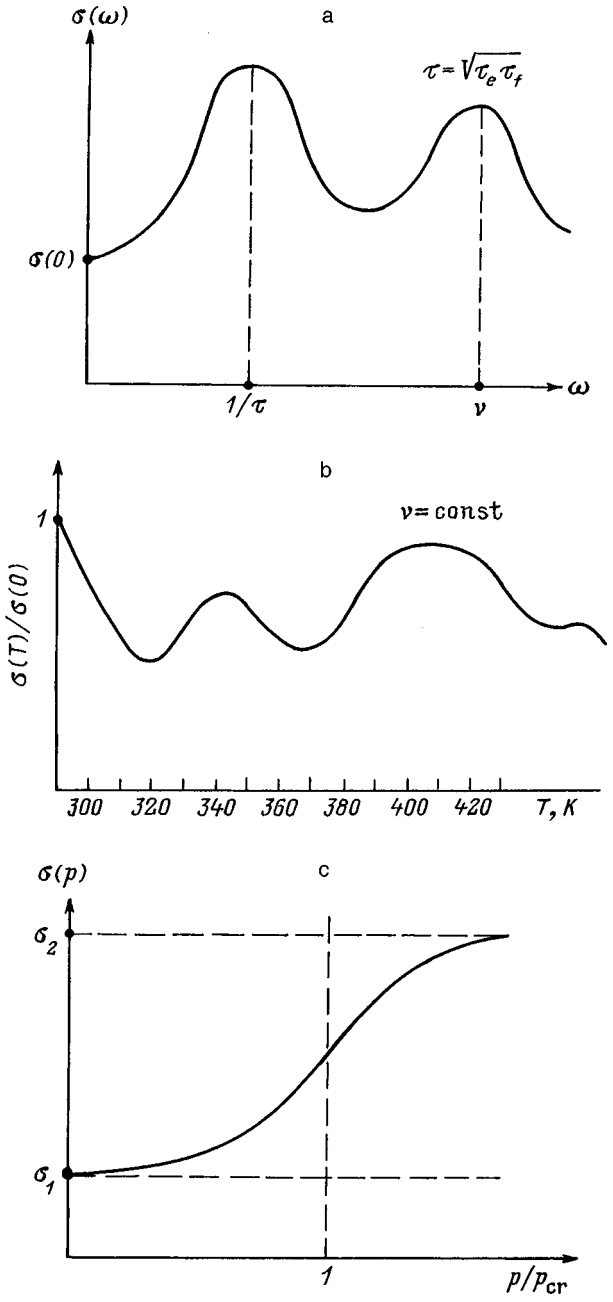

FIG. 2. Schematic behavior of the conductivity as a function of the frequency of the alternating electric field (a), predicted temperature oscillations of the conductivity associated with a pure hopping mechanism of charge transfer (b), and qualitative dependence of the conductivity on the concentration of metal inclusions (c).

$$\sigma_e(\omega,T)=\sigma_{e0}/[1+\omega^2\tau_e^2],\quad \sigma_{e0}=e^2n_e\tau_e/m_e,$$

$$\sigma_f(\omega,T)=\sigma_{f0}/[1+\omega^2\tau_f^2],\quad \sigma_{f0}=e^2n_f\tau_f/m_f,$$

$$\tau_f^{-1}=\tau_{ff}^{-1}+\tau_{fph}^{-1},$$

τ_{ff} is the hole collision time, τ_{fph} is the hole-phonon time, m_e and m_f are the electron and hole masses, respectively, τ_e and τ_f are the electron and hole relaxation times, and n_e and n_f are the electron and hole concentrations per unit volume, we find the exact expression for the conductivity

$$\sigma(\omega,T)=\sigma_{20}[\sigma_{e0}+\sigma_{f0}+\omega^2(\sigma_{e0}\tau_f^2+\sigma_{f0}\tau_e^2)]\times\{g\sigma_{20}(1+\omega^2\tau_e^2)(1+\omega^2\tau_f^2)+(1-g)\times(1+\omega^2/\nu^2)[\sigma_{e0}+\sigma_{f0}+\omega^2(\sigma_{e0}\tau_f^2+\sigma_{f0}\tau_e^2)]\}^{-1},\tag{14}$$

where

$$\sigma_{20}=[\mathfrak{M}e^2\langle v\rangle^2e^{-\varepsilon_0/k_BT}/2\Delta_0\nu]\{1-(\omega/\nu)\sin(\hbar\nu/k_BT)+\cos(\hbar\nu/k_BT)e^{-\hbar\omega/k_BT}\}.\tag{15}$$

It should be specially noted that the temperature dependence of the effective conductivity (or, more precisely, the total conductivity) $\sigma(\omega,T)$ exhibits clearly expressed oscillatory behavior [see expression (15)]. Provided $\hbar\omega\ll k_BT$ and $\omega\ll\nu$, from Eq. (15) we obtain

$$\sigma_{20}(\nu)=[\mathfrak{M}e^2\langle v\rangle^2e^{-\varepsilon_0/k_BT}/2\Delta_0\nu](1-\cos(\hbar\nu/k_BT)).\tag{16}$$

This shows, in particular, that if the tunneling frequency $\nu\to0$, σ_{20} also tends to zero, and the entire electrical conduction mechanism is attributable to the electron-hole mechanism alone. It follows from expression (16) that when $T\to\hbar\nu/k_B$, this formula describes pronounced temperature oscillations of the electrical conductivity. It should be noted that these oscillations are displayed most clearly at low (liquid-helium) temperatures, although they also occur at high temperatures if the Arrhenius above-barrier transition is very efficient but their period then becomes considerably larger than at liquid-helium temperatures.

In order to now estimate the temperature dependence of the conductivity of the composite, we should present explicit expressions for the relaxation times. From Ref. 10, for example (see also Ref. 11), we have: for $T<\theta_D\ll\varepsilon_F$

$$\tau_{ee}^{-1}=[\varepsilon_F/m_e^*\langle a\rangle^2]^{1/2}(T/\varepsilon_F)^2,$$

$$\tau_{ff}^{-1}=[\varepsilon_F/m_f^*\langle a\rangle^2]^{1/2}(T/\varepsilon_F)^2,$$

$$\tau_{fph}^{-1}=\tau_{eph}^{-1}=T^3/\hbar\theta_D^2,\tag{17}$$

for $\theta_D<T\ll\varepsilon_F$

$$\tau_{ee}^{-1}=[\varepsilon_F/m_e^*\langle a\rangle^2]^{1/2}(T/\varepsilon_F)^2,$$

$$\tau_{ff}^{-1}=[\varepsilon_F/m_f^*\langle a\rangle^2]^{1/2}(T/\varepsilon_F)^2,$$

$$\tau_{fph}^{-1}=\tau_{eph}^{-1}=\zeta T/\hbar,\tag{18}$$

where ε_F is the Fermi energy, θ_D is the Debye temperature, $\langle a\rangle$ is the average interatomic distance, and ζ is a constant considerably smaller than unity.

Thus, knowing the temperature dependence of the electron and hole relaxation times, we can use formula (14) to determine the functional dependences of the conductivity of complex composites as a function of the frequency of the applied electric field and the concentration of conducting particles over the entire range of temperatures attainable in practice.

It should be noted that if the electric field frequency is high, effects associated with the nonuniformity of the electric field in the composite begin to play a significant role. In this

case, allowance must be made for the interaction of electrons and holes with photons by introducing quantization of the electric field. However, this problem is beyond the scope of the present study since it was tacitly implied above that frequencies $\omega < 10^{10}$ Hz were being considered.

The behavior of the conductivity of a composite as a function of the parameters of formula (14) is shown schematically in Fig. 2.

To conclude, we give the main results of this study: 1) a theory has been devised for the conductivity of a two-component composite with conducting inclusions in the form of a finely dispersed metal phase; 2) interesting features of the temperature behavior of σ and, in particular, oscillating dependences associated with a pure hopping mechanism of charge transfer have been predicted; 3) the dependence of the conductivity on the concentration of metal additives has been determined.

[1] S. O. Gladkov, Physica B **167**, 159 (1990).
[2] E. Burzo and R. Teteanu, Solid State Commun. **86**, 493 (1993).
[3] I. B. Bergmann, Phys. Pol. **45**, 76 (1995).
[4] S. O. Gladkov, Fiz. Tverd. Tela. (St. Petersburg) **39**, 1622 (1997) [Phys. Solid State **39**, 1446 (1997)].
[5] M. S. Bresler, R. V. Parfen'ev, and S. S. Shalyt, Fiz. Tverd. Tela. (Leningrad) **8**, 1776 (1966) [Phys. Solid State **8**, 1414 (1966)].
[6] S. M. Piccard and B. Derby, Acta Metall. Mater. **38**(12), 239 (1990).
[7] B. I. Shklovskiĭ and A. L. Efros, *Percolation Theory* [in Russian], Nauka, Moscow (1982), 314 pp.
[8] A. A. Abrikosov, *Introduction to the Theory of Normal Metals (Solid State Physics, Suppl. 12)* (Academic Press, New York, 1972) [Russ. original, Nauka, Moscow 1969, 237 pp.].
[9] I. M. Lifshits, M. Ya. Azbel', and M. I. Kaganov, *Electron Theory of Metals* (Consultants Bureau, New York, 1973) [Russ. original, Nauka, Moscow, 1971, 415 pp.].
[10] E. M. Lifshits and L. P. Pitaevskiĭ, *Physical Kinetics* (Pergamon Press, Oxford, 1981) [Russ. original, Nauka, Moscow, 1979, 527 pp.].
[11] S. O. Gladkov, Phys. Rep. **182**(4–5), 211 (1989).
[12] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, N.J., 1963) [Russ. original, Fizmatgiz, Moscow, 1962, 443 pp.].
[13] I. V. Aleksandrov, *Theory of Magnetic Relaxation* [in Russian], Nauka, Moscow (1975), 399 pp.
[14] S. O. Gladkov, Chem. Phys. Lett. **174**, 636 (1990).

Translated by R. M. Durham

Dielectric properties of $(1-x)[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3] \cdot xSrTiO_3$ solid solutions in the vicinity of phase transitions

L. N. Korotkov, S. P. Rogova, and N. G. Pavlova

Voronezh State Technological Academy, 394053 Voronezh, Russia
(Submitted December 15, 1997)
Zh. Tekh. Fiz. **69**, 35–38 (March 1999)

The dielectric properties of polycrystalline $(1-x)[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3] \cdot xSrTiO_3$ solid solutions, where $x=0-0.7$, are studied in the temperature range 150–600 K. Systematic spreading of the ferroelectric phase transition with increasing strontium titanate content is discovered. The dispersion of the dielectric constant at frequencies from 10^{-1} to 10^6 Hz is investigated for a composition with $x \simeq 0.7$. The existence of two relaxation processes characterized by diffuse relaxation time spectra, which broaden with decreasing temperature, is established. It is postulated that a transition to a glasslike state takes place in the material with $x \simeq 0.7$. © *1999 American Institute of Physics.* [S1063-7842(99)00603-0]

Numerous ferroelectrics with a diffuse phase transition, or relaxors,[1] many of which have found extensive application in technology, are presently known. The factors leading to spreading of the ferroelectric phase transition have been debated in the literature,[2–6] and the following approaches to this problem can be singled out. The theory which allows for the coexistence of polar and nonpolar phases over a broad temperature range to account for the diffuse phase transition[3] is most widely accepted. The possible factors which permit such coexistence include spatially inhomogeneous compositional fluctuations,[1,3] slow glasslike dipole librations,[4] the presence of microregions with different degrees of ordering of the ions,[5] and inhomogeneous compositional ordering.[6]

Thus, the study of the factors leading to the relaxor ''behavior'' of ferroelectric materials is of crucial importance. To resolve this question it would be useful to extend the list of objects which exhibit relaxor properties. Systems which display systematic spreading of the ferroelectric phase transition as the composition of the solid solution is varied are of greatest interest to investigators. They include, for example, systems of the $xK_{0.5}Bi_{0.5}TiO_3 \cdot (1-x)PbZrO_3$ (Ref. 7) and $K_{0.5}Bi_{0.5}TiO_3 \cdot SrTiO_3$ (Ref. 8) types, which exhibit, in particular, a significant dependence of the degree of spreading of the phase transition on the concentration of the components.

These solid solutions contain $K_{0.5}Bi_{0.5}TiO_3$ as one component of a ferroelectric with a diffuse phase transition, while the other component is the antiferroelectric $PbZrO_3$ or the virtual ferroelectric $SrTiO_3$, respectively.[9] A decrease in the spreading of the phase transition with increasing concentration of the second component is observed in both cases.[7,8]

It would be interesting to analyze the influence of the composition on the spreading of the phase transition in solid solutions containing all three of the components just enumerated, i.e., $K_{0.5}Bi_{0.5}TiO_3$-$PbZrO_3$-$SrTiO_3$ solid solutions. For this purpose, we synthesized a series of $(1-x)[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3] \cdot xSrTiO_3$ solid solutions, which were obtained using a two-step ceramic technology[10] from the high-purity oxides of bismuth, titanium, strontium, zirconium, and lead and potassium carbonate. The synthesis temperature was 900 °C, and the synthesis time was 3 h. The sintering of the samples was carried out 1200 °C for 1 h.

X-ray powder diffraction analysis, which was performed using a DRON-1 diffractometer, demonstrated the formation of single-phase solid solutions. According to the data from an emission-spectroscopic analysis, the concentration of impurities in the materials synthesized did not exceed 10^{-3} wt %.

The samples were metallized in the form of disks of diameter 10 mm and thickness 1 mm after preliminary polishing by burning a turpentine silver paste at 800 °C for 30 min.

The samples investigated were placed in a thermostat, where the temperature was varied during the experiment from 150 to 600 K and monitored using a differential Alumel-Chromel thermocouple to within ± 1 K.

The temperature dependences of the real $[\varepsilon'(T)]$ and imaginary $[\varepsilon''(T)]$ components of the complex dielectric constant ε^* were obtained using an R-571 bridge at $f=30-10\,000$ Hz and an E7–12 bridge at 1 MHz. Measurements were performed at 0.1 Hz using the apparatus described in Ref. 11. All the measurements were performed in a regime with heating (cooling) at a rate of 1–2 K/min.

The results of the measurements of the temperature dependence of ε' for compositions with different values of x are presented in Fig. 1. It is seen that the maxima on the $\varepsilon'(T)$ curves shift toward lower temperatures and that they decrease in absolute value and spread as the strontium titanate content is increased. These changes are apparently due to the significant spreading of the phase transition with increasing x. The degree of spreading of the phase transition is characterized quantitatively by the spreading parameter G (Ref. 3), which can be determined from the temperature dependence of ε'. According to Ref. 3, in the case of a diffuse ferroelectric phase transition the dependence of ε' on T in the paraelectric phase obeys the law

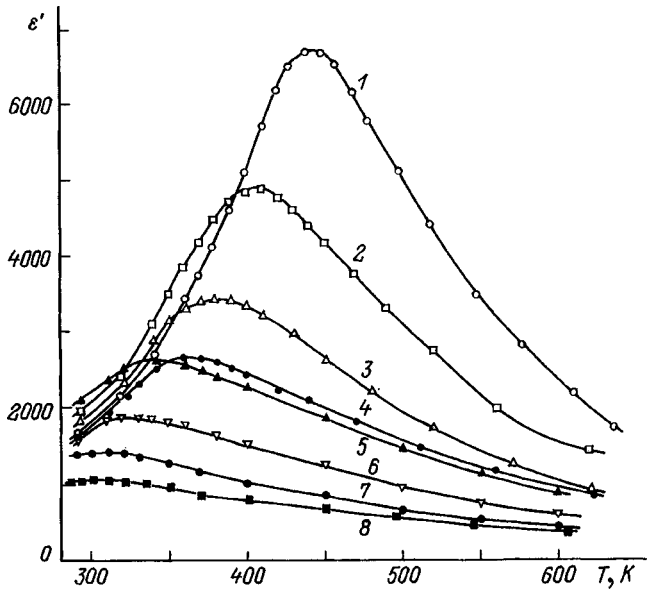

FIG. 1. Temperature dependences of ε' obtained at 10 kHz during heating for $(1-x)[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3] \cdot xSrTiO_3$ solid solutions with $x=0$ (*1*), 0.12 (*2*), 0.2 (*3*), 0.3 (*4*), 0.4 (*5*), 0.5 (*6*), 0.6 (*7*), and 0.7 (*8*).

$$\varepsilon'^{-1} = \varepsilon_m^{-1} + (2\varepsilon_m^{-1} G^{\nu})^{-1} (T - T_m)^{\nu}, \tag{1}$$

where ε_m is the value of ε' at the maximum observed at $T=T_m$, and $\nu=2$.

An analysis of the $\varepsilon'(T)$ curves obtained under the conditions of our experiment shows that $\nu=1.7\pm0.10$ for the compositions investigated and does not depend appreciably on x. At the same time, the spreading parameter G increases significantly as the concentration of $SrTiO_3$ is increased. This can be inferred from Fig. 2, which shows plots of $G(x)$, as well as of $T_m(x)$.

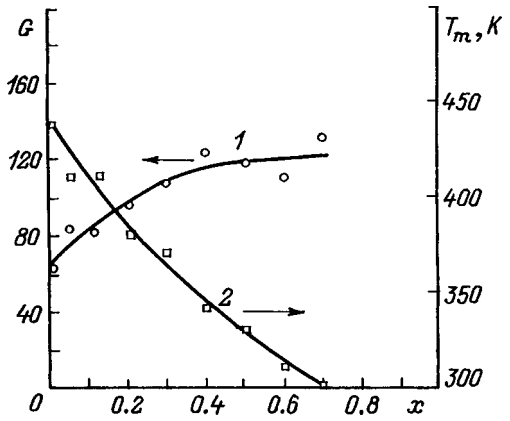

FIG. 2. Concentration dependences of G (*1*) and T_m (*2*).

At the same time, as follows from the data published in Ref. 8, in the case of the ''closely related'' $K_{0.5}Bi_{0.5}TiO_3$-$SrTiO_3$ system there is a decrease in the spreading of the ε' maximum with increasing $SrTiO_3$ concentration, which contrasts sharply with the results presented above. The differences observed can presumably be explained in the following manner.

According to Ref. 12, in some solid solutions the significant spreading of the ferroelectric phase transition due to compositional fluctuations may not occur if the polar phase formed as a result of the phase transition is characterized by high values of the spontaneous polarization and/or the spon-

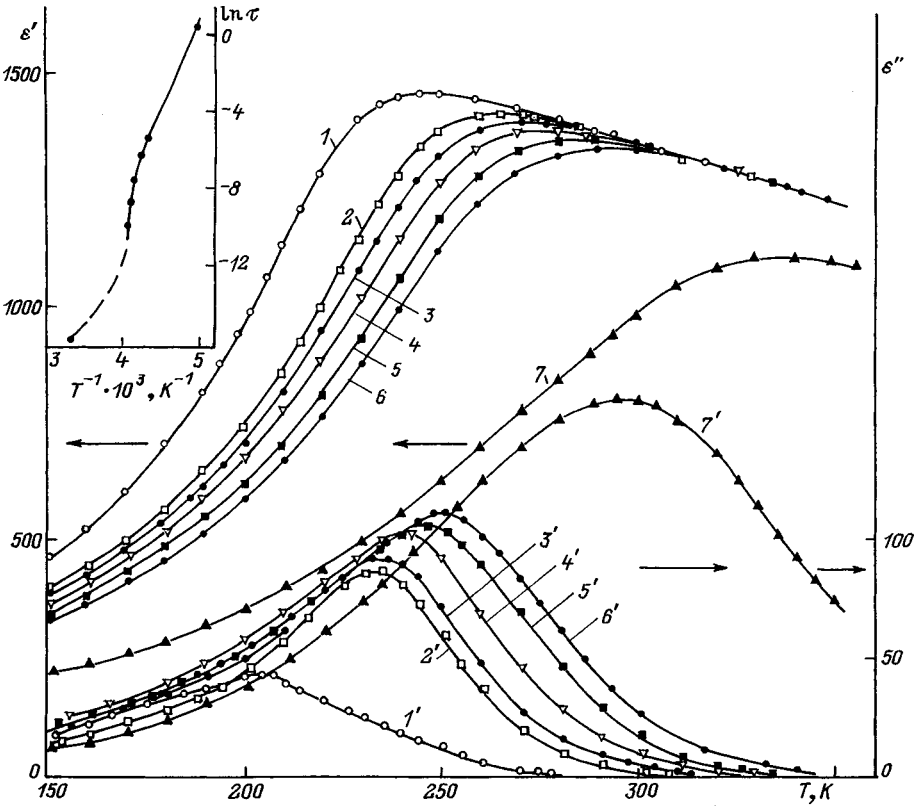

FIG. 3. Temperature dependences of ε' (*1–7*) and ε'' (*1′–7′*) measured at 0.1 (*1*, *1′*), 30 (*2*, *2′*), 80 (3, *3′*), 300 (*4*, *4′*), 1000 (*5*, *5′*), 3000 (*6*, *6′*), and 1 000 000 (*7*, *7′*) Hz for a composition with $x\simeq0.7$. Inset — dependence of ln τ on $1/T$.

taneous strain. This has been observed, for example, for $Na_{0.5}Bi_{0.5}TiO_3$-$PbTiO_3$ solid solutions, in which the spontaneous strain and the spontaneous polarization increase sharply with increasing concentration of the second component.[12]

A similar situation probably occurs in the case of the $K_{0.5}Bi_{0.5}TiO_3$-$PbZrO_3$ (Ref. 7) and $K_{0.5}Bi_{0.5}TiO_3$-$SrTiO_3$ (Ref. 8) systems. In the $[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3]$-$SrTiO_3$ ternary system investigated in the present work the replacement of Pb^{2+} by Sr^{2+} probably leads to decreases in the values of the spontaneous polarization and the spontaneous strain with a resultant increase in the spreading of the phase transition. We note that the increase in the spreading of the phase transition observed with increasing $SrTiO_3$ content is consistent with the arguments advanced in Ref. 13 regarding gradual ''classification,'' i.e., the gradual transition in a solid solution from a ferroelectric with relatively weak spreading of the phase transition to a ferroelectric with a significantly diffuse phase transition and then to dipole glasses. In fact, the $\varepsilon'(T)$ curves obtained for the compositions with $x \simeq 0.6$ and 0.7 in the vicinity of T_m have the form characteristic of dipole glasses, which appear, for example, in mixed crystals of the ferro-antiferroelectric potassium ammonium dihydrogen phosphate.[14] Significant dispersion of both components of the dielectric constant is observed in the vicinity of T_m (Fig. 3). In addition, there is appreciable displacement of the ε' maximum with frequency. For example, for the solid solution with $x \simeq 0.7$, $T_m \simeq 245$ and 335 K in the measurements at 10^{-1} and 10^6 Hz, respectively.

The observed dielectric relaxation process is thermally activated. However, the temperature dependence of the characteristic time τ of the relaxation process cannot be described by the Arrhenius law

$$\tau = \tau_0 e^{U/kT}, \tag{2}$$

where τ_0 is the preexponential factor, U is the activation energy, and k is the Boltzmann constant.

This is illustrated in the inset in Fig. 3, which shows the dependence of $\ln \tau$ on T^{-1} [the experimental points correspond to the temperatures at which the maximum is observed on the $\varepsilon''(T)$ curves for the frequency $f = 1/2\pi\tau$].

A preliminary analysis of the dispersion of ε^* was performed using the Cole–Cole equation[15]

$$\varepsilon^*(\omega) = \varepsilon_\infty + \frac{\varepsilon_S - \varepsilon_\infty}{1 + i(\omega\tau)^{1-\alpha}}, \tag{3}$$

where $\omega = 2\pi f$, ε_S and ε_∞ are the values of the dielectric constant for $\omega \Rightarrow 0$ and $\omega \Rightarrow \infty$, respectively, and α is the relaxation time distribution parameter.

It was discovered that the observed dispersion of the dielectric constant in the vicinity of T_m is caused by at least two relaxation processes, each of which is characterized by a diffuse spectrum of relaxation times, which broadens significantly as the temperature is lowered. This is revealed by the Cole–Cole plots (Fig. 4), which can be used to estimate the distribution parameters $\alpha_1 = 2\varphi_1/\pi$ and $\alpha_2 = 2\varphi_2/\pi$, where the angles φ_1 and φ_2 are determined as is shown in Fig. 4.

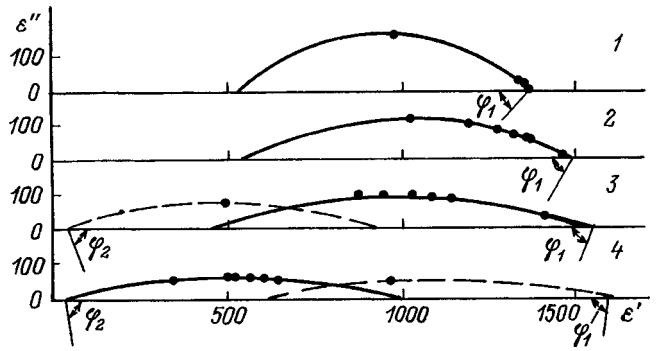

FIG. 4. Cole–Cole plots for a composition with $x \simeq 0.7$ at 300 (*1*), 250 (*2*), 230 (*3*), and 200 K (*4*).

At the same time, the Cole–Cole equation satisfactorily describes the experimental data obtained for the temperature range ~200–300 K (Fig. 4, curves *1–4*). Below ~200 K the dispersion of ε^* cannot be described by the relation (3), since ε'' scarcely depends on ω in that range. The latter finding indicates that the relaxation time distribution density becomes essentially uniform in a range significantly exceeding $f_1^{-1} - f_2^{-1}$, where $f_1 = 10^{-1}$ Hz and $f_2 = 10^6$ Hz are, respectively, the lowest and highest measurement frequencies. A similar effect is observed for both dipole[14] and spin glasses.[16]

It is noteworthy that significant thermal hysteresis of the dielectric constant was observed for $[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3]$ in the vicinity of T_m (Ref. 7). In the case of $0.3[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3] \cdot 0.7SrTiO_3$ only very slight hysteresis of ε' is observed under similar conditions. The discrepancies between the values of ε' measured during heating and cooling did not exceed 5%. Therefore, it can be presumed that the material containing 70% $SrTiO_3$ does not undergo a structural phase transition in the vicinity of T_m.

Thus, it can be concluded on the basis of the body of experimental results that systematic spreading of the ferroelectric phase transition is observed in $(1-x)[0.7PbZrO_3 \cdot 0.3K_{0.5}Bi_{0.5}TiO_3] \cdot xSrTiO_3$ solid solutions as the concentration of strontium titanate is increased. In the case of the composition with $x \simeq 0.7$ there is apparently no structural phase transition, but a transition to a glasslike state does occur. This is indicated by the dispersion of ε^*, which is characteristic of glasslike systems,[14,16] as well as by the essential disappearance of the thermal hysteresis of the dielectric constant.

We thank S. A. Gridnev for following this research and participating in the discussion of the results.

This work was performed with support from the Russian Fund for Fundamental Research (Project No. 96-02-17131).

[1] V. A. Isupov, Izv. Akad. Nauk SSSR, Ser. Fiz. **47**, 559 (1983) [Bull. Acad. Sci. USSR, Phys. Ser. **47**, 136 (1983)].
[2] *Abstracts of the International Seminar on Relaxor Ferroelectrics*, Dubna (1996), 102 pp.
[3] G. A. Smolenskiĭ, V. A. Bokov, and V. A. Isupov *et al.*, *Ferroelectrics and Antiferroelectrics* [in Russian], Nauka, Leningrad (1971), 476 pp.
[4] G. Schmidt, Ferroelectrics **78**, 199 (1988).
[5] C. A. Randall, Jpn. J. Appl. Phys. **29**, 327 (1990).

[6] A. A. Bokov, M. A. Malitskaya, I. P. Raevskiĭ, and V. Yu. Shonov, Fiz. Tverd. Tela (Leningrad) **32**, 2488 (1990) [Sov. Phys. Solid State **32**, 1445 (1990)].
[7] S. A. Gridnev and S. V. Popov, Izv. Ross. Akad. Nauk, Ser. Fiz. **61**, 232 (1997).
[8] S. M. Emel'yanov, I. P. Raevskiĭ, and O. I. Prokopalo, Fiz. Tverd. Tela (Leningrad) **25**, 1542 (1983) [Sov. Phys. Solid State **25**, 889 (1983)].
[9] Yu. N. Venevtsev, E. D. Politova, and S. A. Shipov, *Barium Titanate Ferro- and Antiferroelectrics* [in Russian], Khimiya, Moscow (1985), 256 pp.
[10] S. A. Gridnev, N. G. Pavlova, S. P. Rogova, and L. N. Korotkov, USSR Inventor's Certificate (Patent) No. 1,682,352.
[11] S. A. Gridnev, L. N. Korotkov, and L. A. Shuvalov, in *Relaxation Processes in Dielectrics* [in Russian], VPI, Voronezh (1990), pp. 4–10.
[12] V. A. Isupov, P. L. Strelets, and I. A. Serova *et al.*, Fiz. Tverd. Tela (Leningrad) **6**, 790 (1964) [Sov. Phys. Solid State **6**, 615 (1964)].
[13] V. A. Isupov, Izv. Akad. Nauk SSSR, Ser. Fiz. **54**, 1131 (1990).
[14] S. A. Gridnev, L. N. Korotkov, and L. A. Shuvalov, Ferroelectrics **167**, 99 (1995).
[15] Yu. M. Poplavko, *The Physics of Dielectrics* [in Russian], KPI, Kiev (1972), 313 pp.
[16] S. L. Ginzburg, *Irreversible Phenomena in Spin Glasses* [in Russian], Nauka, Moscow (1989), 152 pp.

Translated by P. Shelnitz

Iterative solution of the inverse problem of dynamic diffraction in inhomogeneous crystals

S. G. Podorov and V. I. Punegov

Syktyvkar State University, 167001 Syktyvkar, Russia
(Submitted May 28, 1997)
Zh. Tekh. Fiz. **69**, 39–42 (March 1999)

An iterative method for solving the inverse problem of dynamic x-ray diffraction in crystalline layers that are inhomogeneous across their thickness is developed. The structural characteristics of an InGaAsSb/AlGaAsSb/GaSb heteroepitaxial system are calculated using the method. © *1999 American Institute of Physics.* [S1063-7842(99)00703-5]

1. INTRODUCTION

One of the central problems in solid-state physics is the investigation of the spatial distributions of structural characteristics of crystals using x-ray diffraction. This problem has taken on crucial significance in the last 20–25 years in connection with the development of new semiconductor devices. Despite the large number of publications devoted to this subject (see, for example, Refs. 1–10), a generally accepted, universal method for solving the inverse problem of x-ray diffraction has still not been developed. Therefore, the search continues for new avenues for progress in this area of solid-state physics.

In Ref. 8 we developed a method for solving the inverse problem within the kinematic approximation. Alongside some other approaches, this method has been used to obtain the spatial structural characteristics of a thin graded single-crystal AlGaAs film on a thick GaAs substrate.[10] However, this method is not applicable to relatively thick heteroepitaxial systems, since the kinematic (Born) approximation is valid only for thin layers. Therefore, the purpose of the present work is to develop a more general approach to the solution of the inverse problem within an iterative procedure.

2. BASIC EQUATIONS

With no loss of generality, let us consider the symmetric Bragg diffraction of x rays from a crystalline layer of thickness l, which is inhomogeneous with respect to the depth z and lies on a thick ideal substrate. Let a monochromatic x-ray plane wave impinge on the surface of the crystal under investigation at an angle θ. The diffraction of x rays in a one-dimensionally distorted crystal is described by the Takagi–Taupin equations

$$\frac{d}{dz}E_0(z)=i\sigma_0E_0(z)+i\sigma_{-g}U^*(z)E_g(z),$$

$$-\frac{d}{dz}E_g(z)=i(\eta-\sigma_0)E_g(z)+i\sigma_gU(z)E_0(z), \qquad (1)$$

where $E_0(z)$ and $E_g(z)$ are the amplitudes of the transmitted and reflected waves.

The parameters in (1) have the following form in the generally accepted notations:[7]

$$\sigma_0=\pi\chi_0/(\lambda\gamma_0); \quad \sigma_{g,-g}=\pi\chi_{g,-g}C/(\lambda|\gamma_{g,0}|);$$

$$\eta=2\pi(\chi_0+\sin(2\theta_0)\Delta\theta)/(\lambda\gamma_0).$$

Here $\mathbf{g}$ is the diffraction vector, and $\Delta\theta=\theta-\theta_0$ is the angular mismatch. The substrate Bragg angle is usually taken as the reference angle θ_0. Distortions of the crystal lattice structure in the system (1) are assigned by the function $U(z)$

$$U(z)=B(z)\exp(-iF(z)),$$

$$F(z)=\mathbf{g}\cdot\mathbf{u}(z)=\int_0^z f(x)\,dx,$$

$$f(z)=-2\pi\Delta d(z)/d^2. \qquad (2)$$

Thus, within the problem under consideration, structural distortions in the test object are specified by the crystal-lattice strain field $\Delta d(z)/d$ and by the presence of defects, whose type, concentration, and dimensions influence the values of the static Debye–Waller factor

$$B(z)=e^{-W(z)}.$$

The boundary conditions for the Bragg case are $E_0(0)=1$ and $E_g(l)=R_s(\eta)$, where $R_s(\eta)$ is the amplitude reflection coefficient of the substrate. Using a known procedure we can reduce the Takagi equations (1) to a nonlinear equation of the Riccati type (Taupin's equation)[11]

$$\frac{d}{dz}R_g(z,\eta)=-i\sigma_gE(z)-i(\eta-f)R_g(z,\eta)$$

$$-i\sigma_{-g}E(z)R_g^2(z,\eta) \qquad (3)$$

with the boundary condition $R_g(z,\eta)|_{z=L}=R_s(\eta)$. Equation (3) underlies the algorithm for solving the problem posed. Here $R_g(z=0,\eta)$ is the amplitude reflection coefficient of the inhomogeneous crystal under investigation.

3. ITERATIVE PROCEDURE FOR SOLVING THE INVERSE PROBLEM

Let us construct the function $R_k[U](\eta)=i\sigma_g\int_0^l e^{i\eta z}U(z)\,dz$, which describes the amplitude reflection coefficient of a kinematic layer of thickness l and is a solution of Eq. (3) when the condition $\sigma_{-g}=0$ is satisfied. Let R_g be the reflection coefficient of an inhomogeneous

layer in the general case of dynamic diffraction. The procedure for finding the function $U(z)$ will be regarded as the solution of the nonlinear functional equation

$$(R_g[U](\eta))R_g[U](\eta)^* = I_g(\eta), \tag{4}$$

where the asterisk denotes the complex conjugate.

The function $U(z)$ is the unknown parameter in this equation. Let us define the operator $A(U)V$:

$$A(U)V = \{R_g[U]\}^*\{aR_k[V] + R_g[U] - aR_k[U]\}, \tag{5}$$

where a is a parameter.

Equation (4) can then be written in the following form:

$$A(U)U = I_g(\eta), \tag{6}$$

whose solution is also the solution of the inverse problem of x-ray diffraction in an irregular crystalline layer. The following iterative method is proposed for solving the functional equation (6):

$$A(U^{(n+1)})U^{(n)} = I_g(\eta).$$

Hence we have

$$U^{(n+1)}(z) = A^{-1}(U^{(n)})I_g(\eta)$$

or

$$U^{(n+1)}(z) = U^{(n)}(z) + \frac{1}{2\pi a i \sigma_g} \times \int_{\Delta\Omega} \frac{I_g - I_g[U^{(n)}]}{I_g[U^{(n)}]} R_g[U^{(n)}] e^{-i\eta z}\, d\eta. \tag{7}$$

The iteration formula (7) is the basic relation for solving the inverse problem of dynamic diffraction in a one-dimensionally distorted crystal.

4. CALCULATION OF THE STATIC DEBYE–WALLER FACTOR AND THE STRAIN PROFILE

From the relation (7) we can find the iterative solution for the static Debye–Waller factor and the strain profile of the structure under investigation. To this end, we perform Fourier transformation in (7) and express the solution in terms of $I_g(\eta)$, $R_g[U^{(n)}]$, and $R_k[U^{(n)}]$. For the static Debye–Waller factor we obtain

$$B^{(n+1)}(z) = \left| \frac{1}{2\pi\sigma_g} \int_{\Delta\Omega} [\{I_g - I_g[U^{(n)}]\} R_g[U^{(n)}] / (aI_g[U^{(n)}]) + R_k[U^{(n)}]] e^{-i\eta z}\, d\eta \right|. \tag{8}$$

The lattice strain profile is found from the iteration formula

$$f^{(n+1)}(z) = \frac{d}{dz} F^{(n+1)}(z) = \mathrm{Re}\left\{ i\left(\frac{dU^{(n+1)}(z)}{dz} \Big/ U^{(n+1)} \right) \right\}, \tag{9}$$

where the derivative has the form

$$\frac{dU^{(n+1)}(z)}{dz} = \frac{-1}{2\pi\sigma_g} \int_{\Delta\Omega} \eta\{R_k[U^{(n)}] + (I_g - I_g[U^{(n)}])R_g[U^{(n)}] / (aI_g[U^{(n)}])\} e^{-i\eta z}\, d\eta. \tag{10}$$

The use of a model of a crystal without defects, for which $B^{(0)} = 1$, as the initial approximation for calculations based on Eq. (8) is proposed. Strictly speaking, Eqs. (8) and (9) are approximate, since the angular integration interval $\Delta\Omega$ is bounded. Because of this constraint the Fourier transformation cannot be considered rigorous. This, in turn, causes the calculated profiles to have an oscillatory character. All this affects the convergence of the iterative procedure. Improvement of the convergence is possible if regularization methods are employed.

5. REGULARIZATION PROCEDURES

The iterative algorithm can be regularized using the convolution of the solutions sought for (8) and (9) with Gauss's function

$$\langle B(z)\rangle = \int_{\Delta x} \exp\left(-\frac{(z-x)^2}{\gamma} \right) B(x)\, dx, \tag{11}$$

$$\langle f(z)\rangle = \int_{\Delta x} \exp\left(-\frac{(z-x)^2}{\gamma} \right) f(x)\, dx. \tag{12}$$

The values of the coefficients γ in (11) and (12) were adjusted so as to smooth the oscillatory behavior of $B(z)$ and $f(z)$ due to the constraints on the angular integration interval in (7)–(10). We note that in calculating the angular distribution of the diffraction intensity the σ and π polarizations of the waves were taken into account and that the convolution with the instrumental function of the monochromator was taken.

6. DETERMINATION OF THE STRUCTURAL CHARACTERISTICS OF AN INHOMOGENEOUS $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}/Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$/(001) GaSb HETEROEPITAXIAL SYSTEM

The computational iterative procedure developed is applicable to the structural diagnostics of an inhomogeneous $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}/Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$/(001) GaSb semiconductor heterostructure grown by liquid-phase epitaxy. An $Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$ layer with a thickness of roughly 3.3 μm was grown on a cleaned surface of a (001) GaSb substrate at 600 °C. Another layer with the composition $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}$, whose thickness was tentatively of the order of 1μm according to the epitaxial technology, was created above that layer.

The measurements of the angular distribution of the diffracted radiation were performed on a Topo high-resolution double-crystal diffractometer from the Japanese manufacturer Rigaku. The double-crystal camera was combined with an RU-200 x-ray tube. The primary beam was collimated and monochromatized using the (440) reflection from a perfect germanium(001) crystal in Cu $K_{\alpha 1}$ radiation. The asymmetry factor had the value $b = 0.095$, which ensured an an-

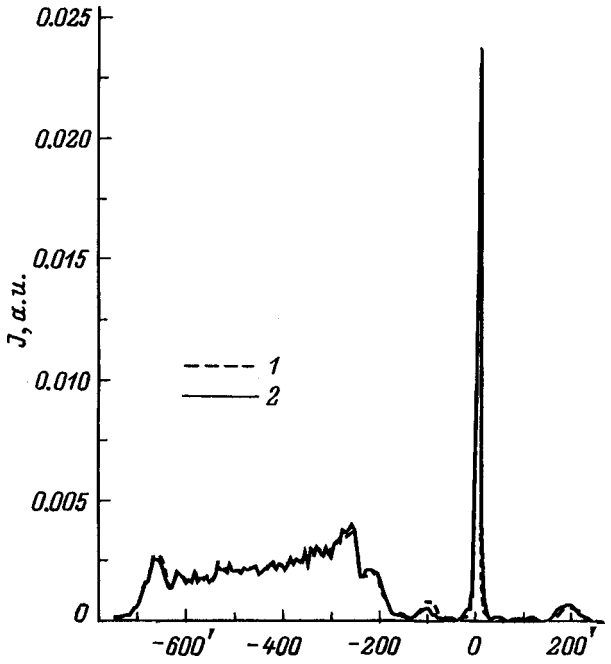

FIG. 1. Calculated (*1*, dashed line) and experimental (*2*, solid line) diffraction reflection curves of an $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}/Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$/(001) GaSb heterostructure.

gular divergence of the primary beam smaller than 1′. The x-ray diffraction photograph of the heteroepitaxial structure was taken at the symmetric (006) reflection with a Bragg angle equal to 49.46°.

The diffraction reflection curve has a form which is typical of inhomogeneous epitaxial structures with a positive lattice strain gradient (Fig. 1). The behavior of the oscillations on the experimental diffraction reflection curve points to the presence of a constant or nearly constant strain gradient in the $Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$ layer. Linear distribution of the components of the solid solution across the thickness of this layer apparently took place during epitaxial growth. The mean interplanar distance of the reflecting atomic planes in the upper $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}$ layer is smaller than the corresponding value for the GaSb substrate. Therefore, the diffraction peak from the $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}$ layer is in the large-angle region at a distance of 200″ from the substrate peak.

The initial computed diagnostics approximation was obtained using a procedure for a model of a crystal with linear variation of the interplanar distance across its thickness.[12] The static Debye–Waller factor was set equal to unity.

The results of the calculations performed by the method described above are shown in Figs. 1–3. In Fig. 1 the theoretical diffraction reflection curve is represented by the dashed line. As expected, the $Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$ epitaxial layer has an essentially constant crystal-lattice strain gradient (Fig. 2). The upper $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}$ layer is also inhomogeneous, its thickness being less than 1 μm. It should also be noted that the upper layer and the graded $Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$ layer are separated by a transition region with a thickness of the order of 0.3 μm, which is formed either during the growth process or as a result of

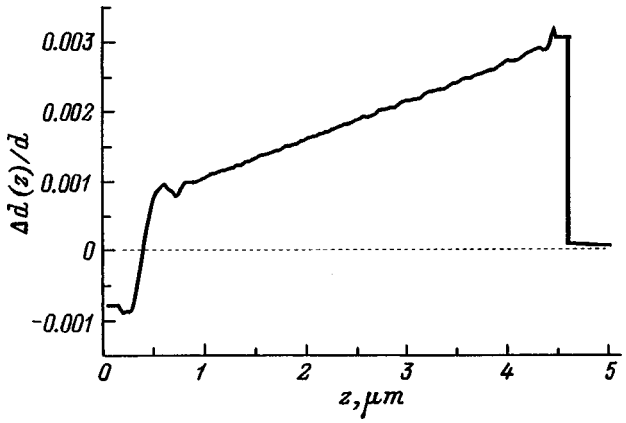

FIG. 2. Strain profile $[\Delta d(z)/d]$ of an $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}/Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$/(001) GaSb heterostructure.

self-diffusion. The boundary between the $Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$ layer and the substrate is fairly sharp.

The results obtained during the calculations for the static Debye–Waller factor indicate that the regions near the heterointerfaces and the surface of the sample have the highest defect density (Fig. 3). These structural features were also observed for other multilayer systems analyzed by various methods.[8–10,13] The mean value of the static Debye–Waller factor of the $Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$ layer is roughly equal to 0.8. Essentially the same crystal perfection was exhibited by a previously investigated graded AlGaAs layer grown by metalorganic vapor-phase epitaxy.[9,10] According to the results obtained, the upper $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}$ layer has a structure with a higher defect density. The strong jump in the degree of amorphization on the boundary between this layer and the $Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$ layer is due to the large mismatch between the lattice parameters of these two compounds. As a result, the strong lattice-parameter mismatch leads to relaxation processes, which partially or completely eliminate the tetragonal strain in the multilayer structure. This, in turn, is accompanied by additional defect formation.

7. CONCLUSION

Thus, one of the simplest methods for numerical solution of the inverse problem of diffraction in a distorted crystal

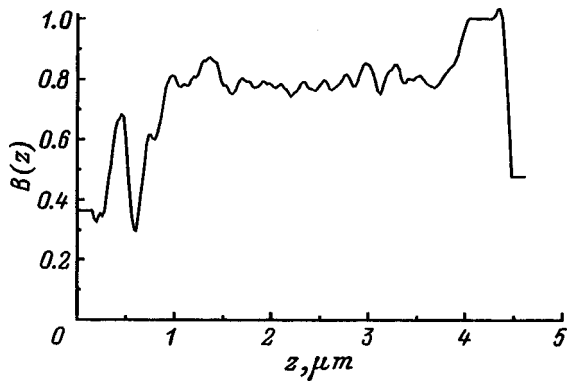

FIG. 3. Variation of the static Debye–Waller factor across the thickness of an $In_{0.22}Ga_{0.78}As_{0.19}Sb_{0.81}/Al_{0.5}Ga_{0.5}As_{0.05}Sb_{0.95}$/(001) GaSb.

structure has been developed. The method permits obtaining information on the distribution of crystal-lattice strains and the degree of amorphization within a fairly short time. Since these structural characteristics are related to the composition of the semiconductor structure investigated, the method opens up additional possibilities for studying relaxation and defect-formation processes in epitaxial multilayer structures as a function of the growth technology.

Because of the features of the method developed, all the solutions were obtained in the form of continuous functions. The method for solving the problem is recurrent; therefore, a good choice for the initial approximation permits significant shortening of the time for the calculation procedure. In addition, there is the problem of nonuniqueness of the solution of the inverse problem.[7] For this reason, careful selection of the starting approximation and utilization of *a priori* information regarding the technology used to fabricate the samples are important steps in structural diagnostics. The algorithm for the method developed calls for calculating Fourier integrals, and large layer thicknesses require an increase in the angular interval during the calculations.

Finally, we note the following fact: the algorithm developed is relatively simple and does not require lengthy complicated calculations. The numerical solution of the inverse diffraction problem on a personal computer with a Pentium-90 processor for 10 iterations takes approximately 10 min when the file for different depths contains 1200 values and the file for different angles contains 400 values.

We thank V. A Kusikov for providing the experimental results.

[1] J. Burget and D. Taupin, Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. **24**, 99 (1968).
[2] J. Burget and R. Collela, J. Appl. Phys. **40**, 3505 (1969).
[3] A. Fukuhara and Y. Takano, Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. **33**, 137 (1977).
[4] A. M. Afanasev, M. V. Kovalchuk, E. K. Kovev, and V. G. Kohn, Phys. Status Solidi A **42**, 415 (1977).
[5] R. N. Kyutt, P. V. Petrashen', and L. M. Sorokin, Phys. Status Solidi A **60**, 381 (1980).
[6] B. C. Larson and J. F. Barhorst, J. Appl. Phys. **51**, 3181 (1980).
[7] A. V. Goncharskiĭ, A. V. Kolpakov, and A. A. Stepanov, *Inverse Problems of X-Ray Diffractometry* [in Russian], Latvian University, Riga (1992), 181 pp.
[8] S. G. Podorov, V. I. Punegov, and V. A. Kusikov, Fiz. Tverd. Tela (St. Petersburg) **36**, 827 (1994) [Phys. Solid State **36**, 454 (1994)].
[9] V. I. Punegov and N. N. Faleev, Fiz. Tverd. Tela (St. Petersburg) **38**, 255 (1996) [Phys. Solid State **38**, 143 (1996)].
[10] V. I. Punegov, K. M. Pavlov, S. G. Podorov, and N. N. Faleev, Fiz. Tverd. Tela (St. Petersburg) **38**, 264 (1996) [Phys. Solid State **38**, 148 (1996)].
[11] D. Taupin, Bull. Soc. Fr. Mineral. Cristallogr. **87**, 469 (1964).
[12] A. V. Kolpakov and V. I. Punegov, Poverkhnost' No. 3, 82 (1988).
[13] K. M. Pavlov, V. I. Punegov, and N. N. Faleev, Zh. Éksp. Teor. Fiz. **107**, 1967 (1995) [JETP **80**, 1090 (1995)].

Translated by P. Shelnitz

Mode-conversion nonreciprocity in a two-section magnetogyrotropic waveguide

D. I. Sementsov and A. M. Shutyĭ

Ul'yanovsk State University, 432700 Ul'yanovsk, Russia
(Submitted December 19, 1996)
Zh. Tekh. Fiz. **69**, 43–46 (March 1999)

It is shown for a two-section magnetogyrotropic waveguide with arbitrary orientation of the magnetization in each of the sections that mode conversion with optimum nonreciprocity can be achieved over a broad range of waveguide-layer thicknesses by varying the orientation of the magnetization in the sections and adjusting their lengths. © *1999 American Institute of Physics.* [S1063-7842(99)00803-X]

The use of magnetogyrotropic waveguides as nonreciprocal devices is one of the high-priority areas in integrated magneto-optics.[1,2] The nonreciprocal properties of a uniformly magnetized film with an isotropic nonmagnetic substrate are significant mainly near the mode cutoff thickness,[3,4] making it difficult to employ them for practical purposes. In magneto-optical waveguides consisting of two or more sections with different magnetization orientations, a significant level of mode-conversion nonreciprocity exists over a fairly broad range of waveguide parameters. The practical utilization of such structures and cascade unidirectional mode converters based on them[5,6] requires a detailed analysis of the waveguide regimes realized in them. This paper examines the influence of the orientation of the magnetization in each of the sections on the nonreciprocity parameter over a broad range of waveguide-layer thicknesses for different section lengths.

1. To investigate nonreciprocal conversion regimes in a two-section magnetogyrotropic waveguide, let us consider a waveguide structure consisting of a substrate, a cladding layer, and two uniformly magnetized waveguide-layer sections fabricated on the basis of a single ferrite-garnet film (Fig. 1). The dielectric constants of the respective layers include the scalar quantities ε_1 and ε_2 and the tensor $\hat{\varepsilon}_f$. Our analysis is based on the coupled-mode approach,[7] under which the part of $\hat{\varepsilon}_f$ that does not depend on magnetization orientation is regarded as the unperturbed part, i.e., $\hat{\varepsilon}_f^{(0)}$, and the dependent part is taken as a small perturbation $\Delta\hat{\varepsilon}_f$. The interaction of two modes with the amplitudes A_μ and A_ν propagating along the z axis in a waveguide structure is described in the two-mode approximation by their coupling equations.[8] The solution of these equations for modes propagating in a uniformly magnetized waveguide section with an initial coordinate z_0 and a final coordinate z can be represented in the following manner:

$$\begin{bmatrix} A_\mu(z) \\ A_\nu(z) \end{bmatrix} = \exp[-i\beta(z-z_0)] \begin{bmatrix} T_- & \gamma t \\ \gamma^* t & T_+ \end{bmatrix} \cdot \begin{bmatrix} A_\mu(z) \\ A_\nu(z) \end{bmatrix}, \quad (1)$$

where the elements of the transmission matrix have the form

$$T_\pm = \cos\varkappa(z-z_0) \pm \frac{i\Delta}{\varkappa} \sin\varkappa(z-z_0),$$

$$t = \frac{i}{\varkappa} \sin\varkappa(z-z_0).$$

Here we have introduced the notations $\varkappa^2 = |\gamma|^2 + \Delta^2$, $2\Delta = \beta_\mu - \beta_\nu + \Delta\beta_\mu - \Delta\beta_\nu$, and $2\overline{\beta} = \beta_\mu + \beta_\nu$, where $\beta_{\mu,\nu}$ are the mode propagation constants, $\Delta\beta_{\mu,\nu}$ are perturbations of the propagation constants, and γ is the mode coupling constant. In the general case the coupling constant is specified by the field profile functions of the respective modes and the perturbation of the dielectric constant:[8]

$$\gamma = k_0 \int \mathcal{E}_\mu^* \Delta\hat{\varepsilon}_f \mathcal{E}_\nu \, dx, \quad (2)$$

where $k_0 = \omega/c$, ω is the frequency, c is the speed of light in free space, and $\mathcal{E}_{\mu,\nu}(x)$ are profile functions, which are normalized to unit power of the energy transmitted per unit width of the waveguide (in the y direction) and determine the distribution of the field of the respective modes in the x direction.

The coupling constant for orthogonally polarized TE and TM modes is found from the expression

$$\gamma = k_0 \int \mathcal{E}_{\mu y}^* (\varepsilon_{yz}\mathcal{E}_{\nu z} + \varepsilon_{yx}\mathcal{E}_{\nu x}) \, dx. \quad (3)$$

The corrections to the mode propagation constants caused by the perturbation of the dielectric constant for the TE_μ mode have the form

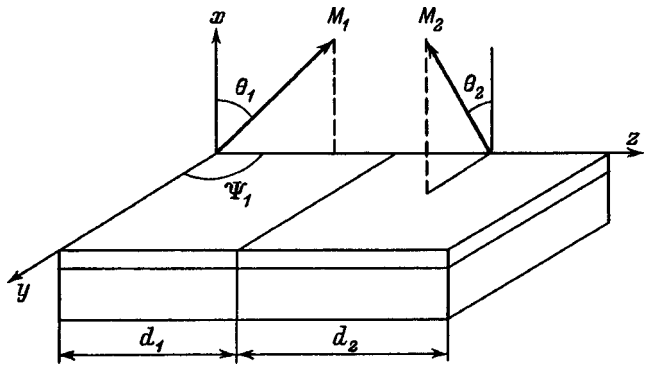

FIG. 1. Two-section unidirectional mode converter.

$$\Delta\beta_{\mu}^{E}=k_0\int \Delta\varepsilon_{yy}\mathcal{E}_{\mu y}^{*}\mathcal{E}_{\mu y}\,dx, \tag{4}$$

and the corrections for the TM_{ν} mode have the form

$$\Delta\beta_{\nu}^{M}=k_0\int [\mathcal{E}_{\nu x}^{*}(\Delta\varepsilon_{xx}\mathcal{E}_{\nu x}+\varepsilon_{xz}\mathcal{E}_{\nu z}) + \mathcal{E}_{\nu z}^{*}(\varepsilon_{zx}\mathcal{E}_{\nu x}+\Delta\varepsilon_{zz}\mathcal{E}_{\nu z})]\,dx. \tag{5}$$

The components of the perturbation $\Delta\varepsilon_{\alpha\beta}$ in these expressions are specified by the polar ($\theta_{1,2}$) and azimuthal ($\psi_{1,2}$) angles of the magnetization in the respective waveguide sections. With consideration of the different orientations of the magnetization in the sections of lengths d_1 and d_2, the amplitudes of the input (μ) and excited (ν) modes propagating from the first to the second section (in the forward direction), are specified at the waveguide exit ($z=d=d_1+d_2$), according to (1), by the expressions

$$A_{\mu}(d)=\exp[i(\Delta_2-\Delta\beta_{\mu})d_2]\left[A_{\mu}(d_1)\left(\cos\varkappa_2 d_2 - i\frac{\Delta_2}{\varkappa_2}\sin\varkappa_2 d_2\right)+iA_{\nu}(d_1)\frac{\gamma_2}{\varkappa_2}\sin\varkappa_2 d_2\right],$$

$$A_{\nu}(d)=\exp[-i(\Delta_2+\Delta\beta_{\nu})d_2]\left[iA_{\mu}(d_1)\frac{\gamma_2^{*}}{\varkappa_2}\sin\varkappa_2 d_2 + A_{\nu}(d_1)\left(\cos\varkappa_2 d_2-i\frac{\Delta_2}{\varkappa_2}\sin\varkappa_2 d_2\right)\right],$$

$$A_{\mu}(d_1)=A_{\mu}(0)\exp[i(\Delta_1-\Delta\beta_{\mu}-\beta_{\mu})d_1]\times\left(\cos\varkappa_1 d_1-i\frac{\Delta_1}{\varkappa_1}\sin\varkappa_1 d_1\right),$$

$$A_{\nu}(d_1)=iA_{\mu}(0)\exp[-i(\Delta_1-\Delta\beta_{\nu}-\beta_{\nu})d_1]\times\frac{\gamma_1^{*}}{\varkappa_1}\sin\varkappa_1 d_1. \tag{6}$$

The subscripts 1 and 2 assign the parameters to the waveguide sections. The derivation of (6) did not take into account effects associated with the reflection and transmission of radiation at the interface between the differently magnetized sections. This is justified by the fact that the difference between the refractive indices of the sections is determined only by the orientation of the magnetic moment and is three to four orders of magnitude smaller than the magnetization-independent refractive index. Expressions for the efficiency of $TE_{\mu}\rightarrow TM_{\nu}$ mode conversion $\eta=|A_{\nu}(d)|^2/|A_{\mu}(0)|^2$ can be obtained on the basis of (6):

$$\eta=\frac{|\gamma_1|^2}{\varkappa_1^2}\sin^2\varkappa_1 d_1\left(\cos^2\varkappa_2 d_2+\frac{\Delta_2^2}{\varkappa_2^2}\sin^2\varkappa_2 d_2\right) + \frac{|\gamma_2|^2}{\varkappa_2^2}\sin^2\varkappa_2 d_2\left(\cos^2\varkappa_1 d_1+\frac{\Delta_1^2}{\varkappa_1^2}\sin^2\varkappa_1 d_1\right) + 2\frac{\sin\varkappa_1 d_1\sin\varkappa_2 d_2}{\varkappa_1\varkappa_2}\left(\mathrm{Re}(\gamma_1\gamma_2^{*})\left[\cos\varkappa_1 d_1\cos\varkappa_2 d_2 - \frac{\Delta_1\Delta_2}{\varkappa_1\varkappa_2}\sin\varkappa_1 d_1\sin\varkappa_2 d_2\right]+\mathrm{Im}(\gamma_1^{*}\gamma_2)\times\left[\frac{\Delta_1}{\varkappa_1}\cos\varkappa_2 d_2\sin\varkappa_1 d_1-\frac{\Delta_2}{\varkappa_2}\cos\varkappa_1 d_1\sin\varkappa_2 d_2\right]\right). \tag{7}$$

To investigate the nonreciprocity of the mode conversion process under consideration for the propagation of radiation in both the forward and reverse (from the second to the first section) directions we assign the section lengths so that the conversion efficiency would be equal to 50% in each of the sections. In this case $\sin\varkappa_1 d_1=\varkappa_1^{+}/\sqrt{2}|\gamma_i|$, and $\cos\varkappa_i d_i=\varkappa_i^{-}/\sqrt{2}|\gamma_i|$, where $i=1,2$, and $\varkappa_i^{\pm}=(|\gamma_i|^2\pm\Delta_i^2)^{1/2}$. The conversion efficiency in this case, which is important from the practical standpoint, has the simple form

$$\eta=\frac{1}{2}-\frac{1}{2}\,\mathrm{Re}[(\Delta_1+i\varkappa_1^{-})(\Delta_2+i\varkappa_2^{-})/\gamma_1^{*}\gamma_2]. \tag{8}$$

2. Let the magnetization be oriented along the radiation propagation direction (the Faraday geometry) in the first section and perpendicularly to it (the Cotton–Mouton geometry) in the second section. Then the real part of the coupling constant in the first section $\mathrm{Re}\,\gamma_1$ is more than an order of magnitude smaller than its imaginary part and can be neglected, and γ_1 is understood to equal $\mathrm{Im}\,\gamma_1$. The mode coupling constant for the second section γ_2 is a real quantity. When the direction of the radiation is reversed, the signs of both coupling constants change (except in the case of $\psi_2=0$, $\theta_2\approx 90°$, which is not considered here because of the small value of γ_2). In the case of the propagation of modes in the reverse direction, the complex conjugate of the coupling constant in the first section should be taken in (8) with resultant reversal of the sign of the second term in this expression. The expressions for the mode conversion efficiency in the forward and reverse directions (which are labeled by $+$ and $-$) ultimately take on the form

$$\eta_{\pm}=\frac{1}{2}\pm(\Delta_1\varkappa_2^{-}+\Delta_2\varkappa_1^{-})/(2\gamma_1\gamma_2). \tag{9}$$

Hence the nonreciprocity parameter $K=\eta_{+}-\eta_{-}$ has the form

$$K=(\Delta_1\sqrt{\gamma_2^2-\Delta_2^2}+\Delta_2\sqrt{\gamma_1^2-\Delta_1^2})/\gamma_1\gamma_2. \tag{10}$$

It follows from (10) that the there is no mode conversion reciprocity when there is complete phase synchronism of the modes in the two sections, i.e., when $\Delta_1=\Delta_2=0$, as well as when $\Delta_1=-\Delta_2\gamma_1/\gamma_2\,\mathrm{sgn}(\gamma_1/\gamma_2)$. The nonreciprocity reaches a maximum when $K=1$ or $K=-1$. In the former case complete switching of the polarization of the input mode to the orthogonal direction takes place when it propagates in the forward direction, whereas there is conservation of the polarization when it propagates in the reverse direction. In the latter case complete mode conversion occurs only in the reverse direction, while there is no mode conversion in the forward direction.

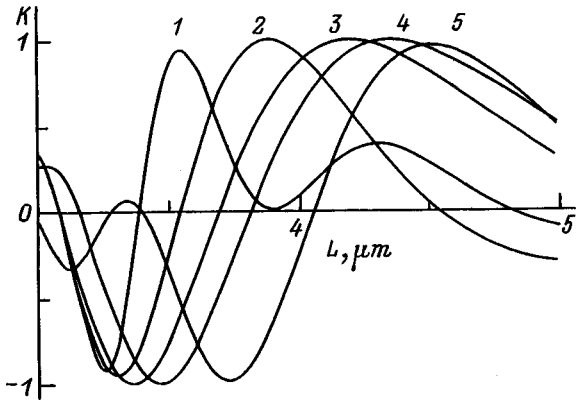

FIG. 2. Dependence of the nonreciprocity parameter K on the waveguide-layer thickness L for $TE_0 \rightarrow TM_0$ conversion and various orientations of the magnetization in a cascade unidirectional mode converter.

If the perturbation of the propagation constant $\Delta\beta_1$ is neglected, then $\Delta_1 = \Delta_2$, and $|K|$ achieves its maximum values when

$$\Delta = \pm[\sqrt{(\gamma_1^2+\gamma_2^2)^2/4+\gamma_1^2\gamma_2^2}-(\gamma_1^2+\gamma_2^2)/2]^{1/2}. \quad (11)$$

In the case of zero phase mismatch in one of the sections, the nonreciprocity level is greatest when the phase mismatch in the other section $\Delta_i = \pm\gamma_i$, which is the condition for 50% conversion efficiency ($\eta_{\max}=0.5$). When $|\Delta_i| > |\gamma_i|$, the maximum attainable efficiency is less than 50%.

3. For a more complete investigation of a two-section mode converter we perform a numerical analysis of the expression (7) for the mode conversion efficiency using a waveguide structure with the parameters $\varepsilon_{xx}^{(0)}=4.5383$, $\varepsilon_{yy}^{(0)}=\varepsilon_{zz}^{(0)}=4.5371$, $\varepsilon_1=3.8$, and $\varepsilon_2=1$, the linear magneto-optical parameter $f=3.07\times10^{-4}$, and the quadratic magneto-optical parameters $g_{44}=2.4\times10^{-4}$ and $\Delta g=-7.3\times10^{-s}$, which correspond to the magnetogyrotropic waveguides for $\lambda=1.15$ μm that are used in practice.[2] The [111] crystallographic axis of the film is normal to the waveguide surface and coincides with the x axis, and the $[11\bar{2}]$ and $[\bar{1}10]$ axes coincide with the y and z axes. A cascade unidirectional mode converter with an assigned waveguide-layer thickness was investigated in Ref. 5 for certain orientations of the magnetization in each section, which produce the strongest mode-conversion nonreciprocity. However, waveguides with an arbitrary waveguide-layer thickness L are used in practice, and since the mode conversion efficiency depends significantly on waveguide-layer thickness, the dependence of the nonreciprocity parameter on L and the possibility of achieving the optimum values $K=\pm1$ for films of arbitrary thickness must be considered. Figure 2 shows the dependence of the nonreciprocity parameter K on the waveguide-layer thickness L for $TE_0 \rightarrow TM_0$ conversion and various magnetization orientations: $\psi_1=90°$ and $\theta_1=20$, 45° (curves *1* and *2*) and $\theta_1=90°$ (curves *3–5*); $\psi_2=0$ and $\theta_2=45°$ (curves *1–3*) and $\theta_2=55$, 70° (curves *4* and *5*). The section lengths were selected so that the largest value of the nonreciprocity parameter would be achieved in the thickness range considered: $d_1=0.725$, 0.347 cm (curves *1* and *2*),

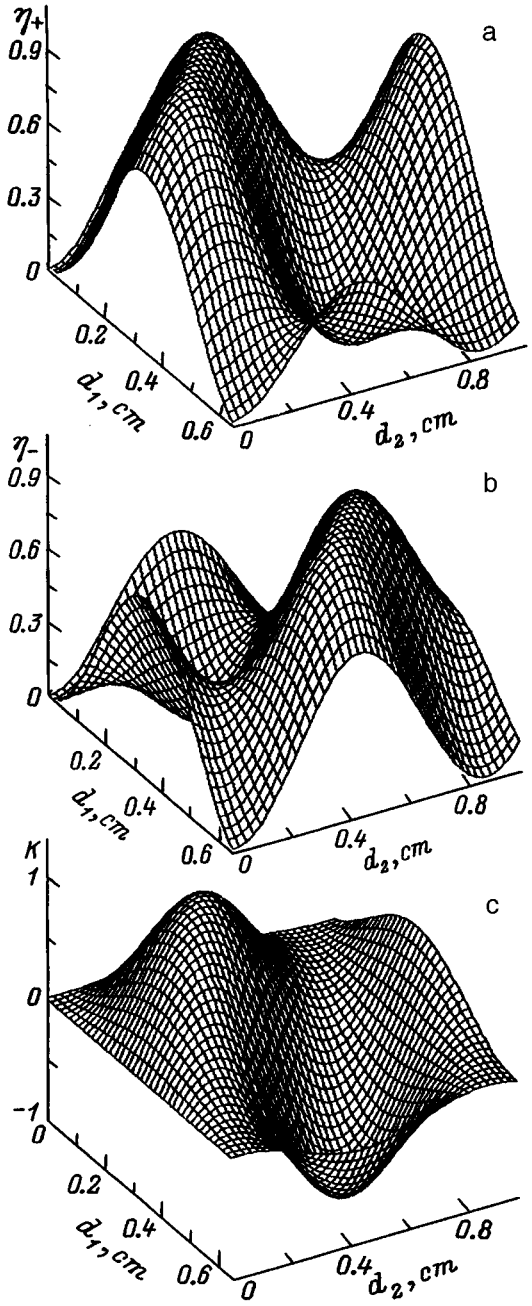

FIG. 3. Conversion efficiencies of modes propagating in the forward η_+ (a) and reverse η_- (b) directions and value of the nonreciprocity parameter K (c) as functions of the section lengths.

$d_1=0.2$ cm (curves *3–5*), $d_2=0.284$ cm (curves *1–3*), $d_2=0.355$, 0.567 cm (curves *4* and *5*). It can be seen that each orientation of the magnetization in the sections of the cascade waveguide structure has corresponding waveguide thicknesses $L_\pm$, at which the nonreciprocity parameter $K\approx\pm1$ (in the case of appropriate selection of the section lengths). The thickness L_0, at which mode-conversion nonreciprocity is scarcely manifested, lies between L_- and L_+. An analysis reveals that the inequality $|K(L)|<|K(L_\pm)|$ holds at thicknesses differing from $L_\pm$. This is very important from the practical standpoint, since an external magnetic field can be used to adjust the orientation of the magnetiza-

tion in the sections in the case of a concrete waveguide-layer thickness to achieve the strongest possible manifestation of mode-conversion nonreciprocity for specific values of d_1 and d_2.

Figure 3 presents the conversion efficiency of modes propagating in the forward (η_+, *a*) and reverse (η_-, *b*) directions and the value of the nonreciprocity parameter K on the plane of the section lengths d_1 and d_2. In the first section the magnetization is oriented along the propagation direction of the modes ($\psi_1=\theta_1=\pi/2$), and in the second section it is perpendicular to it ($\psi_2=0$, $\theta_2=\pi/4$). The plots have been drawn for the film thickness at which $K=1$, i.e., $L=L_+ \approx 4.2$ μm, is achieved for these orientations of the magnetization in the sections. For the parameters of the waveguide structure chosen an increase in the length of the sections leads to reversal of the sign of K, and the nonreciprocity parameter becomes equal to $K=-1$ at certain lengths. Therefore, at the exit from the waveguide the input mode is converted into an orthogonal mode in the case of propagation in the reverse direction and does not change its polarization in the case of propagation in the forward direction. Similar plots constructed for $L=L_0\approx 3.7$ μm show that nonreciprocity is weakly manifested and $|K|\leqslant 0.2$ for any values of d_1 and d_2.

It follows from the analysis presented that the operating regimes of a given integrated-optical element can be controlled and an optimum degree of mode-conversion nonreciprocity can be obtained by varying the orientation of the magnetization in the sections of a cascade waveguide structure and adjusting the lengths of the sections. An extensive group of efficient integrated devices, such as modulators, valves, and isolators, for the near-IR range can be created on the basis of magnetogyrotropic cascade waveguides with pronounced nonreciprocal properties.

[1] A. M. Prokhorov, G. A. Smolenskiĭ, and A. N. Ageev, Usp. Fiz. Nauk. **143**, 33 (1984) [Sov. Phys. Usp. **27**, 339 (1984)].
[2] A. K. Zvezdin and V. A. Kotov, *Magneto-optics of Thin Films* [in Russian], Nauka, Moscow (1988), 192 pp.
[3] D. I. Sementsov and A. M. Shutyĭ, Opt. Spektrosk. **79**, 507 (1995) [Opt. Spectrosc. **79**, 470 (1995)].
[4] D. I. Sementsov and A. M. Shutyĭ, Zh. Tekh. Fiz. **65**(2), 156 (1995) [Tech. Phys. **40**, 200 (1995)].
[5] J. P. Castera and G. Hepner, IEEE Trans. Magn. **MAG-13**, 1583 (1977).
[6] H. Hemme, H. Dotsch, and H.-P. Menzler, Appl. Opt. **26**, 3811 (1987).
[7] *Guided-Wave Optoelectronics*, T. Tamir (Ed.), Springer-Verlag, Berlin–New York (1988); Mir, Moscow (1991), 575 pp.
[8] D. I. Sementsov, A. M. Shutyĭ, and O. V. Ivanov, Radiotekh. Elektron. **41**, 421 (1996).

Translated by P. Shelnitz

Dependences of the refractive indices on the proton concentration in H:$LiNbO_3$ waveguides

Yu. N. Korkishko and V. A. Fedorov

Moscow Institute of Electronic Technology (Technical University), 103498 Moscow, Russia
(Submitted January 24, 1997; resubmitted August 4, 1997)
Zh. Tekh. Fiz. **69**, 47–57 (March 1999)

Proton-exchanged optical waveguides in lithium niobate crystals exhibit a rich variety of structures and phases. It is established that seven $H_xLi_{1-x}NbO_3$ crystalline phases with a lithium niobate structure may form under various conditions of proton exchange and post-exchange annealing. A method is proposed to determine the proton concentration in the various phases identified. Relationships are established between the structural parameters, the proton concentration, and the ordinary and extraordinary refractive indices of various $H_xLi_{1-x}NbO_3$ crystalline phases. The results can explain various optical phenomena observed in proton-exchanged waveguides and permit prediction of the characteristics of light-guide structures.
 [S1063-7842(99)00903-4]

1. INTRODUCTION

Lithium niobate $LiNbO_3$ is a ferroelectric widely used in integrated optics and acoustoelectronics. One of the main methods used to fabricate optical waveguides in these crystals, apart from the diffusion of titanium from films, is proton exchange.[1] Rice and Jackel[2,3] reported detailed studies of $H_xLi_{1-x}NbO_3$ solid solutions produced by proton exchange in powders and constructed an $LiNbO_3$–$HNbO_3$ phase diagram. They established the existence of two rhombohedral phases and one monoclinic phase in this system. Relationships were obtained between the crystal-lattice parameters and the proton concentration in the unstrained α- and β-$H_xLi_{1-x}NbO_3$ rhombohedral phases having the same structure as lithium niobate.[2,3] In particular, a thermogravimetric analysis revealed that $x \leqslant 0.12$ for the α phase of $H_xLi_{1-x}NbO_3$, whereas the β phase occurs in the concentration range $0.55 \leqslant x \leqslant 0.75$. However, it should be noted that the results obtained for powders cannot be directly applied to the strained solid solutions of the same composition now being used in practical applications. In fact, the stresses created at the interface between the proton-exchanged layer and the substrate as a result of lattice mismatch may significantly modify the equilibrium phase diagram constructed for powders. Recent investigations of proton-exchanged waveguides in a crystal of lithium tantalate ($LiTaO_3$), which is isomorphous to lithium niobate, revealed rich structural and phase diversity.[4–6]

Our previous investigations[7–13] permitted the identification of seven different crystalline phases of $H_xLi_{1-x}NbO_3$, viz., α, κ_1, κ_2, β_1, β_2, β_3, and β_4, which can be formed in proton-exchanged layers (Fig. 1). Under certain conditions of proton exchange and post-exchange annealing, one, two, or even three different crystalline phases of $H_xLi_{1-x}NbO_3$, may form as separate thin layers in proton-exchanged $LiNbO_3$ optical waveguides.

Table I shows that an increase in the proton concentration in the proton-exchange melts leads to the following phase transitions: $\alpha \rightarrow \beta_1 \rightarrow \beta_2 \rightarrow \beta_3 \rightarrow \beta_4$. All five phases can be formed only in *Z*-cut $LiNbO_3$, whereas only four phases can occur in *X*-cut lithium niobate (α, β_1, β_2, and β_3), and only two (α and β_1) can occur in *Y*-cut crystals. Waveguides containing the α phase have a graded refractive-index profile and a small refractive-index increment for extraordinary rays at the surface $\Delta n_e < 0.03$. This crystalline phase most likely corresponds to the equilibrium α phase identified by Rice and Jackel[2,3] and can be formed by proton exchange in a weakly acidic melt, for example in a benzoic acid melt containing 2.5–3.5 mol % lithium benzoate at process temperatures between 200 and 350 °C (Ref. 14).

Optical waveguides containing the other three phases, viz., β_1, β_2, and β_3, which can be formed by direct low-temperature proton exchange, have a graded refractive-index profile with $\Delta n_e \geqslant 0.09$. However, waveguides containing the β_4 phase exhibit a complex refractive-index profile. In the range $0.12 \leqslant \Delta n_e \leqslant 0.15$ the index profile is graded, whereas at lower values of Δn_e the refractive-index increment exhibits a steep drop.[12,13]

It should be specially stressed that proton-exchanged waveguides containing different crystalline phases may have essentially identical refractive-index profiles, but other important parameters, such as their electro-optic and nonlinear coefficients, optical losses, stability of the parameters, and some other parameters, may differ substantially. This explains the appreciable differences in the properties of integrated-optics devices fabricated using waveguides having the identical index profiles, which led some researchers to conclude that proton exchange is not a reproducible method. In fact, waveguides having identical index profiles may belong to different crystalline phases and thus exhibit different physical properties.

Only the β_1-$H_xLi_{1-x}NbO_3$ phase exists independently at the surface of wafers whereas the β_2 and β_3 phases can exist only together with β_1, which occurs as a separate layer beneath the surface phase.[10,12,15] The β_4 phase can appear

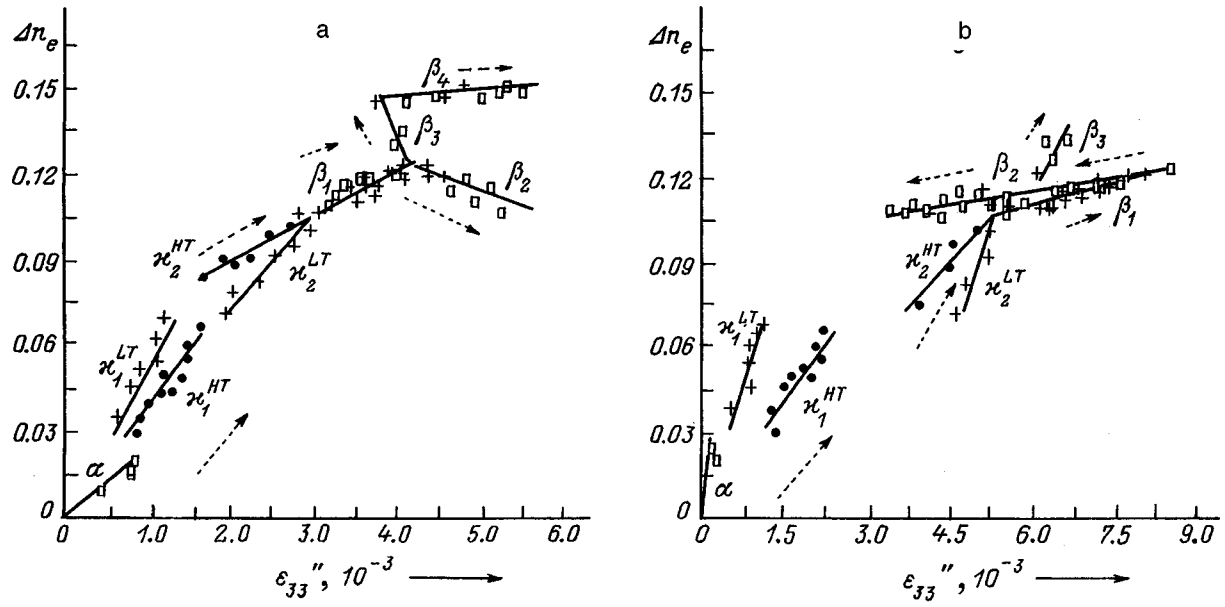

FIG. 1. Structural phase diagram of proton-exchanged $H_xLi_{1-x}NbO_3$ waveguide layers formed on *Z*- (a) and *X*-cut (b) samples: the arrows indicate the direction of increasing proton concentration; □ — direct exchange, $T=200-300\,^\circ$C; ●, + — annealing at 300 °C and 400 °C, respectively.

only in conjunction with layers of the β_3 and β_1 phases located in the bulk of the crystal.[10,13]

Post-exchange annealing can produce new phases, viz., κ_1 and κ_2, which have low-temperature (κ_i^{LT}) and high-temperature (κ_i^{HT}) modifications.[9,10,13,15] The low-temperature modifications are formed by annealing at $T<340\,^\circ$C, whereas the high-temperature modifications are formed by annealing at $T>400\,^\circ$C. A characteristic feature of the existence of these phases is that, apart from κ_1^{HT}, they occupy only part of the waveguide region: the phase with a higher proton concentration is situated closer to the surface. As a result of annealing, it is possible to have a situation where a four-layer structure forms with a β_1 phase which has not disappeared completely from the surface and κ_2, κ_1, and α phases, which are distributed consecutively as separate graded-index layers. Further annealing at temperatures below 340 °C leads to the systematic disappearance of the β_1, κ_2^{LT}, and κ_1^{LT} phases, and, ultimately, only the α phase remains. Thus, the β_1, κ_2^{HT}, and κ_1^{HT} phases disappear successively during high-temperature annealing ($T>400\,^\circ$C).

Trigonal $H_xLi_{1-x}NbO_3$ phases are not the only phases which can be formed in proton-exchanged waveguide layers in $LiNbO_3$. When proton exchange takes place on $(0\bar{1}4)$-cut $LiNbO_3$ in melts of benzoic or pyrophosphoric acid or ammonium dihydrogen phosphate, a monoclinic η phase is formed.[8,15–17]

Experimental investigations of the refractive-index increment Δn_e as a function of the proton concentration in proton-exchanged $LiNbO_3$ optical waveguides have been carried out by many researchers using various methods of analysis. The available published results are presented in Table II. Although most authors report a nonlinear relationship between Δn_e and the concentration, the numerical val-

TABLE I. Proton exchange conditions for the formation of various $H_xLi_{1-x}NbO_3$ phases.

Phase at surface	$LiNbO_3$ cut	Benzoic acid + lithium benzoate, wt. % (300 °C) (Refs. 11 and 14)	$KHSO_4$ solution in glycerol (220 °C)	Other proton sources
α	*X*, *Y*, *Z*	>2.5		
β_1	*X*, *Y*, *Z*	1−2.5	0−1 g/l	Stearic, palmitic, oleic acid
β_2	*X*, *Z*	0	2−4 g/l	Myristic acid
β_3	*X*, *Z*		8−15 g/l	Lauric acid
β_4	*Z*		>20 g/l	Pyrophosphoric acid, ammonium dihydrogen phosphate, potassium bisulfate
κ_2^{HT}	*X*, *Y*, *Z*	⋯	⋯	$ZnSO_4-K_2SO_4-Na_2SO_4-KHSO_4$ ($C_{KHSO_4}=0.04-0.07$ mol/kg)
κ_1^{HT}	*X*, *Y*, *Z*	⋯	⋯	$ZnSO_4-K_2SO_4-Na_2SO_4-KHSO_4$ ($C_{KHSO_4}=0.01-0.04$ mol/kg)
η	$(0\bar{1}4)$	0	>2 g/l	Pyrophosphoric, sulfuric acid, ammonium dihydrogen phosphate, potassium bisulfate

TABLE II. Data on proton concentrations in various $H_xLi_{1-x}NbO_3$ phases.

Cut	Proton source	Annealing temperature, °C	$H_xLi_{1-x}NbO_3$ phase	Δn_e	x	Method of determination	References
Z	BA	⋯	β_2	⋯	0.4−0.5	SIMS	Ref. 18
		275	β_1	0.11	0.32		
X	BA	⋯	β_2	0.12	0.85 [a]	SIMS	Ref. 19
Z	BA	400	from κ_2^{HT} to κ_1^{HT}	⋯	<0.22	SIMS	Ref. 20
Z	BA	⋯	β_2	⋯	0.3	SIMS	Ref. 21
		400	β_1		0.26		
			κ_1^{HT}		<0.2		
Z	BA	275	β_1	0.11	0.32	SIMS	Ref. 22
		400	κ_1^{HT}	0.047	0.14		
Z, X	BA	400	β_2	0.11	0.65−0.75	NAA	Ref. 23
			β_1	0.105	0.6		
			κ_1^{HT}	0.07	0.4		
			κ_1^{HT}	0.02	0.15		
Z, X	BA	⋯	β_2	⋯	0.65−0.75	NAA	Ref. 24
Z	BA	⋯	β_2	⋯	0.75	AAA	Ref. 25
Z	BA	⋯	β_2	⋯	0.33	AAA	Ref. 26
X	BA+1% LB	300	κ_2^{LT}	0.081	0.315	RBS	Ref. 27
		320	κ_1^{LT}	0.039	0.315		
		340	α or κ_1^{LT}	0.039	0.275		
		400	κ_1^{HT}	0.0035	0.020		
Z	BA	300−370	α	0−0.02	0−0.12	RBS	Ref. 17
			$\kappa_1^{LT}+\kappa_1^{HT}$	0.02−0.075	0.12−0.4		
			$\kappa_2^{LT}+\kappa_2^{HT}$	0.075−0.105	0.4−0.55		
			β_1	~0.11	0.55−0.65		
			β_2	~0.11	0.65−0.8		
Z	PA	⋯	β_4	0.132	0.68	RS	Ref. 28
		200	β_4 or β_3	0.124	0.68		
		260	β_1	0.114	0.65		
		280	β_1	0.113	0.64		
		300	β_1	0.110	0.55		
		320	β_1	0.104	0.55		
		340	κ_2^{LT}	0.099	0.51		
		360	$\kappa_2^{LT}+\kappa_2^{HT}$	0.092	0.41		
Z	PA	⋯	β_4	0.148	0.75	RS	Ref. 29
		250	β_3 or β_1	0.12	0.66		
Z	PA	⋯	β_4	>0.14	0.9	RBS	Ref. 30
		300	β_3 or β_1	⋯	0.5		
		300	β_1	⋯	0.4		

Note: [a]Only the lithium concentration was determined. BA — benzoic acid, LB — lithium benzoate, PA — pyrophosphoric acid, SIMS — secondary-ion mass spectroscopy, RBS — Rutherford backscattering spectroscopy, AAA — atomic absorption analysis, RS — Raman scattering, NAA — nuclear activation analysis.

ues obtained for the hydrogen concentration differ substantially in different publications, even for structures fabricated under identical conditions. This is not surprising since these investigations neglected the existence of different crystalline phases in proton-exchanged $LiNbO_3$ waveguides. Quite clearly, the matrix effect observed in all the instrumental methods of analysis used differs for different phases, making it difficult to compare the amplitudes of the signals being analyzed with a standard and restricting the determination of the proton concentration.

In the present paper we propose a method for determining the dependence of Δn_e on the proton concentration in the various crystalline phases formed in proton-exchanged waveguides in lithium niobate. The method is based on analyzing the change in the area under the refractive-index curve $\Delta n_e(z)$ during annealing.

The methods of fabricating and investigating the proton-exchanged waveguides were described in detail in our previous publications.[8–13]

2. RESULTS AND DISCUSSION

It is interesting to note that the area under the plot of the distribution of the extraordinary-refractive-index increment does not remain constant during annealing. Various authors (see Refs. 17 and 31–33) cited this change in area as evidence of a nonlinear relationship between the refractive-index increment and the proton concentration in proton-exchanged $LiNbO_3$ waveguides.

The most significant change in area under the $\Delta n_e(z)$ curve is observed for waveguides initially containing predominantly the β_2 and β_4 phases. We know that the β_2, β_3, and β_4 phases contain interstitial protons, which are manifested by the presence of a depolarized –O–H bond in the

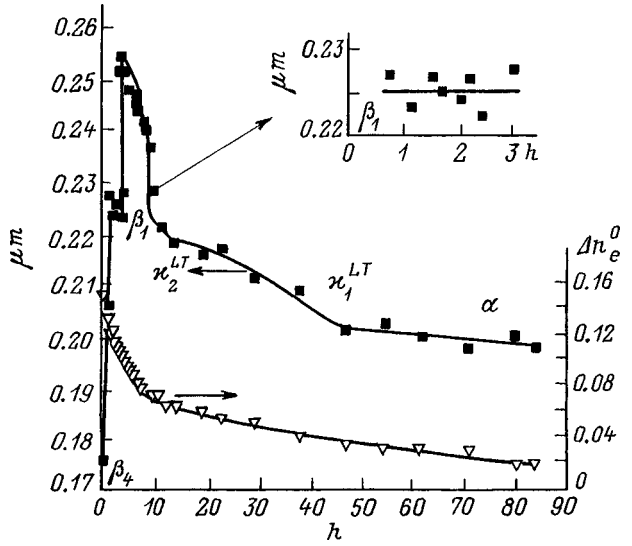

FIG. 2. Dependences of the area under the $\Delta n_e(z)$ profile and the refractive-index increment at the surface Δn_e^0 on the annealing time at $T=340\,^{\circ}\mathrm{C}$. Initial sample — Z-cut $LiNbO_3$, treatment in an ammonium dihydrogen phosphate melt at 220 °C for 3 h (β_4-$H_xLi_{1-x}NbO_3$ phase at the surface).

infrared spectra at 3240 cm^{-1}. During annealing, the protons leave interstitial positions, which have little effect on the refractive indices,[20,21] and occupy positions in oxygen planes, which are ''active'' in affecting the refractive index.[20,21] As a result, the area under the $\Delta n_e(z)$ profile increases substantially. Figure 2 gives the area under the extraordinary-refractive-index increment as a function of annealing time for a proton-exchanged $LiNbO_3$ waveguide initially containing the β_4 phase at the surface.

Quite clearly, the relationships between Δn_e and the hydrogen concentration x differ for different crystalline phases. However, as we can see, it is possible to determine the hydrogen concentration in all phases if we know the dependence $\Delta n_e(z)$ for at least one phase.

We make three assumptions, which in our view are fully justified.

1. We assume that the refractive-index increment Δn_e in each crystalline phase containing no interstitial hydrogen (α-, κ_1-, κ_2- and β_1-$H_xLi_{1-x}NbO_3$) depends linearly on the proton concentration C_H:

$$C_H^i = A(i)\cdot\Delta n_e^i + E(i) \quad \text{or} \quad x = B(i)\cdot\Delta n_e^i + D(i), \tag{1}$$

where $A(i)$, $B(i)$, $E(i)$, and $D(i)$ are constants for each i-$H_xLi_{1-x}NbO_3$ phase.

2. The second assumption is that the total quantity of hydrogen Q remains constant during annealing in a dry atmosphere. Then, the protons only undergo redistribution over the depth of the waveguide layer. However, we note that Zavada *et al.*[21] reported that the annealing of proton-exchanged waveguides in a moist atmosphere leads to an increase in the total hydrogen content in the surface layer as a result of gas-phase diffusion; however, our experiments were carried out in a dry atmosphere. It was observed that when samples containing $H_xLi_{1-x}NbO_3$ phases (β_1, κ_1, κ_2, or α) without interstitial hydrogen are annealed, the area under the infrared absorption peak near 3500 cm^{-1} remains constant, confirming the validity of the assumption. Thus, we can write

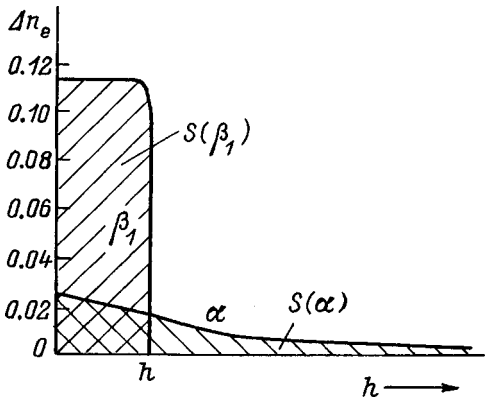

FIG. 3. Transformation of the refractive-index profile during annealing of a β_1-$H_xLi_{1-x}NbO_3$ waveguide.

$$Q = S\int_0^{\infty} C_H(z)\,dz = S\sum_i^N \int_{z_{i-1}}^{z_i} (A(i)\cdot\Delta n_e^i(z) + E(i))\,dz, \tag{2}$$

where $i=1,\ldots,N$ denotes a crystalline phase without interstitial hydrogen ($N_{max}=4$) formed in proton-exchanged waveguides in the form of a separate layer, z_{i-1} and z_i are the coordinates of the upper and lower concentration boundaries of the i phase, respectively ($z_0=0$, and for the α phase, $z_i=\infty$), and $S=\text{const}$ is the surface area of the sample.

Waveguides containing the β_1 phase have a stepped refractive-index profile (Fig. 3) and

$$Q = S\cdot C_H\cdot h, \tag{3}$$

where h is the waveguide depth.

We established experimentally that the area under the $\Delta n_e(z)$ curve remains constant in all stages of annealing when the β_1 phase is present in the waveguide (inset to Fig. 2). Since $Q=\text{const}$, we can write

$$\int_0^h (B(\beta_1)\cdot\Delta n_e^{\beta_1}(z) + D(\beta_1))\,dz$$

$$= B(\beta_1)\int_0^h \Delta n_e^{\beta_1}(z)\,dz + D(\beta_1)\cdot h = \text{const}, \tag{4}$$

which can be satisfied only if $E(\beta_1)=0$ and $D(\beta_1)=0$.

Proton-exchanged $LiNbO_3$ waveguides containing the α phase have the lowest hydrogen concentration, and thus the weakest matrix effect is observed for these waveguides when the proton concentration is determined by various instrumental methods of analysis (such as secondary-ion mass spectrometry, Rutherford backscattering spectroscopy, and so on). Most investigations indicate that in the α phase $x\leqslant 0.12$ is obtained for $\Delta n_e\leqslant 0.025$ (Table II), in good agreement with Refs. 2 and 3. On the basis of this reasoning, we take $B(\alpha)=0.12/0.025=4.8$, and obviously $D(\alpha)=0$.

After determining the dependence of Δn_e on x for the α phase, it is possible to determine the dependences of Δn_e on x for all phases without interstitial photons by comparing the

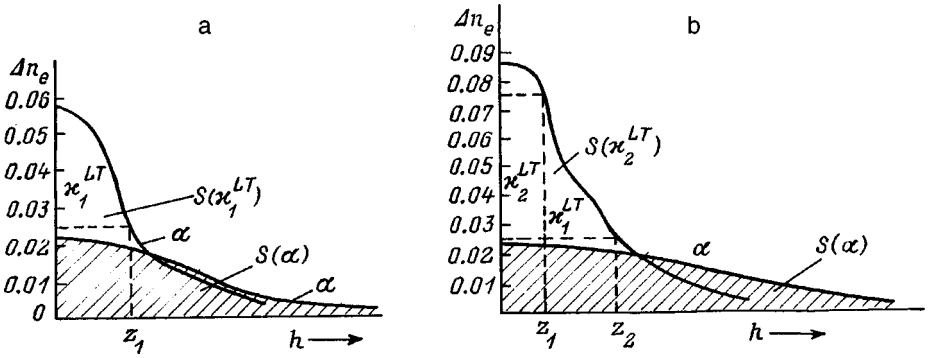

FIG. 4. Refractive-index profiles for proton-exchanged $LiNbO_3$ waveguides with the κ_1^{LT} (a) and κ_2^{LT} phases (b) on the surface: $S(\kappa_1^{LT})$ and $S(\kappa_2^{LT})$ — areas under the refractive-index profiles; $S(\alpha)$ — areas under the refractive-index profiles for the same samples annealed until the α phase appears at the surface.

areas under the experimental $\Delta n_e(z)$ profile obtained for the same sample in various stages of annealing. It was found that if a proton-exchanged $LiNbO_3$ waveguide contained the β_1 phase, long-term annealing at $T \leqslant 340\,^\circ C$ could bring it into the α phase (Fig. 3)

$$B(\beta_1)\int_0^h \Delta n_e^{\beta_1}(z)dz = B(\alpha)\int_0^\infty \Delta n_e^{\alpha}(z)\,dz. \tag{5}$$

Studies of many samples showed that the area under the $\Delta n_e(z)$ profile for a waveguide containing the β_1 phase was $17\pm 2\%$ greater than that for the same sample annealed until conversion to the α phase took place. We thus obtain $B(\beta_1)=4.15$.

We established experimentally that the area under the $\Delta n_e(z)$ profile for waveguides containing the κ_1^{LT} and κ_2^{LT} phases varies during annealing and depends on the refractive-index increment at the surface of the structure Δn_e^0. We use $S(\kappa_2^{LT})$, $S(\kappa_1^{LT})$, and $S(\alpha)$ to denote the area under the $\Delta n_e(z)$ curves for waveguides at whose surfaces the κ_2^{LT}, κ_1^{LT} and α phases form, respectively. Annealing clearly causes the following phase transformations: $\kappa_2^{LT}\rightarrow\kappa_1^{LT}\rightarrow\alpha$. The experiments showed that the ratio of the areas under the $\Delta n_e(z)$ profiles for the same sample $K(\kappa_2^{LT})=S(\alpha)/S(\kappa_2^{LT})$ depends on the depth of the initial waveguide, whereas $K(\kappa_1^{LT})=S(\alpha)/S(\kappa_1^{LT})$ depends only on the refractive-index increment at the surface of the sample. We also find $K(\kappa_1^{LT})=0.9$ at the upper boundary of the κ_1^{LT} phase, where $\Delta n_e^0=0.075$, and $K(\kappa_1^{LT})=1$ at the lower boundary ($\Delta n_e^0=0.025$).

3. The third assumption is that no concentration jumps occur at crystalline phase boundaries, i.e., at the $\alpha-\kappa_1^{LT}$, $\kappa_1^{LT}-$, $\kappa_2^{LT}-$, and $\kappa_2^{LT}-\beta_1$ phase boundaries. This assumption is fully justified because no Δn_e jumps occur at these phase boundaries.[10,12,13,15] Making this assumption, we can determine the dependence of Δn_e on x in the κ_1^{LT} phase by comparing the areas under the $\Delta n_e(z)$ curves for the same sample annealed to states in which the κ_1^{LT} phase and then the α phase form at the surface of the proton-exchanged $LiNbO_3$ waveguide (Fig. 4a). In this case, the material balance may be written as follows:

$$\left[B(\kappa_1^{LT})\int_0^{z_1}\Delta n_e^{\kappa_1^{LT}}(z)\,dz + D(\kappa_1^{LT})\cdot z_1 + B(\alpha)\int_{z_1}^\infty \Delta n_e^{\alpha}(z)\,dz\right]_{\kappa_1^{LT}/\alpha} = \left[B(\alpha)\int_0^\infty \Delta n_e^{\alpha}(z)\,dz\right]_{\alpha}. \tag{6}$$

Here z_1 is the coordinate of the $\alpha-\kappa_1^{LT}$ phase boundary (Fig. 5a) which, like the area under the $\Delta n_e(z)$ curve, can easily be determined from the experimental refractive-index profile. Taking into account the experimentally established fact that

$$\left[\int_0^\infty \Delta n_e^{\alpha}(z)\,dz\right]_{\alpha} = K\left[\int_0^{z_1}\Delta n_e^{\kappa_1^{LT}}(z)\,dz + \int_{z_1}^\infty \Delta n_e^{\alpha}(z)\,dz\right]_{\kappa_1^{LT}/\alpha}, \tag{7}$$

for a two-layer κ_1^{LT}/α structure, we obtain (Fig. 4a)

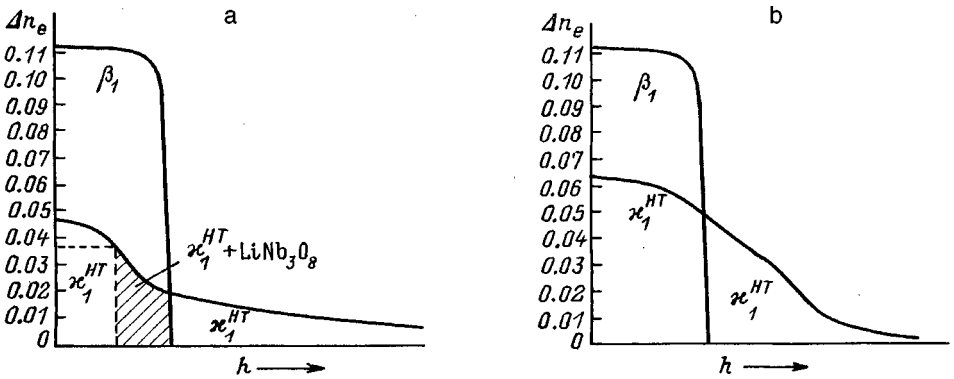

FIG. 5. Refractive-index profiles for proton-exchanged $LiNbO_3$ waveguides annealed at 400 °C: a — case where the $LiNb_3O_8$ phase is formed, b — no $LiNb_3O_8$ phase is formed.

$$(B(\kappa_1^{\mathrm{LT}})-K\cdot B(\alpha))\int_0^{z_1}\Delta n_e^{\kappa_1^{\mathrm{LT}}}(z)\,dz+B(\alpha)$$

$$\times(1-K)\int_{z_1}^{\infty}\Delta n_e^{\alpha}(z)\,dz+D(\kappa_1^{\mathrm{LT}})z_1=0. \tag{8}$$

It is widely known that optical waveguides containing the κ_1^{LT} phase at the surface have a refractive-index profile described by the Gaussian function

$$\Delta n_e=\Delta n_e^0\exp\left(-\frac{z^2}{d^2}\right). \tag{9}$$

Thus, the depth of the $\kappa_1^{\mathrm{LT}}-\alpha$ phase boundary is defined as $z_1=d\sqrt{\ln(\Delta n_e^0/\Delta n_e^0(\alpha))}$, where $\Delta n_e^0(\alpha)=0.025$ is the refractive-index increment at the upper boundary of the α phase. We then have

$$\int_0^{z_1}\Delta n_e\,dz=\frac{\sqrt{\pi}}{2}\Delta n_e d\,\mathrm{erf}\sqrt{\ln(\Delta n_e^0/\Delta n_e^0(\alpha))}, \tag{10}$$

$$\int_{z_1}^{\infty}\Delta n_e\,dz=\frac{\sqrt{\pi}}{2}\Delta n_e(1-\mathrm{erf}\sqrt{\ln(\Delta n_e^0/\Delta n_e^0(\alpha))}). \tag{11}$$

Substituting Eqs. (9), (10), and (11) into Eq. (8), we obtain

$$B(\kappa_1^{\mathrm{LT}})\mathrm{erf}\sqrt{\ln(\Delta n_e^0/\Delta n_e^0(\alpha))}+D(\kappa_1^{\mathrm{LT}})\frac{2}{\Delta n_e^0\sqrt{\pi}}$$

$$\times\sqrt{\ln(\Delta n_e^0/\Delta n_e^0(\alpha))}=B(\alpha)((K-1)$$

$$+\mathrm{erf}\sqrt{\ln(\Delta n_e^0/\Delta n_e^0(\alpha))}). \tag{12}$$

By solving the system of equations (12) for the experimental profiles, we determined the constants for the κ_1^{LT} phase: $B(\kappa_1^{\mathrm{LT}})=4.56$ and $D(\kappa_1^{\mathrm{LT}})=0.003$. When the constants $B(\alpha)$, $B(\kappa_1^{\mathrm{LT}})$, and $D(\kappa_1^{\mathrm{LT}})$ are known, it is possible to repeat the procedure for the same sample with a determination of $\Delta n_e(z)$ in the annealing stages when the κ_2^{LT} and α phases form at the waveguide surface, respectively. Using the assumption put forward above that $B(\kappa_1^{\mathrm{LT}})\,\Delta n_e+D(\kappa_1^{\mathrm{LT}})=B(\kappa_2^{\mathrm{LT}})\,\Delta n_e+D(\kappa_2^{\mathrm{LT}})$ at the upper boundary of the κ_1^{LT} phase (or the lower boundary of the κ_2^{LT} phase), where $\Delta n_e=0.075$ (Refs. 9,12,13, and 15), we can determine the constants $B(\kappa_2^{\mathrm{LT}})$ and $D(\kappa_2^{\mathrm{LT}})$ by numerically solving the following material balance equation:

$$\left(\int_0^{z_1}(B(\kappa_2^{\mathrm{LT}})\cdot\Delta n_e^{\kappa_2^{\mathrm{LT}}}(z)+D(\kappa_2^{\mathrm{LT}}))\,dz\right.$$

$$+\int_{z_1}^{z_2}(B(\kappa_1^{\mathrm{LT}})\cdot\Delta n_e^{\kappa_1^{\mathrm{LT}}}(z)+D(\kappa_1^{\mathrm{LT}}))\,dz$$

$$\left.+\int_{z_2}^{\infty}B(\alpha)\cdot\Delta n_e^{\alpha}(z)\,dz\right)_{\kappa_2^{\mathrm{LT}}}=\left(\int_0^{\infty}B(\alpha)\cdot\Delta n_e^{\alpha}(z)\,dz\right)_{\alpha}. \tag{13}$$

Here z_1 and z_2 are the coordinates of the $\kappa_2^{\mathrm{LT}}-\kappa_1^{\mathrm{LT}}$ and $\kappa_1^{\mathrm{LT}}-\alpha$ phase boundaries, respectively (Fig. 4b). These values are determined from the experimental refractive-index profiles under the conditions $\Delta n_e(z_1)=0.075$ and $\Delta n_e(z_2)=0.025$. A similar approach was used to determine the hydrogen concentration in the high-temperature modifications of the κ_1 and κ_2 phases: κ_1^{HT} and κ_2^{HT}. We note only some specific features. In an earlier study[34] we showed that when proton-exchanged waveguides are annealed at temperatures above 400 °C, lithium oxide evaporates from the surface of the wafers, and, as a result, a lithium-oxide-depleted $LiNb_3O_8$ phase forms in the region of the original waveguide where $\Delta n_e<0.037$ (Fig. 5a). Long-term annealing leads to the appearance of $LiNb_3O_8$ at the surface of the wafer, causing destruction of the surface.[34] However in regions with $\Delta n_e>0.037$ no $LiNb_3O_8$ forms even in samples annealed at $T>400$ °C (Fig. 5b).[34] This phase likewise does not form in proton-exchanged $LiNbO_3$ waveguides annealed at temperatures below 340 °C. These proton-exchanged $LiNbO_3$ waveguides with $\Delta n_e<0.075$ are homogeneous, and contain only the κ_1^{HT} phase (Fig. 5b). Note that unlike κ_1^{LT}, the κ_1^{HT} phase has the lowest proton concentration; therefore, $D(\kappa_1^{\mathrm{HT}})=0$, and for the κ_1^{HT} phase we can write $x=B(\kappa_1^{\mathrm{HT}})\cdot\Delta n_e$.

The experiments show that the area under the $\Delta n_e(z)$ profile begins to decrease abruptly as soon as the monoclinic $LiNb_3O_8$ phase appears in the sample. However, for cases where a waveguide initially containing the β_1-$H_xLi_{1-x}NbO_3$ phase is annealed at $T>400$ °C to form a still single-phase κ_1^{HT}-$H_xLi_{1-x}NbO_3$ waveguide, in which $\Delta n_e>0.037$ over the entire region initially occupied by the β_1 phase, we can write the following material balance condition:

$$B(\beta_1)\cdot S(\beta_1)=B(\kappa_1^{\mathrm{HT}})\cdot S(\kappa_1^{\mathrm{HT}}), \tag{14}$$

where $S(\beta_1)$ and $S(\kappa_1^{\mathrm{HT}})$ are the areas under the $\Delta n_e(z)$ profiles for the same sample measured in the annealing stages when the β_1 and κ_1^{HT} phases form at the waveguide surface, respectively.

A similar procedure was used to determine $B(\kappa_2^{\mathrm{HT}})$ and $D(\kappa_2^{\mathrm{HT}})$. Since the dependence of Δn_e on Δn_0 for all the phases is known,[13,15] we can plot Δn_0 as a function of x.

In this approach we assumed a linear relationship between the extraordinary-refractive-index increment and the proton concentration in the phases containing no interstitial protons, viz., α-, κ_1-, κ_2-, and β_1-$H_xLi_{1-x}NbO_3$. This assumption seems quite justified since all the protons are ''optically active.''[21] However, in the β_2-, β_3-, and β_4-$H_xLi_{1-x}NbO_3$ phases the protons occupy both optically active positions in the oxygen planes and optically inactive interstitial positions. In this case, the dependence of Δn_e on x is clearly nonlinear even within a single phase. However, since the range of concentrations in which the phases exist is narrow, using a linear dependence of Δn_e on x as a first approximation does not lead to any serious errors.

Let us consider a waveguide which is formed in a benzoic acid melt and thus contains the β_2 phase on the surface, where $\Delta n_e^0=0.11$. Annealing at 330 °C for several minutes causes a $\beta_2\rightarrow\beta_1$ phase transition and increases Δn_e^0 (Refs. 35–37). After a complete $\beta_2\rightarrow\beta_1$ phase transition has taken place, further annealing causes Δn_e^0 in the homogeneous β_1 phase to begin to decrease, and for a waveguide with $\Delta n_e^0=0.011$ there is a 30% increase in the area under the $\Delta n_e(z)$

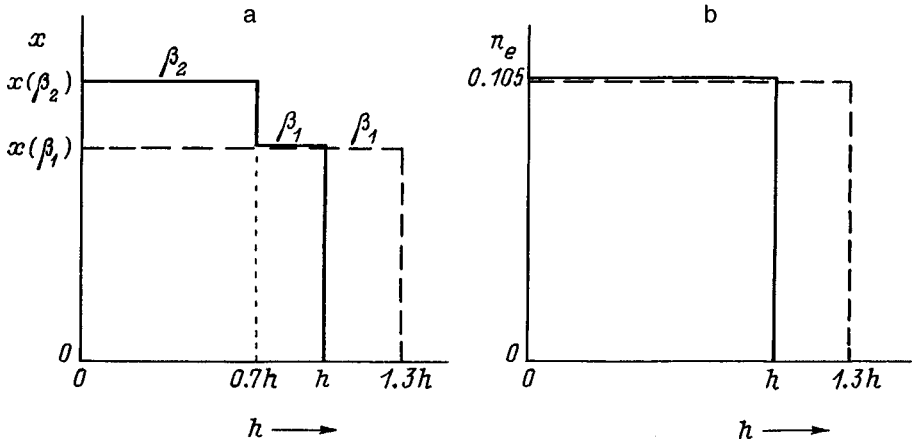

FIG. 6. Transformation of the proton concentration profile (a) and the refractive-index profile (b) as a function of annealing time for a β_2/β_1 proton-exchanged $LiNbO_3$ waveguide.

profile (and thus in the waveguide depth since the refractive-index profile continues to remain stepped). The successive recording of rocking curves using precision polishing of the initial β_2/β_1 sample shows that the β_2 phase occupies approximately 70% of the total waveguide volume (the remaining 30% is occupied by the β_1 phase) (Fig. 6). A similar result was obtained using optical measurements. In fact, in a fairly deep ($h \geqslant 3\ \mu$m) two-layer β_2/β_1 structure on *X*- and *Z*-cut $LiNbO_3$, not only extraordinarily polarized waveguide modes but also ordinarily polarized modes are excited (TE modes on the *X*-cut and TM modes on the *Z*-cut samples) with $n_{\text{eff}} \approx 2.25$, which is substantially lower than the ordinary refractive index of the $LiNbO_3$ substrate.[38,39] Subsequently, similar modes in *X*-cut $LiNbO_3$ fabricated in a benzoic acid melt were also observed by Rickerman *et al.*[40] It was noted that these modes have a very high absorption coefficient[40] ($\alpha > 2\ \text{mm}^{-1}$). The existence of ordinarily polarized modes can easily be explained by the fact that in a two-layer structure with $\Delta n_e(\beta_1) = \Delta n_e(\beta_2)$ the condition $n_0(\beta_2) > n_0(\beta_1)$ is easily satisfied although $n_0(\beta_2) < n_0(\text{LiNbO}_3)$ and $n_0(\beta_1) < n_0(\text{LiNbO}_3)$.[13,15] In this case, the β_1 phase layer acts as the substrate for the ordinarily polarized mode (Fig. 7). The high optical losses are caused by ''leakage'' of the mode into the substrate because of the small thickness of the β_1 layer (Fig. 7).

Three TE modes are observed in a two-layer β_2/β_1 waveguide on a *Z*-cut slab with a total thickness of 7 μm. The refractive-index profile reconstructed from the measured effective refractive indices is stepped with a step depth of 5 μm (Fig. 7), which amounts to $\approx 70\%$ of the total waveguide depth. Thus, the optical measurements confirm the data on the depth of the layers obtained by precision polishing accompanied by the recording of rocking curves. We now have at our disposal the information required to write the material balance equation:

$$x(\beta_2)0.7h + x(\beta_1)0.3h = x(\beta_1)1.3h \tag{15}$$

or $x(\beta_2) = 1.43x(\beta_1)$ for $\Delta n_e^0(\beta_2) = 0.11$. It is quite clear that $x(\beta_2) = x(\beta_1)$ for $\Delta n_e^0 = 0.125$. We obtained two extreme points on the plot of Δn_e as a function of *x* for the β_2 phase. By drawing a straight line between them, we obtain

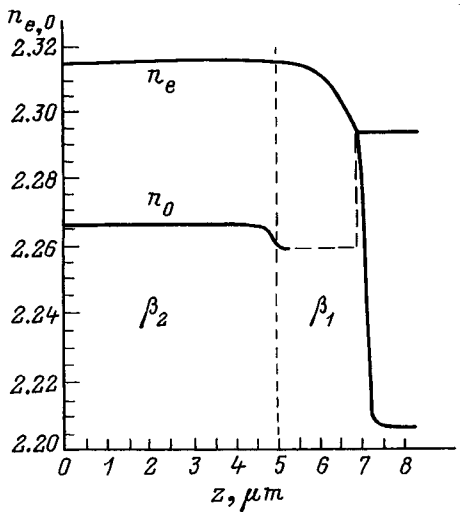

FIG. 7. Extraordinary- and ordinary-refractive-index profiles of a two-layer β_2/β_1-$H_xLi_{1-x}NbO_3$ waveguide structure obtained using a $KHSO_4$ solution in glycerol (4 g/l) at $T = 215\,^\circ$C for 80 h.

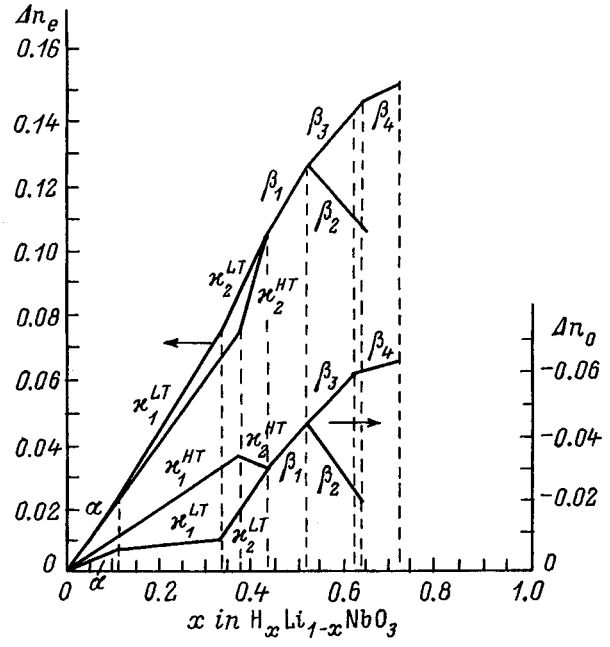

FIG. 8. Extraordinary and ordinary refractive indices as a function of the proton concentration in various $H_xLi_{1-x}NbO_3$ phases.

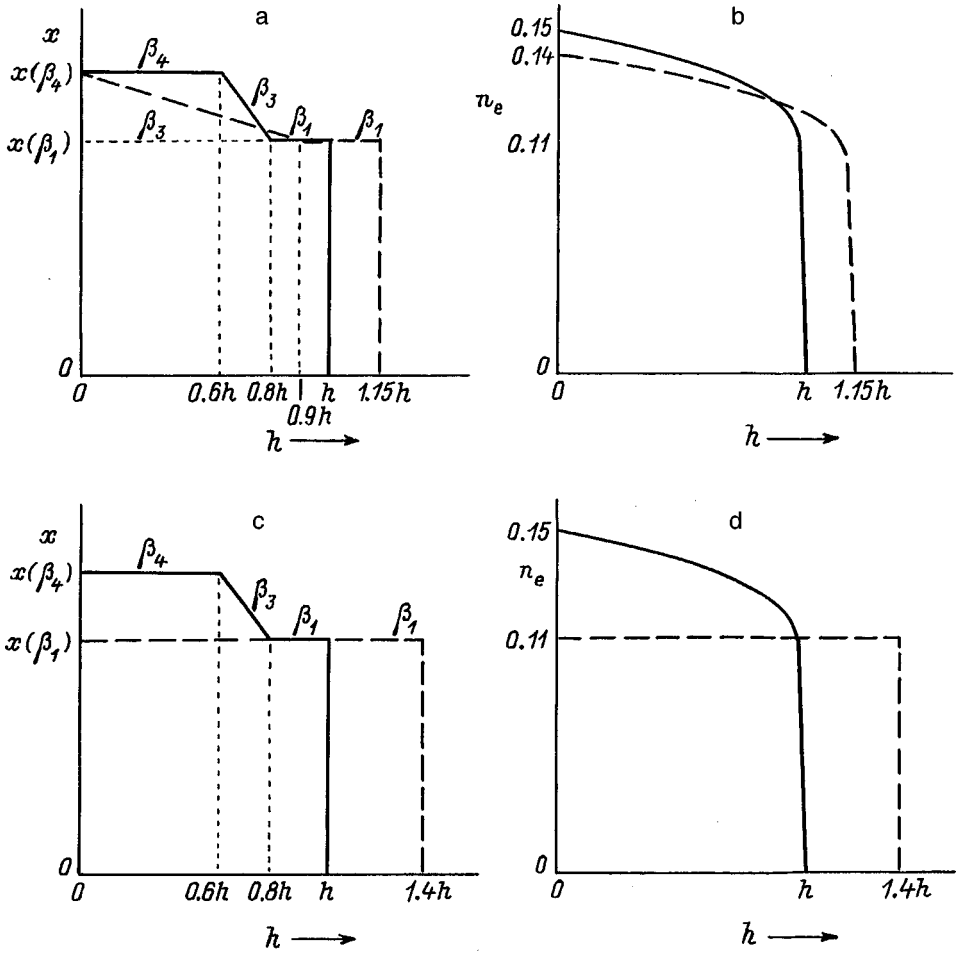

FIG. 9. Transformation of the proton concentration profiles (a, c) and the refractive-index profile (b, d) as a function of annealing time for a $\beta_4/\beta_3/\beta_1$ proton-exchanged $LiNbO_3$ waveguide: a, b — first anneal, c, d — second anneal.

an approximate dependence of the refractive index on total proton concentration for the β_2 phase (Fig. 8).

A similar procedure was used to estimate the proton concentration in the β_4 and β_3 phases. The spatial configuration of the phase layers was determined by precision polishing and recording the rocking curves after each polishing step. Polishing a waveguide with $\Delta n_e^0 = 0.15$ fabricated in an ammonium dihydrogen phosphate melt and then recording the rocking curves showed that the β_4 phase occupies approximately 60% of the total waveguide depth, the β_3 phase occupies 20%, and the remaining 20% is filled with the β_1 phase (Figs. 9a and 9b). Polishing also shows that the region occupied by the β_4 phase is almost homogeneous since Δn_e and the strains ε''_{33} do not change during polishing, as long as the β_4 phase remains on the surface. The region occupied by the β_1 phase is also homogeneous, but the refractive index and, accordingly, the proton concentration vary in the region occupied by the β_3 phase. As a first approximation, we assume that the functions $\Delta n_e(z)$ and $x(z)$ are linear in the β_3 phase.

Annealing this waveguide until the β_4 phase disappears completely increases the waveguide thickness by 15% (Figs. 9a and 9b). In this case, the β_3 phase occupies approximately 80% of the total depth (Figs. 9a and 9b). Thus, the material balance equation may be written as follows:

$$x(\beta_4)0.6h + \frac{x(\beta_4)+x(\beta_1)}{2}0.2h + x(\beta_1)0.2h = 1.15\left(\frac{x(\beta_3)+x(\beta_1)}{2}0.8h + x(\beta_1)0.2h\right). \quad (16)$$

Further annealing until a single-phase β_1 waveguide is formed with $\Delta n_e^0 = 0.11$ increases the waveguide thickness by 40% compared with the initial waveguide (Figs. 9c and 9d). The material balance equation has the form:

$$x(\beta_4)0.6h + \frac{x(\beta_4)+x(\beta_1)}{2}0.2h + x(\beta_1)0.2h = 1.4x(\beta_1)h, \quad (17)$$

and for the β_4 phase with $\Delta n_e = 0.15$ we obtain $x = 0.71$. From Eq. (16) we obtain $x(\beta_3) = 0.66$ for $\Delta n_e = 0.145$.

By drawing straight lines between these points, we obtain an approximate dependence of the extraordinary refractive index on proton concentration for the β_2, β_3, and β_4 phases, which contain interstitial protons (Fig. 8). Knowing the dependence of Δn_e on Δn_0 (Refs. 13 and 15), we can construct the dependences of Δn_0 on x (Fig. 8).

A possible explanation for the phase diversity observed may be that the coherent conjugation between the incipient proton-exchanged layers and the lithium niobate substrate

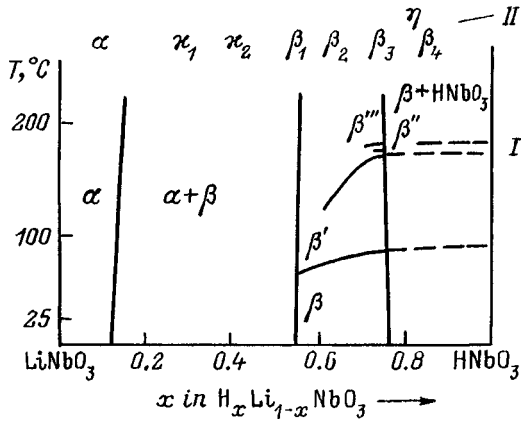

FIG. 10. Phase diagrams of equilibrium[3] (*I*) and strained (*II*) $H_xLi_{1-x}NbO_3$ solid solutions.

and the differences between their lattice parameters give rise to an excess elastic energy compared with the equilibrium state, which splits the equilibrium β phase observed in powders[2,3] into four different phases, viz., β_1, β_2, β_3, and β_4 (Fig. 10). The difference between the lattice parameters of the equilibrium α-$H_xLi_{1-x}NbO_3$ phase and $LiNbO_3$ is extremely small; therefore, we can postulate that the α phase created in proton-exchanged waveguides is the equilibrium α phase studied in Refs. 2 and 3. In the concentration range corresponding to an $(\alpha+\beta)$ phase mixture in an equilibrium unstrained system (in powders), κ_1 and κ_2 phases having various temperature modifications are formed in strained waveguide layers (Fig. 10).

In addition to the rhombohedral $H_xLi_{1-x}NbO_3$ phases, we also estimated the proton concentration in the monoclinic η phase, which forms on the surface of $(0\bar{1}4)$-cut $LiNbO_3$ treated in highly acidic melts, such as pyrophosphoric acid or ammonium dihydrogen phosphate. In the first stage of proton exchange in these melts (when the waveguide thickness is less than 1.5 μm), the rocking curves reveal only one additional peak, which corresponds to a homogeneous proton-exchanged layer containing only the η phase. As the proton-exchange time increases, the stresses at the layer/substrate interface begin to relax by forming a sublayer consisting of the β_1 phase.[16]

In order to estimate the proton concentration, we produced a homogeneous waveguide containing only the η phase ($h=1.4$ μm). Annealing this structure at 220 °C for 15 h causes a complete $\eta\rightarrow\beta_1$ phase transition, and in this case the balance equation is

$$x(\eta)h(\eta)=A(\beta_1)S(\beta_1), \tag{18}$$

where $h(\eta)$ is the depth of the initial waveguide and $S(\beta_1)$ is the area under the $\Delta n_e(z)$ curve for the sample annealed to the β_1 phase.

Here we used such a low annealing temperature because when waveguides on $(0\bar{1}4)$-cut slabs containing the η phase were annealed at higher temperatures, for some as yet unknown reasons, the crystallographic planes rotated appreciably.[16]

Experiments carried out using ten samples prepared in melts of ammonium dihydrogen phosphate and pyrophosphoric acid give $x(\eta)=0.65\pm0.02$. Note that this value is lower than the hydrogen concentration in the rhombohedral β_4 phase.

3. CONCLUSIONS

To conclude, we note that the proposed method of determining the proton concentration in the various $H_xLi_{1-x}NbO_3$ phases formed in proton-exchanged optical waveguides in lithium niobate crystals has been used for the first time to determine the concentration limits of the phases and to establish the concentration dependence of the refractive index in the various crystalline phases. Knowledge of the structural phase diagram of the waveguide layers permits prediction of the properties of integrated-optics devices based on proton-exchanged $LiNbO_3$ waveguides.

[1] J. L. Jackel, C. E. Rice, and J. J. Veselka, Appl. Phys. Lett. **41**, 607 (1982).
[2] C. E. Rice and J. L. Jackel, Mater. Res. Bull. **19**, 591 (1984).
[3] C. E. Rice, J. Solid State Chem. **64**, 188 (1986).
[4] V. A. Fedorov and Yu. N. Korkishko, Ferroelectrics **160**, 185 (1994).
[5] V. A. Fedorov and Yu. N. Korkishko, Proc. SPIE **2150**, 333 (1994).
[6] K. El Hadi, P. Baldi, S. Nouh *et al.*, Opt. Lett. **20**, 1698 (1995).
[7] V. A. Fedorov and Yu. N. Korkishko, Proc. SPIE **2291**, 243 (1994).
[8] Yu. N. Korkishko and V. A. Fedorov, IEEE J. Sel. Top. Quantum Electron. **2**(2), 187 (1996).
[9] Yu. N. Korkishko and V. A. Fedorov, Opt. Mater. **5**, 175 (1996).
[10] Yu. N. Korkishko, V. A. Fedorov, M. P. De Micheli *et al.*, Appl. Opt. **35**, 7056 (1996).
[11] M. P. De Micheli, D. B. Ostrowsky, Yu. N. Korkishko, and P. Bassi, *Insulating Materials for Optoelectronics* (World Scientific, 1995), Chap. 12, pp. 329–366.
[12] Yu. N. Korkishko and V. A. Fedorov, Ferroelectrics **183**, 245 (1996).
[13] Yu. N. Korkishko and V. A. Fedorov, Proc. SPIE **2700**, 186 (1996).
[14] S. Chen, P. Baldi, M. P. De Micheli *et al.*, IEEE J. Lightwave Technol. **LT-12**(5), 862 (1994).
[15] Yu. N. Korkishko and V. A. Fedorov, *Ion Exchange in Single Crystals for Integrated Optics and Optoelectronics* (Cambridge International Science Publishers, 1996).
[16] Yu. N. Korkishko and V. A. Gan'shin, Zh. Tekh. Fiz. **58**, 692 (1988) [Sov. Phys. Tech. Phys. **33**, 423 (1988)].
[17] M. Howerton, W. K. Burns, P. R. Skeath, and A. S. Greenblatt, IEEE J. Quantum Electron. **QE-27**, 593 (1991).
[18] S. W. Wilson, S. W. Novak, J. M. Zavada *et al.*, J. Appl. Phys. **66**, 6055 (1989).
[19] R. Richter, T. Bremer, P. Hertel, and E. Kratzig, Phys. Status Solidi A **114**, 765 (1989).
[20] J. M. Zavada, H. C. Casey Jr., Ch. Ho. Chen, and A. Loni, Appl. Phys. Lett. **62**, 2769 (1993).
[21] J. M. Zavada, H. C. Casey, R. J. States *et al.*, J. Appl. Phys. **77**, 2697 (1995).
[22] A. Loni, R. M. De La Rue, J. M. Zavada, and S. W. Wilson, Electron. Lett. **27**, 1245 (1991).
[23] K. Ito and K. Kawamoto, Jpn. J. Appl. Phys. **31**, 3882 (1992).
[24] C. Canali, A. Carnera, G. Della Mea *et al.*, J. Appl. Phys. **59**, 2643 (1986).
[25] A. Loni, R. M. De La Rue, and J. M. Winfield, J. Appl. Phys. **61**, 64 (1987).
[26] A. Loni, R. M. De La Rue, M. A. Foad, and J. M. Winfield, in *Proceedings of the IGWO*, 1989, pp. 258–262.
[27] M. Rottschalk, T. Bachmann, and A. Witzmann, Nucl. Instrum. Meth. Phys. Res., Sect. B **61**, (1991).
[28] I. Savatinova, S. Tonchev, and M. Kuneva, Appl. Phys. A: Solids Surf. **56**, 81 (1993).

[29] C. Ziling, L. Pokrovski, N. Terpugov *et al.*, J. Appl. Phys. **73**, 3125 (1993).
[30] G. R. Paz-Pujalt, D. D. Tuschel, G. Braunstein *et al.*, J. Appl. Phys. **76**, 3981 (1994).
[31] S. T. Vohra, A. R. Mickelson, and S. E. Asher, J. Appl. Phys. **65**, 1429 (1989).
[32] J. Nikolopoulos and G. L. Yip, Proc. SPIE **1583**, 71 (1991).
[33] J. Nikolopoulos and G. L. Yip, IEEE J. Lightwave Technol. **LT-9**, 864 (1991).
[34] V. A. Ganshin and Yu. N. Korkishko, Opt. Commun. **86**, 523 (1991).
[35] Yu. N. Korkishko, V. A. Fedorov, and M. P. De Micheli, Electron. Lett. **31**, 1603 (1995).
[36] Yu. N. Korkishko and V. A. Fedorov, Pis'ma Zh. Tekh. Fiz. **21**(18), 61 (1995) [Tech. Phys. Lett. **21**(9), 752 (1995)].
[37] V. V. Atuchin, C. C. Ziling, I. Savatinova, M. N. Armenise, and V. M. N. Passaro, J. Appl. Phys. **78**, 6936 (1995).
[38] V. A. Ganshin and Yu. N. Korkishko, J. Opt. Commun. **13**(1), 2 (1991).
[39] V. A. Gan'shin and Yu. N. Korkishko, Zh. Tekh. Fiz. **62**(1), 98 (1992) [Sov. Phys. Tech. Phys. **37**, 49 (1992)].
[40] F. Rickerman, D. Dip, B. Gather, and E. Krätzig, Phys. Status Solidi A **150**, 763 (1995).

Translated by R. M. Durham

Linear intermode conversion of electromagnetic wave energy in a magnetically soft gyrotropic material

A. S. Antonov, A. N. Lagar'kov and I. T. Yakubov

Scientific-Research Center for Applied Problems in Electrodynamics, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted May 27, 1997; resubmitted February 26, 1998)
Zh. Tekh. Fiz. **69**, 58–63 (March 1999)

The excitation of axial radio-frequency (rf) magnetic induction by an axial rf current is observed in a conductor with circular magnetic anisotropy when a weak magnetizing field is applied. The conductor is an amorphous cobalt-based wire, which exhibits azimuthal magnetic anisotropy. It serves as the central conductor in a coaxial line. The axial rf magnetic induction produces an emf in an induction coil coaxial to the conductor. The induction coil is part of a matched receiving circuit. The power conversion coefficient is as high as tens of percent. The measurements demonstrate the high sensitivity of the conversion coefficient to an external field. The theory of ferromagnetic resonance faithfully describes the results of the observations.
 [S1063-7842(99)01003-X]

1. INTRODUCTION

If a gyrotropic wire is the central conductor of a coaxial line, the structure of the fields in the gap within the coaxial line is a combination of the fields of quasi-TE and quasi-TM waves.[1] In fact, a radio-frequency (rf) current passed through an anisotropic ferromagnetic wire magnetized by a longitudinal field can excite not only a circular rf magnetic field, but also a commensurate axial field. For this to occur, the precessing magnetic moment vector of the sample must have both a longitudinal component and a transverse component with respect to the wire axis. Then the axial component of the rf magnetic moment appears. It can be detected experimentally in the form of an emf induced in a coil wound about the wire. Since the ratio between the fields of the waves depends on the mean direction of the precessing magnetic moment in the wire, the observed effect depends on external field.

This idea can be realized best when a conductor with circular magnetic anisotropy is employed. Such conductors include amorphous wires based on cobalt with a small negative magnetostriction constant (see, for example, Ref. 2). Such a wire exhibits a magnetic moment component in fields commensurate with the anisotropy field. In the present work an amorphous ferromagnetic wire with circular anisotropy was the central conductor of a short-circuited coaxial line. The coefficient for the transmission of power from the coaxial line to an induction coil coaxial to the conductor was measured as a function of the constant external magnetic field applied along the wire.

It is significant that in weak magnetizing fields the measured conversion coefficient exhibits high sensitivity to the field amplitude. This is a manifestation of ''giant magnetic impedance,'' which has been widely investigated in connection with problems in creating highly sensitive magnetic-field sensors, magnetic recording devices, etc.[3,4] Giant magnetic impedance is based on the high sensitivity of the impedance of a ferromagnetic conductor to the magnetizing field because of the skin effect.

We believe that the effect observed can serve as the basis for creating a highly sensitive magnetic-field sensor. The effect that we measured is caused by precession of the magnetic moment vector and is strictly linear with respect to the rf power supplied to the conductor. It differs fundamentally from the Matteuci effect and its inverse, viz., the Wiedemann effect, in which the magnetization of the sample is reversed along a complete hysteresis loop,[2] i.e., with a fairly large amount of power supplied. These effects, which are displayed under the action of a longitudinal external variable field in the former case and in the circular field of a flowing alternating current in the latter case, are nonlinear.

This paper is composed according to the following plan. The results of measurements of the conversion coefficient as a function of the external constant magnetic field are presented first with a description of the basic scheme of the experimental setup. This is followed by a treatment employing the theory of ferromagnetic resonance phenomena, which yielded a formula for the conversion coefficient. The theory satisfactorily describes the measurement results.

2. EXPERIMENTAL RESULTS

Figure 1 presents the system for measuring the coefficient for the transmission of the power of electromagnetic modes excited in coaxial line *1* with gyrotropic conductor *3* of length 6 mm and diameter 30 μm to measuring coil *5*. The wire was an amorphous ferromagnetic conductor with the composition $(Fe_6Co_{94})_{72.5}Si_{12.5}B_{15}$ from Unitika Ltd., which was prepared by rapid cooling from a melt. The magnetic properties of amorphous conductors based on Fe and Co are usually described within a model which presumes that an amorphous conductor has two regions, viz., a near-axis region (core) and a shell, with differently directed easy anisotropy axes.[5,6] For the amorphous Co-based conductor with

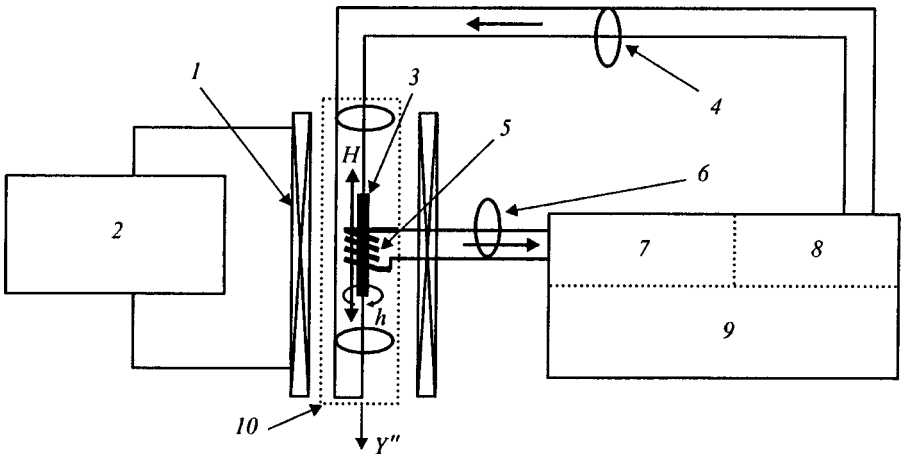

FIG. 1. System for measuring the conversion coefficient: *1* — coil of an electromagnet; *2* — stabilized current source; *3* — amorphous wire; *4, 6* — coaxial lines; *5* — measuring coil; *9* — P4-37 complex transmission coefficient meter combined with generator *7* and indicator *8*; *10* — measuring cell.

a negative magnetostriction constant considered in our case, the body of experimental data from different investigators provides evidence that there is a region near the wire axis that is uniformly magnetized along the axis, while the magnetization is directed azimuthally in the outer shell. We shall bear in mind this model in our treatment of ferromagnetic-resonance phenomena in a wire.

The effect described was observed over the entire frequency range of the panoramic complex transmission coefficient meter *9* used in the experiment from 0.5 to 1250 MHz. A frequency of 50 MHz was chosen for the analysis of the experimental data. Measuring coil *5* with a length of 3 mm and a diameter of 1 mm had 15 turns of copper wire. The power of the signal in coaxial line *2* excited by the measuring coil was displayed by the indicator of a P4-37 instrument. Measuring cell *10* was placed in the uniform constant magnetic field created by coil *1*. Coil *1* was powered by stabilized current source *2*, by which we could vary the field to within 10^{-4} Oe. The direction of the constant magnetic field, which amounted to 0.32 Oe at the site of the sensitive element, i.e., the wire, was determined preliminarily, and the wire was oriented along that direction. This was done so that the stray field could be compensated by the coil field. The field in the coil could be varied in the range from -100 to 100 Oe. The reference power of the electromagnetic modes excited in coaxial line *1* did not exceed 10 μW.

Figure 2 presents the results from measuring the transmission coefficient $T=10\log P/P_0$, where P and P_0 are the powers of the recorded and reference signals, respectively, as a function of the external magnetic field applied along the wire. The field was varied smoothly from the maximum value of one direction to the maximum value of the opposite direction and then in the reverse order. The measurements were performed for two values of the reference signal power differing by 100 fold. The curves practically coincide, attesting to the linearity of the conversion of the current in the wire into coil current. As the field is varied in the range from approximately -1 to 1 Oe, the magnetization-reversal process has a reversible character, i.e., it is realized by rotation of the magnetic moment. When the field range is increased, a single jump in the longitudinal magnetization takes place near 1.3 Oe. When a level of $8-10$ Oe is achieved, the system reaches saturation, i.e., the magnetic moment is oriented strictly along the sample axis.

3. MAGNETIC PERMEABILITY TENSOR

As was indicated above, the Unitika wire is a magnetically soft amorphous object based on cobalt with a small negative magnetostriction constant, the structure of whose near-surface layer is characterized by circular anisotropy due to the residual quenching stresses created during its preparation. The effective anisotropy field in this layer is created both by the circular anisotropy indicated and the anisotropy of the sample shape. When an axial magnetic field **H** is applied, the orientation of the magnetic moment $\mathbf{M}_0$ can be determined by minimizing the free energy U with respect to the angle Θ formed by $\mathbf{M}_0$ and the easy axis (Fig. 3). In the case under consideration the free energy is represented in the form of the sum of the structural anisotropy energy U_A, the Zeeman energy U_Z, and the demagnetizing field energy U_M. We write U, allowing a possible deviation of the easy axis by an angle ψ from the direction normal to the wire axis:

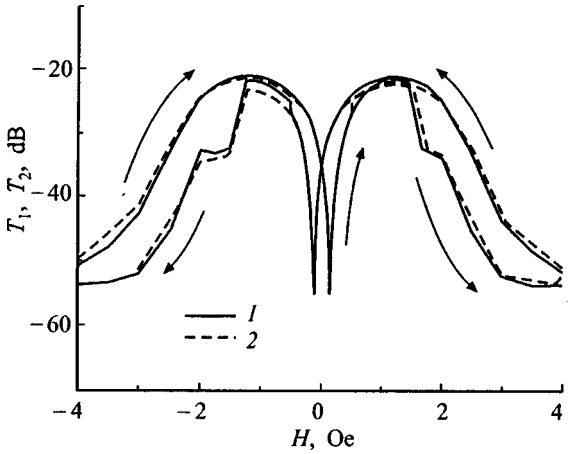

FIG. 2. Measured dependence of the power transmission coefficient on the strength of the external constant magnetic field: T_1 — solid line ($P_0=10$ μW), T_2 — dashed line ($P_0=0.1$ μW). The arrows point out the sequence of variation of the field upon reversal of the magnetization.

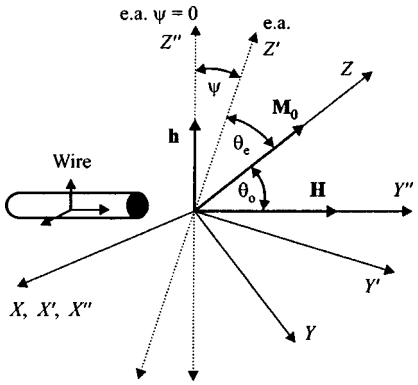

FIG. 3. Coordinate axes considered in the film model: Z' — easy axis deflected by the angle ψ from the azimuthal direction in the wire Z''.

$$U = U_A + U_Z + U_M = K_1 \sin^2\Theta - M_0 H \sin(\Theta + \psi) + \frac{1}{2} M_0 N_Z \sin^2(\Theta + \psi).$$

Here K_1 is the anisotropy constant, and N_Z is the demagnetization factor. Setting $\partial U/\partial\Theta$ equal to zero, we obtain the equation for the equilibrium angle Θ_e

$$\sin 2\Theta_e + \frac{M_0^2 N_Z}{2K_1} \sin 2(\Theta_e + \psi) = \frac{M_0 H}{K_1} \cos(\Theta_e + \psi). \tag{1}$$

We introduce the notations $H_A = H_K + N_Z M_0$ and $H_K = 2K_1/M_0$. If $\psi = 0$, then $\sin\Theta_e = H/H_A$ for $H \leqslant H_A$ and $\sin\Theta_e = 1$ for $H \geqslant H_A$.

Let us estimate the influence of the form factor N_Z on the value of the equilibrium angle Θ_e. For the wires which we used $K_1 \approx 125$ erg/cm^3, and the saturation magnetization $M_0 \approx 500$ G.[2,3] With consideration of the ratio between the length and diameter of the wire $n = l/d = 200$, the contributions of structural anisotropy H_K and the demagnetizing field to H_A, which influence the equilibrium angle, are approximately identical: $H_K \approx N_Z M_0 \approx 0.5$ Oe.

At our frequencies of the variable magnetic field h the thickness of the skin layer δ, which is specified by the conductivity of the conductor $\sigma = 6.7 \times 10^{15}$ s^{-1}, but calculated without consideration of its magnetic permeability, is comparable to the radius of the wire. If we, nevertheless, assume that the inequality $\delta \ll a$ holds, we can solve the problem of the magnetic susceptibility tensor of the surface layer of the wire, treating it as a ferromagnetic field with uniaxial anisotropy lying in its plane. Such an approximation can be considered well founded for magnetizing fields greater than a certain strength, at which the magnetic permeability, reaching values of $10^3 - 10^4$, significantly reduces the thickness of the skin layer, and, moreover, just these conditions are of practical interest.

Now, in a first approximation we assume that the demagnetizing factors for an rf field are the same as for a thin film, while the factors for a constant magnetizing field are the same as for an oblong-ellipsoidal wire. We can utilize the method of effective magnetizing factors, for which it is now convenient to direct the Z axis along the equilibrium magnetization (Fig. 3).

Thus, a simple model can be considered. It is assumed that there is no domain structure in our short samples. In fact, the experimental data provide evidence that the length of our samples is much shorter than the ''critical'' length needed for the formation of a domain structure.[2,7] This length amounts to several centimeters.

In the $\{X,Y,Z\}$ coordinate frame considered the magnetic susceptibility tensor has the general form[1]

$$\underset{\leftrightarrow}{\chi} = \begin{vmatrix} \chi_x & i\chi_a & 0 \\ -i\chi_a & \chi_y & 0 \\ 0 & 0 & \chi_z \end{vmatrix}, \tag{2}$$

$$\chi_{x,y} = \frac{\omega_M \omega_{x,y}}{\omega_0^2 - \omega^2}, \quad \chi_a = \frac{\omega_M \omega}{\omega_0^2 - \omega^2}, \tag{3}$$

$$\omega_x = \omega_H + \gamma N_{22} M_0, \quad \omega_y = \omega_H + \gamma N_{11} M_0, \tag{4}$$

$$\omega_H = \gamma H - \gamma N_{33} M_0 - i\alpha\omega, \quad \omega_M = 4\pi\gamma M_0,$$

$$\omega_0^2 = \omega_x \omega_y. \tag{5}$$

The parameter α specifies the dissipation, and γ is the gyromagnetic ratio. The effective demagnetizing factor tensor can be represented in the form $\underset{\leftrightarrow}{\mathbf{N}} = \underset{\leftrightarrow}{\mathbf{N}}^F + \underset{\leftrightarrow}{\mathbf{N}}^A$, where $\underset{\leftrightarrow}{\mathbf{N}}^F$ is responsible for demagnetization across the skin layer and $\underset{\leftrightarrow}{\mathbf{N}}^A$ is responsible for the structural anisotropy and the shape anisotropy due to the finite width of the film (which is equal to length of the wire l). The corresponding anisotropy field is

$$\mathbf{H}_A' = \underset{\leftrightarrow}{\mathbf{N}}^A \mathbf{M}_0. \tag{6}$$

In the $\{X,Y,Z\}$ coordinate frame the components of $\underset{\leftrightarrow}{\mathbf{N}}$ have the form

$$N_{11} \approx 4\pi, \quad N_{22} = -(H_A'/M_0)\sin^2\Theta_e,$$

$$N_{33} = (H_A'/M_0)\cos^2\Theta_e.$$

The value of H_A' is close to H_A, and we shall disregard the difference between them. In our experiment the measuring coil picks up the variation of the induction of the axial magnetic field $b_{y''} = \mu_{y''z''} h_{z''}$ with time, which is the response to the rf magnetic field directed perpendicularly to the external constant field, i.e., along the Z'' axis in the $\{X'',Y'',Z''\}$ coordinate frame, which is associated with the axes of the ellipsoidal wire. Going over to the axes of the $\{X'',Y'',Z''\}$ system according to the rules for the transformation of tensor components and neglecting the small quantities, we write out the final result in the form of the corresponding components of the magnetic permeability

$$\mu_{y''z''} = 4\pi\chi_{y''z''} = \frac{\omega_M \sin\Theta_e \cos\Theta_e}{\omega_A \cos^2\Theta_e - i\alpha\omega - \omega^2/\omega_M}, \tag{7}$$

$$\mu_{z''z''} = 1 + 4\pi\chi_{z''z''} = 1 + \frac{\omega_M \sin^2\Theta_e}{\omega_A \cos^2\Theta_e - i\alpha\omega - \omega^2/\omega_M}. \tag{8}$$

Here $\omega_H=\gamma H$, $\omega_A=\gamma H_A$, and $\omega=2\pi f$. The ferromagnetic resonance frequency is specified by the following expression:

$$\omega_0^2=\omega_M\omega_A[1-(\omega_H/\omega_A)^2]. \tag{9}$$

The ''natural'' ferromagnetic resonance frequency (for $H=0$) $\tilde{\omega}_0=(\omega_H\omega_A)^{1/2}\cong 1.5\times 10^9$ rad/s. The hierarchy of characteristic frequencies of the problem has the form

$$\omega_M \geqslant \tilde{\omega}_0 \geqslant \omega_A \geqslant \omega_H. \tag{10}$$

The circular frequency of the rf field $\omega\approx 3\times 10^8$ rad/s significantly exceeds $\omega_A\approx 2.2\times 10^7$ rad/s. Therefore, the condition for ferromagnetic resonance at the frequency ω corresponds to a field H_0 slightly smaller than the anisotropy field H_A.

The inhomogeneity of the skin layer in the transverse direction strongly broadens the ferromagnetic resonance line.[8] The linewidth $\Delta H=(16\pi/3)(A\sigma\omega_A/c^2)^{1/2}$ amounts to about 0.4 Oe ($A=2\times 10^{-6}$ erg/cm is the exchange-coupling constant). Therefore, the ferromagnetic resonance line is strongly broadened and, therefore, was not resolved in the measurements.

The component $\mu_{y''z''}$ ($\mu_{z\varphi}$ in the cylindrical coordinate frame attached to the wire) is responsible for the field conversion under investigation. It is proportional to the product $\cos\Theta_e\sin\Theta_e$. Hence it follows that $\mu_{y''z''}$ vanishes when $H=0$ and when $H\geqslant H_A$. In fact, in the absence of the field H, the vector $\mathbf{M}_0$ is oriented along the Z'' axis, and the rf field $\mathbf{h}$ does not excite precession. In the other limiting case, where the magnetizing field H is great, the axial component of the variable magnetic moment $m_{y''}$ is equal to 0, since only $m_{z''}$ and $m_{x''}$ are excited.

The measured conversion coefficient is, in fact, vanishingly small when $H=0$, but in fields exceeding H_A it gradually decreases with increasing H rather than vanishes. In addition, when the magnetization is reversed, hysteresis is observed in a range of fields exceeding ± 1.5 Oe. These two experimental findings provided a basis for introducing the effective angle ψ, which is a measure of the deviation of the easy axis from the direction normal to the wire axis (Fig. 3), into the formulas describing magnetostatic equilibrium. The nature of this angle can vary. Technological factors which cause the easy axis to have just such an orientation cannot be ruled out. Another explanation can be based on features of the magnetic structure of a conductor of the composition under investigation, particularly on the existence of an inner core with longitudinal magnetization. The magnetic charges appearing on the ends of the wire create a scattered field, in which the magnetic moment of the shell acquires a direction differing from the circular even in the absence of an external field. Of course, this is a qualitative explanation. A rigorous treatment of the internal magnetic structure is associated with considerable difficulties and is beyond the scope of this paper. We note that to describe the measurement results it was sufficient to introduce a small angle ψ equal to -0.1 rad.

Both conversion schemes, viz., current in the wire into current in the coil and current in the coil into current in the wire, were realized in the experiment. The first scheme, which was considered above, required a calculation of the magnetic permeability $\mu_{z\varphi}$. In the second scheme a calculation of $\mu_{\varphi x}$ is needed. In this case the initiating rf field $\mathbf{h}$ is directed along the wire axis Z, and the rf magnetic moment m_φ is the parameter sought. An analysis similar to the one performed above reveals that

$$\mu_{\varphi z}=\mu_{z\varphi}. \tag{11}$$

This corresponds to the observation results.

Let us turn to the question of how the possible domain structure on the surface of the wire would be manifested. The method for describing ferromagnetic resonance in a multidomain sample takes into account the coupling of modes in neighboring domains, which is mediated by the demagnetizing fields appearing in the domain walls, and then averages the magnetization over many domains.[1] The ground state of a multidomain system is characterized by the fact that the magnetization vectors of neighboring domains $\mathbf{M}_{01}$ and $\mathbf{M}_{02}$, which are initially antiparallel to the Z'' axis, acquire identical components along the Y'' axis in the field $\mathbf{H}$. Precession in the field $\mathbf{h}=\mathbf{e}_{y''}h$ creates an rf magnetization $\mathbf{m}$. Averaging of the rf component $m_{z''}$ over many domains would lead to the disappearance of $h_{z''}$ and the emf induced by it. Therefore, our results together with the results presented in Ref. 2 provide evidence that the samples which we used are too short for the formation of a domain structure.

4. CONVERSION COEFFICIENT

In magnetizing fields that are weaker than the effective anisotropy field H_A the axial rf current i_z excites the mutually coupled fields h_φ, e_z and h_z, e_φ. These fields excite an emf in the coil coaxial to the wire. The magnitude of the emf is proportional to the component of the surface resistance tensor

$$(R_S)_{z\varphi}=\mathrm{Re}\,\zeta_{z\varphi}=\mathrm{Re}\sqrt{\omega\tilde{\mu}/4\pi i\sigma}\sin\vartheta_e\cos\vartheta_e. \tag{12}$$

Here ω is the circular frequency of the rf field, and $\zeta_{z\varphi}=\zeta_{y''z''}$ is the off-diagonal component of the surface impedance tensor.[9] The value of $\tilde{\mu}$ is specified using (7) by the relation

$$\mu_{z\varphi}=\mu_{y''z''}=\tilde{\mu}\sin\vartheta_e\cos\vartheta_e,$$

$$\tilde{\mu}=\frac{\omega_M}{\omega_A\cos^2\Theta_e-i\alpha\omega-\omega^2/\omega_M}.$$

The experimentally realized active part of the system, i.e., the segment of the ferromagnetic wire and the conductor short-circuiting the waveguide line, comprise a system which lacks azimuthal symmetry and requires a tedious description. For simplicity, we consider a short-circuited coaxial waveguide with a sample serving as the central conductor. After obtaining the conversion coefficient, we see that this simplification is justified.

The conversion coefficient is proportional to the ratio of the energy absorbed by the wire in the required mode to the energy absorbed in the waveguide:[10]

$$\xi_{z\varphi}=\frac{c}{4\pi}(R_s)_{z\varphi}I\int_L h_\varphi^2[\mathrm{Re}\,ZI^2]^{-1}. \quad (13)$$

Here $h_\varphi=h_\varphi(a)$ is the amplitude of the magnetic field strength on the surface of a wire of radius a. The integration is carried out over the contour L of a cross section of the wire. The integral in Eq. (13) equals $8\pi I^2/ac$, where I is the amplitude of the rf current. The energy absorbed in the waveguide and the source is proportional to the real part of the impedance $\mathrm{Re}\,Z\approx R_0+\mathrm{Re}\,Z_{\mathrm{inp}}$, where R_0 is the resistance of the coaxial transmission line, $Z_{\mathrm{inp}}=-iZ_c\tan(k_z l)$ is the input impedance of the short-circuited segment containing the sample, $Z_c=(2/c)(k_z/k_0)\ln(b/a)$ is the characteristic impedance, k_z is the longitudinal wave number, $k_0=\omega/c$, and b is the radius of the outer coaxial conductor. For our short wires $\tan(k_z l)\approx k_z l$. Taking this into account, we write the following expression for the conversion coefficient:

$$\xi_{z\varphi}=(R_s)_{z\varphi}k_0[Z_0-a\,\mathrm{Re}(ik_z)\ln(b/a)]^{-1}. \quad (14)$$

The lowest mode is close to a TM mode. The following expression for its longitudinal wave number was derived in Ref. 11 under the conditions of a strong skin effect, i.e., when $\delta\ll a$:

$$k_z^2=k_0^2\{1+(1+i)\,\delta_0\sqrt{\tilde{\mu}}\sin^2\vartheta_e[2a\ln(b/a)]^{-1}\}, \quad (15)$$

where $\delta_0=c/\sqrt{2\pi\sigma\omega}$. It was taken into account here that for large values of $\tilde{\mu}$ the diagonal component of the magnetic permeability $\mu_{\varphi\varphi}=\mu_{z''z''}\approx\tilde{\mu}\sin^2\vartheta_e$.

Then we obtain the final expression for the conversion coefficient

$$\xi_{z\varphi}=\mathrm{Re}[(1-i)\sqrt{\tilde{\mu}}\sin\vartheta\cos\vartheta]\times\{P+\mathrm{Re}[(1-i)\sqrt{\tilde{\mu}}\sin^2\vartheta]\}^{-1}, \quad (16)$$

where $P=(a/l)\sqrt{8\pi\sigma/\omega}\ln(b/a_0)$ and a_0 is the radius of the central rod in the coaxial transmission line.

The conversion coefficient depends weakly on the geometric configuration of the electrical circuit outside the wire. This dependence $[\propto\ln(b/a_0)]$ reflects the dependence of the characteristic impedance on the geometric dimensions which are transverse to the wave vector.

5. COMPARISON OF THEORY AND EXPERIMENT

If the ratio of the energy of the field induced in the measuring coil to the energy absorbed in the wire is construed as the conversion coefficient, then the theory gives a quantity equal to tens of percent. In the present experiment the signal from a load in a secondary circuit, which depends on the structure of that circuit, whose analysis is fairly complex, is recorded. However, we can take advantage of the fact that the structure of the secondary circuit does not significantly influence the relative dependence of the signal on magnetizing field. The relative dependence of the logarithm of the conversion coefficient $\xi_{z\varphi}$ calculated from Eq. (16) on external field is presented in Fig. 4 [$T_{z\varphi}=10\log(C\xi_{z\varphi})$, where C is a fitting parameter]. This dependence corresponds well to the measured dependence of T_1.

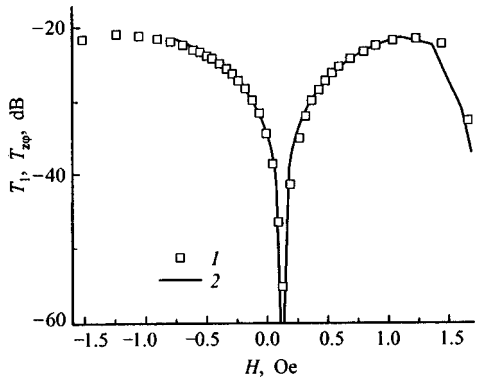

FIG. 4. Theoretical (*1*) and experimental (*2*) dependences of the power transmission coefficient on external field. Half of a magnetization-reversal cycle is shown; $\psi=-0.1$, $\alpha=0.05$, and $\omega/2\pi=50$ MHz.

At a zero value of H the calculated conversion coefficient vanishes. Although the measured value is very small (about -55 dB), it is still finite primarily because of disturbances from the primary circuit. The high-field portion of the theoretical curve faithfully reproduces the change in the character of the dependence appearing when the magnitude of the magnetizing field passes through the magnitude of the effective anisotropy field. For this reason, it was necessary to introduce the angle of deviation of the easy axis from the circular direction ψ.

Figure 4 presents only half of a magnetization-reversal cycle. The introduction of ψ naturally permitted the description of the hysteresis observed upon complete reversal of the magnetization (Fig. 2). In fact, upon demagnetization a transition to metastable states is unavoidable when the magnetic moment vector passes through the easy axis. Inasmuch as the value of ψ is primarily the result of the technological treatment of the sample, we shall regard it as a fitting parameter.

The dependence of the conversion coefficient on external field is very sharp at small fields. This can be of interest from the standpoint of creating rf magnetic field sensors.

We express our thanks to I. Morozov and T. Furmanova for their great assistance in setting up and performing the experiment.

This work was supported by the Russian Fund for Fundamental Research.

[1] A. G. Gurevich and G. A. Melkov, *Magnetization Oscillations and Waves*, CRC Press, Boca Raton (1996), 462 pp.
[2] M. Vazquez and A. Hernando, J. Phys. D **29**, 939 (1996).
[3] F. B. Humphrey, K. Mohri, J. Yamasaki *et al.*, *Magnetic Properties of Amorphous Metals*, A. Hernando, V. Madurga, M. S. Sanches, and M. Vazquez (Eds.), Elsevier, Amsterdam (1987).
[4] K. Mohri, F. B. Humphrey, K. Kawashima *et al.*, IEEE Trans. Magn. **MAG-26**, 1789 (1990).
[5] K. Mohri, K. Kozhawa, K. Kawashima, and L. V. Panina, IEEE Trans. Magn. **MAG-28**, 3150 (1992).
[6] R. S. Beach and A. E. Berkowitz, J. Appl. Phys. **64**, 3652 (1994).
[7] N. A. Usov, A. S. Antonov, A. M. Dykhne, and A. N. Lagar'kov, Élektrichestvo No. 2, 55 (1998).
[8] *Ferromagnetic Resonance; The Phenomenon of Resonant Absorption of a High-Frequency Magnetic Field in Ferromagnetic Substances*, S. V. Vonsovski (Ed.), Pergamon, Oxford–New York Press (1966), 343 pp.

[9] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon, Oxford (1984).
[10] L. M. Brekhovskikh, *Waves in Layered Media*, Academic Press, New York (1960), 343 pp.
[11] A. Antonov, A. Granovsky, A. Lagarkov *et al.*, Physica A (Amsterdam) **241**, 420 (1997).

Translated by P. Shelnitz

Ion-induced sputtering of a metal in the form of large clusters

V. I. Matveev

Department of Thermal Physics, Uzbekistan Academy of Sciences, 700135 Tashkent, Uzbekistan

S. F. Belykh and I. V. Berevkin

Institute of Electronics, Academy of Sciences of the Republic of Uzbekistan, 700143 Tashkent, Uzbekistan

(Submitted November 20, 1997)
Zh. Tekh. Fiz. **69**, 64–68 (March 1999)

A model is proposed for the ion-induced sputtering of a metal in the form of large clusters with a number of atoms $N \geqslant 5$. The model is based on simple physical assumptions and is consistent with experiment. As an example, calculations are made of the relative cluster yield as a function of the number of atoms in the cluster as a result of the bombardment of various metals by singly charged 5 keV argon ions. A comparison is made with experimental data.
 [S1063-7842(99)01103-4]

1. INTRODUCTION

Studies of the sputtering products of a metal under ion bombardment are interesting from both the fundamental and applied points of view. Applied aspects mainly involve technological applications in microelectronics and in space and fusion technologies. Fundamental aspects are associated with the development of the physical mechanisms for sputtering, which reflect the dynamics of the evolution of collision cascades in irradiated targets having a specific internal structure. Sputtering has been the subject of numerous publications (see Refs. 1 and 2, for example). Sputtering products consist of various numbers of target atoms, although the vast majority of these are usually isolated neutral atoms (see Refs. 3 and 4, for instance). The theory of sputtering based on an analysis of isolated target atoms is well-developed[1,2] and relies to a considerable extent on the so-called cascade sputtering mechanism proposed by Sigmund.[5] Mechanisms of sputtering in the form of two or more bound target atoms (clusters) are currently under discussion[1,2] since they do not satisfactorily describe the formation of large clusters with a number of atoms $N \geqslant 5$. Hopes of performing calculations ''from first principles'' now rest on computer simulation using molecular-dynamics methods (see the calculations reported in Ref. 6), but this of course requires a good knowledge of the interatomic and ion–atom interaction potentials and the structure of the solid. In addition, these calculations are technically highly complex, especially as the number of atoms in the cluster increases, and this necessitates making approximations which lead to insufficiently controllable results.

In the present paper we consider a model of metal sputtering in the form of large clusters with $N \geqslant 5$ based on simple physical assumptions, which agrees with experiment.

2. MODEL

Let us consider a body formed of atoms, each situated in an oscillator well of depth δ and having the natural frequency ω. In other words, we shall use a truncated oscillator potential.[1)] The characteristic oscillation period is $T = 2\pi/\omega$. Also, let the velocity of an incident ion be such that within the time $\tau \ll T$ the ion (or a fast recoil atom) undergoes a large number of collisions as it moves in the metal, and, as a result, the metal atoms acquire various momenta $\mathbf{q}_i$, where i labels the atom. The inequality $\tau \ll T$ allows us to formulate this differently: the result of an ion passing through a system of oscillators reduces to the instantaneous and simultaneous transfer of the momentum $\mathbf{q}_i$ to each oscillator ($i = 1,2,\ldots,N$), where N is the number of oscillators. Moving forward, we note that all the $\mathbf{q}_i (i = 1,2,\ldots,N)$ are assumed to be independent and all directions equally probable.

This block of N atoms may move as an entity or may separate as a cluster if: 1) all the atoms have the same average momentum $\mathbf{k}/N$ so that the total momentum of the cluster is $\mathbf{k}$; 2) the kinetic energy of the block as a whole is sufficient to overcome the binding energy of the cluster with the remaining metal; 3) the energy of the excited oscillators is insufficient to break up the block.

It is easiest to use a quantum-mechanical approach to calculate the probability of these events, which correspond to the correlated detachment of a block of atoms. The wave function of a system of N oscillators will be represented as the product of the N wave functions of the oscillators. The arbitrary state of a three-dimensional isotropic oscillator having the number i (ith atom) and the average momentum $\mathbf{k}/N$ will be described by the wave function[2)]

$$\exp\left(i \frac{\mathbf{k}}{N} \mathbf{r} \right) |n_{ix}, n_{iy}, n_{iz}\rangle. \tag{1}$$

The latter can be characterized by the three quantum numbers n_{ix}, n_{iy}, and n_{iz}, which are such that the energy of the state can be expressed in terms of the principal quantum number $n_i = n_{ix} + n_{iy} + n_{iz}$ as follows:

$$\varepsilon_{n_i} = \omega\left(n_i + \frac{3}{2} \right) + \frac{k^2}{2mN^2}. \tag{2}$$

Before the momentum $\mathbf{q}_i$ is imparted to an oscillator, it is assumed to be in the ground state $|0,0,0\rangle$. If the momentum $\mathbf{q}_i$ is instantaneously imparted to the oscillator, the probability of observing any of the degenerate states (1) having the energy (2) has the form

$$w_{n_i}=\sum_{n_{ix}+n_{iy}+n_{iz}=n_i}|\langle n_{ix},n_{iy},n_{iz}| \times\exp\left\{i\left(\mathbf{q}_i-\frac{\mathbf{k}}{N}\right)\mathbf{r}|0,0,0\rangle\right|^2, \tag{3}$$

where the summation is performed over all values of n_{ix}, n_{iy}, and n_{iz} with the assigned sum $n_{ix}+n_{iy}+n_{iz}=n_i$, and $\mathbf{r}$ denotes the oscillator coordinates.

Fermi's results[7] are used to calculate the square of the matrix element (3) and for summation (see also Ref. 8). Then, the probability of observing the ith oscillator (ith atom) in any state with the principal quantum number n_i and the momentum $\mathbf{k}/N$ equals

$$w_{n_i}=\frac{1}{n_i!}\left[\frac{1}{2\alpha^2}\left(\mathbf{q}_i-\frac{\mathbf{k}}{N}\right)^2\right]^{n_i}\exp\left\{-\frac{1}{2\alpha^2}\left(\mathbf{q}_i-\frac{\mathbf{k}}{N}\right)^2\right\}, \tag{4}$$

where $\alpha^2=m\omega$ and m is the oscillator (atomic) mass.

If we have a system of oscillators, each of which is excited as a result of the instantaneous transfer of the momentum $\mathbf{q}_i(i=1,2,\ldots,N)$ with the probability (4) to any state having the principal quantum number $n_i(i=1,2,\ldots,N)$ and the same momentum $\mathbf{k}/N$, the probability of observing a system of N oscillators in any state having the energy

$$E_n=\omega\left(n+N\frac{3}{2}\right)+\frac{k^2}{2mN}, \tag{5}$$

where $n=\Sigma_{i=1}^{N}n_i$, moving as an entity with the momentum $\mathbf{k}=N(\mathbf{k}/N)$ has the form

$$W_N^{(n)}=\sum_{n_1+n_2+\ldots+n_N=n}\left[\frac{1}{2\alpha^2}\left(\mathbf{q}_1-\frac{\mathbf{k}}{N}\right)^2\right]^{n_1} \times\left[\frac{1}{2\alpha^2}\left(\mathbf{q}_2-\frac{\mathbf{k}}{N}\right)^2\right]^{n_2} \times\ldots\times\left[\frac{1}{2\alpha^2}\left(\mathbf{q}_N-\frac{\mathbf{k}}{N}\right)^2\right]^{n_N}\frac{1}{n_1!n_2!\ldots n_N!} \times\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{N}\left(\mathbf{q}_i-\frac{\mathbf{k}}{N}\right)^2\right\}, \tag{6}$$

where the summation is performed over all possible values of $n_1, n_2, \ldots, n_N$ under the condition $n_1+n_2+\ldots+n_N=n$, where n has the meaning of the principal quantum number of a system of N oscillators (in our case, a cluster of N atoms). As a result, we have

$$W_N^{(n)}=\frac{1}{n!}\left[\frac{1}{2\alpha^2}\sum_{i=1}^{N}\left(\mathbf{q}_i-\frac{\mathbf{k}}{N}\right)^2\right]^n \times\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{N}\left(\mathbf{q}_i-\frac{\mathbf{k}}{N}\right)^2\right\}. \tag{7}$$

This is the probability of a cluster of N atoms escaping as an entity with the momentum $\mathbf{k}$ in the state of excitation n. Equation (7) is then summed over all states of vibrational excitation of the cluster up to a certain number n_0, at which the energy stored in the excited oscillators is sufficient to break up the cluster. A sufficient condition for this is $n_0\approx\delta/\omega$, under which the oscillation energy of all the oscillators is sufficient to eject a single atom from a well of depth δ. Thus, we need to calculate

$$W_N(n_0)=\sum_{n=0}^{n_0}W_N^{(n)},$$

where $n_0\gg1$; therefore, we take the sum in the form

$$f(x)=e^{-x}\sum_{n=0}^{n_0}\frac{1}{n!}x^n\approx e^{-x}\sum_{n=0}^{\infty}\frac{1}{n!}x^n\exp\left\{-\frac{n}{n_0}\right\} =\exp\left\{x\left[\exp\left(-\frac{1}{n_0}\right)-1\right]\right\}\approx\exp\left\{-\frac{x}{n_0}\right\}. \tag{8}$$

Thus,

$$W_N(n_0)=\exp\left\{-\frac{1}{n_0}\frac{1}{2\alpha^2}\sum_{i=1}^{N}\left(\mathbf{q}_i-\frac{\mathbf{k}}{N}\right)^2\right\}. \tag{9}$$

The probability (9) should then be averaged over all possible values of $\mathbf{q}_i(i=1,2,\ldots,N)$. For the distribution of $\mathbf{q}_i$ we naturally assume that all $\mathbf{q}_i$ are independent, but that the directions of $\mathbf{q}_i$ are equally probable, and we take the average over the angles of the vectors $\mathbf{q}_i$

$$\overline{W}_N(n_0)=\int\frac{d\Omega_{\mathbf{q}1}}{4\pi}\int\frac{d\Omega_{\mathbf{q}2}}{4\pi}\ldots\int\frac{d\Omega_{\mathbf{q}_N}}{4\pi}W_N(n_0) =\prod_{i=1}^{N}\left[\frac{\exp\left\{-\frac{\beta}{n_0}\left(q_i-\frac{k}{N}\right)^2\right\}-\exp\left\{-\frac{\beta}{n_0}\left(q_i+\frac{k}{N}\right)^2\right\}}{4q_i\frac{k}{N}\frac{\beta}{n_0}}\right], \tag{10}$$

where $\beta=1/(2\alpha^2)$.

The following calculations are simplified appreciably, if we assume that all $\mathbf{q}_i$ have the same length, i.e., $|\mathbf{q}_i|=q$. This implies that, on the average, all $\mathbf{q}_i$ have the same magnitude but are randomly directed. Then we have

$$\overline{W}_N(n_0)=\left[\frac{\exp\left\{-\frac{\beta}{n_0}\left(q-\frac{k}{N}\right)^2\right\}-\exp\left\{-\frac{\beta}{n_0}\left(q+\frac{k}{N}\right)^2\right\}}{4q\frac{k}{N}\frac{\beta}{n_0}}\right]^N. \tag{11}$$

Formula (11) describes the probability of a cluster of N atoms escaping if the kinetic energy of the cluster is suffi-

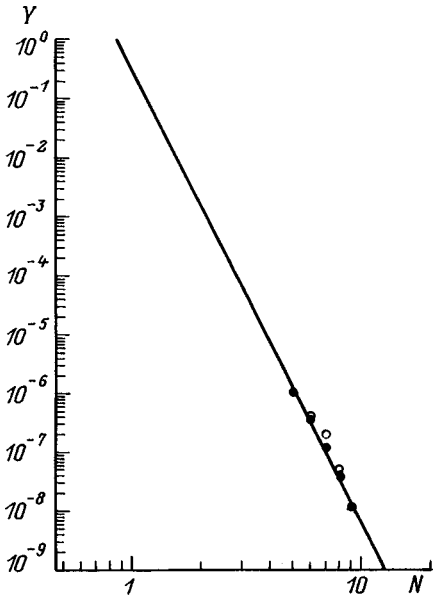

FIG. 1. Relative cluster yield Y as a function of the number of atoms N in cluster for a tantalum target bombarded by singly charged 5 keV argon ions: ○ — theory, Eq. (19), ● — experiment[3], curve — power law.[3]

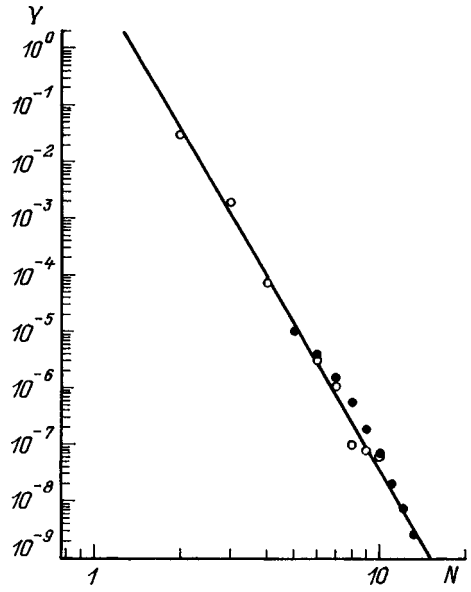

FIG. 2. Relative cluster yield Y as a function of the number of atoms N in cluster for an aluminum target bombarded by singly charged 5 keV argon ions: ○ — theory, Eq. (19), ● — experiment[3], curve — power law.[3]

cient to overcome the binding energy of N atoms with the remaining metal atoms. This binding energy is proportional to the surface area S_N over which the block of N atoms is in contact with the remaining metal. We assume that it is a hemisphere with its center lying on the surface of the metal before sputtering. The radius of the hemisphere is clearly related to the number of atoms in the cluster:

$$R_N=\left[N\frac{3}{2\pi d}\right]^{1/3}, \tag{12}$$

where d is the number of atoms per unit volume.

Then, the binding energy of the cluster, which is proportional to S_N, is given by

$$U_N=\sigma S_N=\sigma 2\pi R_N^2=\sigma\left(\frac{3}{2\pi d}\right)^{2/3}N^{2/3}=\delta N^{2/3}, \tag{13}$$

where δ has the meaning of the binding energy per atom.

After overcoming the binding energy, a cluster which has acquired the momentum $\mathbf{k}$ before detachment (before overcoming the binding energy) will move with the kinetic energy T_N given by

$$T_N=\frac{k^2}{2mN}-\sigma S_N. \tag{14}$$

Setting $T_N=0$, from Eq. (14) we find the minimum value of k_{0N} above which the cluster can overcome the binding energy:

$$k_{0N}=[2mN\sigma S_N]^{1/2}=[2m\delta]^{1/2}N^{5/6}=k_{01}N^{5/6}, \tag{15}$$

where $k_{01}=(2m\delta)^{\frac{1}{2}}$ has the meaning of the minimum momentum needed for the detachment of a single atom.

Note that having rewritten expression (15) in the form

$$k_{0N}=k_{01}N^{\nu}, \tag{16}$$

where the exponent ν depends on the cluster shape, we can study clusters of different shape (other than spherical, for which $\nu=5/6$) by varying ν.

Next, in order to obtain the probability of a cluster of N atoms being detached as an entity, we need to integrate the probability (11) over all possible values of $\mathbf{k}$, i.e., with the condition $|\mathbf{k}|>k_{0N}$ and $\mathbf{k}$ directed outward, which corresponds to a solid angle equal to 2π

$$\overline{W}_N(n_0)=\int_{|\mathbf{k}|>k_{0N}}\overline{W}_N(n_0)\,d^3k=\int_{k>k_{0N}}^{\infty}2\pi k^2\,dk\,\overline{W}_N(n_0). \tag{17}$$

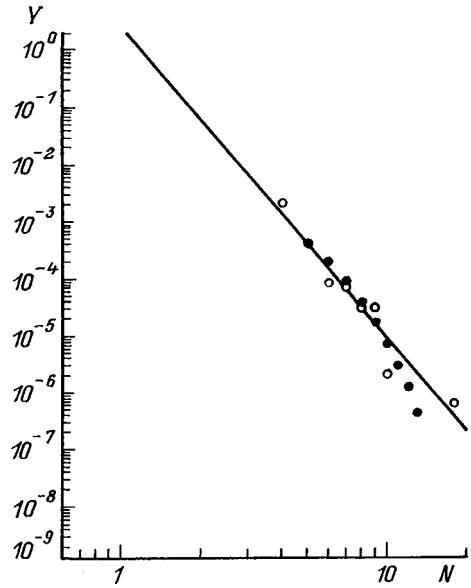

FIG. 3. Relative cluster yield Y as a function of the number of atoms N in cluster for a silver target bombarded by singly charged 5 keV argon ions: ○ — theory, Eq. (19), ● — experiment[3], curve — power law.[3]

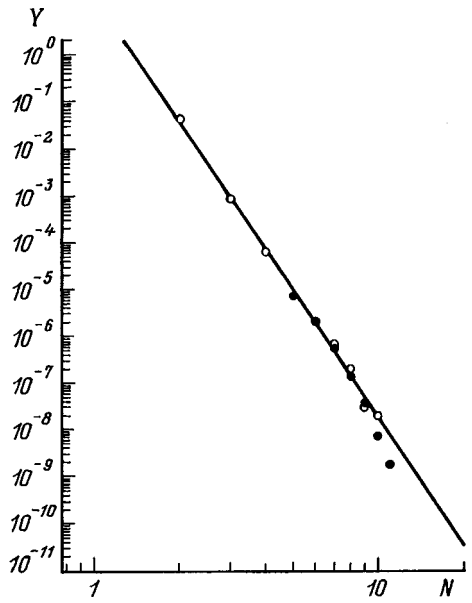

FIG. 4. Relative cluster yield Y as a function of the number of atoms N in cluster for a niobium target bombarded by singly charged 5 keV argon ions: ○ — theory, Eq. (19), ● — experiment[3], curve — power law.[3]

By integrating (17), we can reduce the probability to a form in which the expression in the integrand has the meaning of the spectral distribution of the clusters over their kinetic energy T_N, for which we rewrite Eq. (14) in the form

$$T_N=\frac{1}{2mN}(k^2-k_{0N}^2). \tag{18}$$

As a result of a replacement of variables in Eq. (17), we obtain the final expression for the probability of the detachment of a cluster of N atoms

$$\overline{W}_N(n_0)=2\pi\int_0^\infty kmN\,dT_N \times\left[\frac{\exp\left\{-\frac{\beta}{n_0}\left(q-\frac{k}{N}\right)^2\right\}-\exp\left\{-\frac{\beta}{n_0}\left(q+\frac{k}{N}\right)^2\right\}}{4q\frac{k}{N}\frac{\beta}{n_0}}\right]^N, \tag{19}$$

where $k=\sqrt{2mNT_N+(k_{0N})^2}$, $\beta/n_0=(2m\delta)^{-1}=(k_{01})^{-2}$, and a lower limit is imposed on the values of q by the condition for cluster detachment $Nq^2/(2m)\geqslant(k_{0N})^2/(2mN)$ or, in accordance with Eq. (15), $q\geqslant k_{01}/N^{\frac{1}{6}}$.

Thus, we have calculated the probability of a cluster becoming detached as an entity (block) without the atoms changing places with one another. If we also take this possibility into account, it leads to the appearance of a small parameter corresponding to overlap of the oscillator functions centered at different points. Thus, we can neglect these permutations.

3. COMPARISON WITH EXPERIMENTS

It should be noted that the procedure for integrating over d^3k in Eq. (17) or over dT_N in Eq. (19) from the dimensionless probability (11) less than unity gives an unnormalized dimensional probability. However, this is unimportant because the relative probabilities of the escape of clusters with various numbers of atoms are usually measured experimentally.

TABLE I. Values[9] of the binding energy δ (eV) and the variable parameter q (a.u.) for various ion–target combinations (5 keV Ar^+ bombarding ions).

Target	Binding energy δ, eV	q, a.u.
Ag	2.96	200
Al	3.34	120
Nb	7.47	400
Ta	8.65	550

In addition, it should also be stressed that we have considered the ejection of large clusters with $N\gg1$, which, according to our model, are formed in the early stage of evolution of a collision cascade. Whereas the entire evolution period of a cascade must be taken into account to calculate the total sputtering (including small clusters), to make a comparison with experiment, the probability (19) should first be normalized to the probability of the detachment of a cluster with $N=5$ (more precisely, any $N\geqslant5$ can be selected but $N=5$ is more convenient for our purposes)

$$Y_N=\frac{\overline{W}_N(n_0)}{\overline{W}_5(n_0)}. \tag{20}$$

The experimental data should also be normalized similarly. Then, if necessary we can convert to any convenient arbitrary units. We obtained the calculated data plotted in Figs. 1–4 in just this way together with the experimental results presented in Ref. 3. For the calculations we assumed that q is a variable parameter. Table I gives values of the binding energy δ (eV) and the variable parameter q (a.u.) for various ion–target combinations.

Thus, as in Refs. 1 and 2, we conclude that the mechanisms responsible for sputtering in the form of small ($N<5$) and large ($N>5$) clusters differ substantially. Whereas small clusters are formed either directly as a result of being knocked out or as a result of the merging of isolated atoms (or smaller clusters) over the entire evolution time of the collision cascade (usually $\geqslant10^{-12}$ s), large clusters are ejected as an entity in the early stages of evolution of the collision cascade ($\leqslant10^{-13}$ s).

[1)] Nevertheless, we assume that all the levels are equidistant. Clearly, this equivalence is only impaired near the truncation boundary, and if the number of bound levels in this potential is fairly high (and we shall assume that the number of levels is $\sim10^2$–10^3), the deviation from equidistance is insignificant.

[2)] In order to avoid any misunderstanding, we stress that the function (1) is not the state function of the continuous spectrum (truncated oscillator potential) of an isolated oscillator but merely reflects the fact that any given oscillator, being in the bound state $|n_{ix},n_{iy},n_{iz}\rangle$, moves as an entity with the average momentum $\mathbf{k}/N$ together with the remaining $(N-1)$ oscillators. Here and everywhere below, we use the atomic units $h=m_e=e=1$.

[1] *Fundamental and Applied Aspects of Sputtering of Solids* [Russ. trans., Mir, Moscow (1989) 399 pp.].

[2] *Interaction of Charged Particles with Solids and Surfaces*, NATO ASI Series. Series B, Physics, Vol. 271, edited by A. Gras-Martí, H. M. Urbassek, N. R. Arista, and F. Flores (Plenum Press, New York, 1991) [Russ. trans., Vysshaya Shkola, Moscow, 1994, 752 pp.].
[3] A. Wucher and W. Wahl, Nucl. Instrum. Methods Phys. Res. B **115**, 581 (1996).
[4] S. R. Coon, W. F. Calaway, and M. Y. Pellin, Nucl. Instrum. Methods Phys. Res. B **90**, 518 (1994).
[5] P. Sigmund, Phys. Rev. **184**, 383 (1969).
[6] A. Wucher and B. Y. Garrison, J. Chem. Phys. **105**, 5999 (1996).
[7] E. Fermi, *Collected Papers (Note e Memorie), Vol. 1: Italy 1921–1938* (University of Chicago Press, Chicago, 1962) [Russ. trans., Nauka, Moscow, 1971, 818 pp.].
[8] L. D. Landau and E. M. Lifshits, *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Pergamon Press, Oxford, 1977) [Russ. original, 4th ed., Nauka, Moscow, 1989, 768 pp.].
[9] C. Kittel, *Introduction to Solid-State Physics*, 5th ed. (Wiley, New York, 1976) [Russ. trans., Nauka, Moscow, 1978, 792 pp.].

Translated by R. M. Durham

BRIEF COMMUNICATIONS

Vavilov–Cherenkov effect in a chiral waveguide

K. A. Barsukov and A. A. Smirnova

St. Petersburg Electrical-Engineering University, 197376 St. Petersburg, Russia
(Submitted March 25, 1997)
Zh. Tekh. Fiz. **69**, 69–71 (March 1999)

The Vavilov–Cherenkov radiation appearing in a circular regular ideal waveguide with a chiral medium is considered. The features of the emission spectrum excited by a charge moving along the waveguide axis are investigated. Expressions for the energy loss of a moving charge are obtained. © *1999 American Institute of Physics.* [S1063-7842(99)01203-9]

The electrodynamics of chiral media have been actively discussed in recent years (see, for example, Ref. 1, and the literature cited therein). Here, special interest has been noted in the application of chiral media in the microwave range. For example, the features of the propagation of electromagnetic waves in waveguides with various configurations were investigated in Refs. 2–4. One of the effective excitation methods involves utilization of the Vavilov–Cherenkov effect.[5] The features of this method for exciting waveguides are explored below in the example of a circular waveguide with a chiral medium.

Let us consider a regular waveguide with a circular transverse section of radius a, whose axis coincides with the $0z$ axis whose transverse section lies in the $x0y$ plane. The waveguide walls have infinite conductivity, and its volume is filled by a chiral medium, whose material equations can be written in the form[1]

$$\mathbf{D}=\varepsilon\mathbf{E}-i\chi\mathbf{H}, \qquad \mathbf{B}=\mu\mathbf{H}+i\chi\mathbf{E}, \tag{1}$$

where ε and μ are the dielectric constant and the magnetic permeability of the medium, and χ is the chirality parameter.

A point charge q moves with a constant velocity $\mathbf{v}$ along the waveguide axis, creating a charge density and a current density of the form $\rho=q\,\delta(z-vt)\,\delta(x)\,\delta(y)$ and $\mathbf{j}=\rho\mathbf{v}$ or, in the time-Fourier representation,

$$\rho_\omega=\frac{q}{2\pi v}\exp\left(-i\frac{\omega}{v}z\right)\delta(x)\,\delta(y), \quad \mathbf{j}_\omega=\rho_\omega\mathbf{v}. \tag{2}$$

A convenient technique for representing fields is to introduce the potentials[3] associated with field sources using the equation

$$\Delta\Psi_{1,2_\omega}+k^2n_{1,2}^2\Psi_{1,2_\omega}=4\pi\sqrt{\varepsilon/\mu}\,i \times\{-1/n_{1,2}\omega/v\,\rho_\omega+k/cn_{1,2}j_{\omega z}\}, \tag{3}$$

where $k=\omega/c$, $n_{1,2}=n\pm\chi$, $n=\sqrt{\varepsilon\mu}$, the subscript 1 refers to the upper sign, and the subscript 2 refers to the lower sign. The dependence of all the quantities on z in the form of the factor $\exp(-i(\omega/v)z)$ is omitted here and below.

Then, due to the symmetry of the problem $\Psi_{1,2}=\Psi_{1,2}(r)$, and the field vectors can be expressed in terms of the potentials by the following relations:[6]

$$E_{\omega r}=-\frac{i}{2}\frac{\omega}{v}\frac{\partial}{\partial r}(\Psi_{1\omega}/\varkappa_1^2+\Psi_{2\omega}/\varkappa_2^2),$$

$$H_{\omega r}=\frac{1}{2}\sqrt{\varepsilon/\mu}\frac{\omega}{v}\frac{\partial}{\partial r}(\Psi_{1\omega}/\varkappa_1^2-\Psi_{2\omega}/\varkappa_2^2),$$

$$E_{\omega\varphi}=\frac{k}{2}\frac{\partial}{\partial r}(-n_1\Psi_{1\omega}/\varkappa_1^2+n_2\Psi_{2\omega}/\varkappa_2^2),$$

$$H_{\omega\varphi}=-\frac{ik}{2}\sqrt{\varepsilon/\mu}\frac{\partial}{\partial r}(n_1\Psi_{1\omega}/\varkappa_1^2+n_2\Psi_{2\omega}/\varkappa_2^2),$$

$$E_{\omega z}=(\Psi_{1\omega}+\Psi_{2\omega})/2,$$

$$H_{\omega z}=i\sqrt{\varepsilon/\mu}(\Psi_{1\omega}+\Psi_{2\omega})/2, \tag{4}$$

where $\varkappa_{1,2}=(\omega/v)\sigma_{1,2}$, $\sigma_{1,2}=\sqrt{\beta^2n_{1,2}^2-1}$, and $\beta=v/c$.

We shall seek the solution of the problem in the form of the sum of the partial solution of the inhomogeneous system of equations (3) for an infinite chiral medium and the general solution of the homogeneous system (3), assuming that $\beta n_1>1$, $\beta n_2>1$, and

$$\Psi_{1,2_\omega}=iq/\pi\sqrt{\mu/\varepsilon}\,\varkappa_{1,2}/\omega n_{1,2}^2K_0(\varkappa_{1,2}r)+\alpha_{1,2}J_0(\varkappa_{1,2}r). \tag{5}$$

The first term in (5) describes the field of a charge in an infinite chiral medium and is found as in the case of ordinary achiral media (see, for example, Ref. 6), and the coefficients $\alpha_{1,2}$ are determined from the boundary conditions on the waveguide walls

$$E_{\omega\varphi}=0, \quad E_{\omega z}=0 \quad \text{at} \quad r=a. \tag{6}$$

The corresponding substitution of (5) into (6) permits satisfaction of the conditions (6) and determination of the unknown coefficients $\alpha_{1,2}$ in the form

$$\alpha_{1,2}=\Delta_{1,2}/\Delta, \tag{7}$$

where

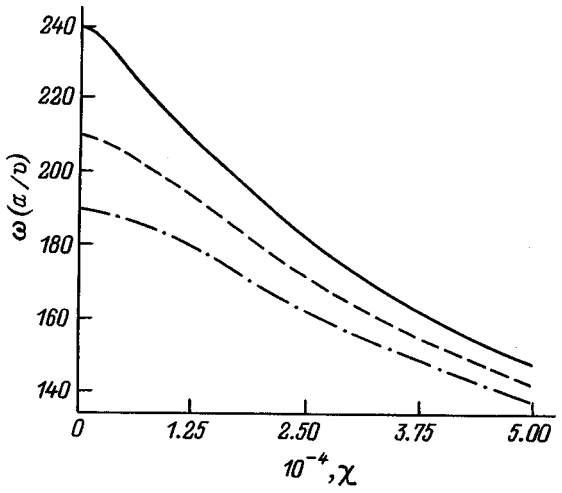

FIG. 1.

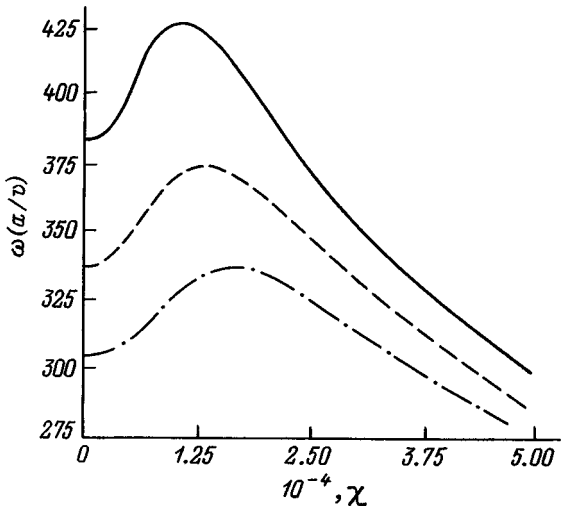

FIG. 2.

$$\Delta = J_0(\varkappa_1 a)J_1(\varkappa_2 a)n_2/\varkappa_2 + J_0(\varkappa_2 a)J_1(\varkappa_1 a)n_1/\varkappa_1, \tag{8}$$

$$\begin{aligned}\Delta_1 = &-iq/\pi\sqrt{\mu/\varepsilon}\,1/\omega[\{\varkappa_1^2/n_1 K_0(\varkappa_1 a)\\ &+\varkappa_2^2/n_2 K_0(\varkappa_2 a)\}J_1(\varkappa_2 a)n_2/\varkappa_2\\ &+\{\varkappa_1 K_1(\varkappa_1 a)-\varkappa_2 K_1(\varkappa_2 a)\}J_0(\varkappa_2 a)],\\ \Delta_2 = &-iq/\pi\sqrt{\mu/\varepsilon}\,1/\omega[\{\varkappa_1^2/n_1 K_0(\varkappa_1 a)\\ &+\varkappa_2^2/n_2 K_0(\varkappa_2 a)\}J_1(\varkappa_2 a)n_1/\varkappa_1\\ &+\{\varkappa_1 K_1(\varkappa_1 a)+\varkappa_2 K_1(\varkappa_2 a)\}J_0(\varkappa_1 a)].\end{aligned} \tag{9}$$

The Vavilov–Cherenkov radiation in a chiral waveguide, as in the ordinary case, has a discrete spectrum, which can be determined from the condition $\Delta = 0$ or from (8):

$$\varkappa_1 n_2 J_0(\varkappa_1 a)J_1(\varkappa_2 a) + \varkappa_2 n_1 J_0(\varkappa_2 a)J_1(\varkappa_1 a) = 0. \tag{10}$$

The complete expression for the potentials $\Psi_{1,2}$ is found using inverse Fourier transformation:

$$\Psi_{1,2}(r,z) = \int_c \Psi_{1,2_\omega}(r)\exp(-i\omega/v(z-vt))\,d\omega, \tag{11}$$

where the integration contour c is chosen along the real ω axis with circumvention of the simple poles specified by the zeros of the relation (10) and with the circumvention direction dictated by the emission condition. It is also assumed that the insulator does not have dispersion. This assumption usually holds well in the radio-frequency region.

Performing this integration by substituting (5) into (11) and calculating the residues, we obtain

$$\begin{aligned}\Psi_1(r,z) &= \sum_m F_m\left(\frac{J_0(\varkappa_{2m}a)}{J_0(\varkappa_1 a)} - 1\right),\\ \Psi_2(r,z) &= \sum_m F_m\left(\frac{J_0(\varkappa_{1m}a)}{J_0(\varkappa_{2m}a)} - 1\right),\end{aligned} \tag{12}$$

where

$$F_m = \frac{q}{a^2\varepsilon\{J_1(\varkappa_{1m}a)J_1(\varkappa_{2m}a)(\beta^2 n_1 n_2 - 1)/\sigma_1\sigma_2 - J_0(\varkappa_{1m}a)J_0(\varkappa_{2m}a)}$$

and the summation is carried out over all m, the numbers of the roots of Eq. (10).

The expression (12) together with (4) completely specifies the field of a moving charge in a chiral waveguide. For the case of $\beta n_1 > 1$, $\beta n_2 < 1$, we must replace $\varkappa_2$ by $\varkappa_2' = i\varkappa_2$ and σ_2 by $\sigma_2' = i\sigma_2$ in Eqs. (5)–(12).

As can be seen from (10), in the general case a charge moving in a waveguide excites hybrid waves, which can separate into quasi-TE waves with a longitudinal component of the electric field proportional to χ at small intensities and into quasi-TM waves with a similar feature in the longitudinal component of the magnetic field.

At small values of χ (10) can be expanded into a series in χ for a quasi-TM field far from resonance with consideration of the proximity of the Cherenkov frequencies to the achiral case, and with accuracy to χ^4 we have

$$\omega_m = \omega_{0m}\left(1 - \frac{\beta^2(\beta^2 n^2 + 1)}{2\sigma^4}\chi^2\right), \tag{13}$$

where

$$\omega_{0m} = \frac{\nabla\nu_{0m}}{a\sigma}, \qquad \sigma = \sqrt{\beta^2 n^2 - 1}, \tag{14}$$

and ν_{0m} is the mth root of the equation $j_0(\nu_{0m}) = 0$.

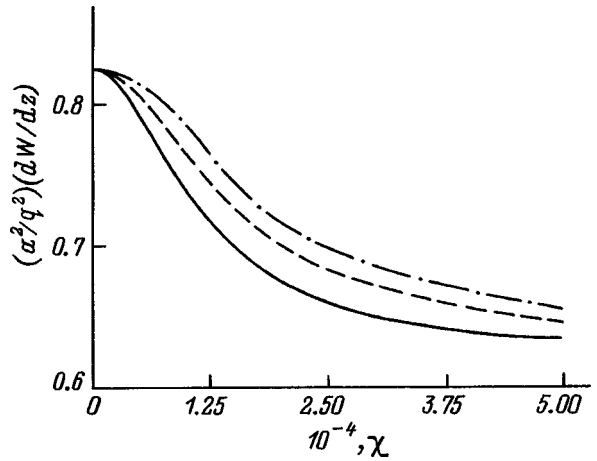

FIG. 3.

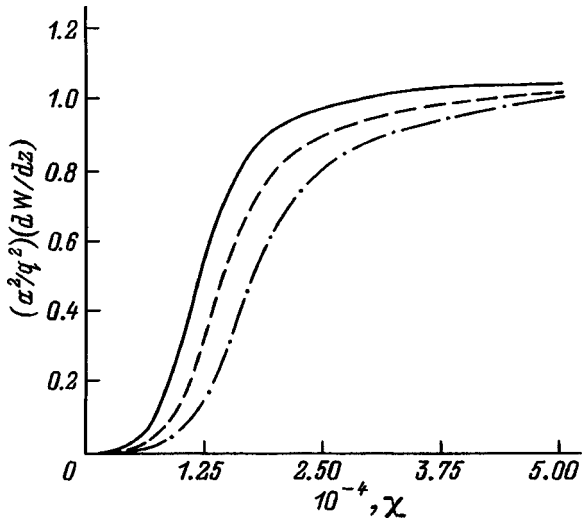

FIG. 4.

Similarly, for the second group of modes (quasi-TM modes) the relation (13) has the form

$$\omega_m' = \omega_{0m}' \left(1 + \frac{\beta^2(\beta^2 n^2 + 3)}{2\sigma^4} \chi^2 \right), \tag{15}$$

where $\omega_{0m}' = \nabla \mu_{0m} / a\sigma$ and μ_{0m} is the mth root of the equation $J_1(\mu_{0m}) = 0$.

The computer solution of Eq. (10) provides the dependence of the frequencies of the quasi-TM and quasi-TE modes on χ. For small values of χ the chiral effect will be appreciable at the emission threshold when $\beta n = 1$. As an illustration, Figs. 1 and 2 present plots of the frequency of the Vavilov–Cherenkov radiation as a function of the chirality parameter χ for the lowest quasi-TM and TE modes, respectively, $n = 3$, and $\sigma = 1 \times 10^{-2}$ (solid curve), 1.14×10^{-2} (dashed curve), and 1.26×10^{-2} (dot-dashed curve).

The energy losses of a charge per unit length of its trajectory can be found most simply, as usual, from the decelerating force exerted by the radiation field acting on the charge: $dW/dz = qE_z|_{r=0}^{z=vt}$. Two cases should also be distinguished here. In the first, which is defined by the conditions $\beta n_1 > 1$ and $\beta n_2 > 1$ and thus corresponds to two Cherenkov cones in an infinite chiral medium, the expression for the losses has the form

$$\frac{dW}{dz} = -\frac{q}{2} \sum_m F_m \frac{(J_0(\varkappa_{1m} a) - J_0(\varkappa_{2m} a))^2}{J_0(\varkappa_{1m} a) J_0(\varkappa_{2m} a)}. \tag{16}$$

In the second case, where $\beta n_1 > 1$ and $\beta n_2 < 1$ and only one Cherenkov cone is generated in an infinite medium, we have

$$\frac{dW}{dz} = -\frac{q}{2} \sum_m F_m' \frac{(J_0(\varkappa_{1m} a) - J_0(\varkappa_{2m}' a))^2}{J_0(\varkappa_{1m} a) J_0(\varkappa_{2m}' a)},$$

where

$$F_m' = \frac{q}{a^2 \varepsilon \{ J_1(\varkappa_{1m} a) I_1(\varkappa_{2m}' a)(\beta^2 n_1 n_2 - 1)/\sigma_1 \sigma_2' - J_0(\varkappa_{1m} a) I_0(\varkappa_{2m}' a) }, \tag{17}$$

and $\varkappa_{1m} a$ and $\varkappa_{2m} a$ are roots of the equation

$$J_0(\varkappa_1 a) I_1(\varkappa_2' a) n_2 / \varkappa_2' + I_0(\varkappa_2' a) J_1(\varkappa_1 a) n_1 / \varkappa_1 = 0.$$

Figures 3 and 4 present the dependence of $a^2 dW/q^2 dz$ on the chirality parameter of the lowest quasi-TM and quasi-TE modes on χ for $n = 3$ and $\sigma = 1 \times 10^{-2}$ (solid curve), 1.14×10^{-2} (dashed curve), and 1.26×10^{-2} (dot-dashed curve). It is seen that upon passage to the achiral case ($\chi \to 0$) the TE field vanishes, and only the TM mode remains.

In the limiting case of $\chi \to 0$ formulas (16) and (17) transform into the familiar expression for an achiral medium[7]

$$\frac{qW}{dz} = \sum_m \frac{2q^2}{a^2 \varepsilon J_1^2(\varkappa_m a)},$$

where m is the number of the root of the equation $J_0(\varkappa_m a) = 0$.

[1] Nader Engheta, Dwight L. Jaggard, and Marek W. Kowarz, IEEE Trans. Antennas Propag. **40**, 367 (1992).
[2] Jan A. M. Skein, IEEE Trans. Microwave Theory Tech. **MTT-38**, 1488 (1990).
[3] Cornel Eftimiu and L. Wilson Pearson, Radio Sci. **24**, 351 (1989).
[4] Richard Hollinger, Vasundara V. Varadan, and Vijay K. Varadan, Radio Sci. **26**, 351 (1991).
[5] V. I. Kanavets and A. N. Sandapov, Itogi Nauki Tekh. No. 17, 82 (1985).
[6] T. P. Khan, J. Phys. A **3**, 246 (1970).
[7] C. Muzikar, Czech. J. Phys. **5**, 9 (1955).

Translated by P. Shelnitz

Influence of nonparallelism of the surfaces of a piezoelectric layer on the frequency band of an electroacoustic piezoelectric transducer

V. V. Petrov and S. A. Lapin

N. G. Chernyshevskiĭ Saratov State University, 410071 Saratov, Russia
(Submitted May 19, 1997)
Zh. Tekh. Fiz. **69**, 72–73 (March 1999)

The influence of nonparallelism of the surfaces of a piezoelectric transducer on its impedance and frequency band is analyzed theoretically. The model of a wedge-shaped piezoelectric layer taken for the analysis consists of a set of n plane-parallel elements arranged in succession along the length of the transducer with a gradual increase in the thickness of the piezoelectric layer. The analysis performed demonstrates the possibility of significantly expanding the band of working frequencies of a piezoelectric transducer when its electroacoustic conversion efficiency is reduced. In some cases, however, a decrease in conversion efficiency can be employed to optimize an acoustic device, for example, in creating a filter-type piezoelectric transducer, where electromagnetic energy is converted into acoustic energy successively from one cell of the filter to another. The proposed method for expanding the frequency band can be useful, for example, in creating high-frequency acousto-optic Bragg cells. © *1999 American Institute of Physics.* [S1063-7842(99)01303-3]

A method for expanding the frequency band of acousto-optic (AO) Bragg cells by employing a multielement transducer with several variable parameters, viz., lattice spacing, electrode length, and thickness of the piezoelectric layer, was discussed in Ref. 1. The derived frequency dependence of the spacing in the multielement structure required for ''exact'' autotuning of an acoustic beam to the Bragg angle in an assigned frequency range was described in that paper. Under such an approach the constraint on the AO interaction length is removed, suggesting a possibility for significantly increasing the diffraction efficiency. It was theorized that the frequency band of a piezoelectric transducer could be matched to the AO interaction band by varying the thickness of the piezoelectric layer in addition to the spacing of the structure along the length of the transducer. In this case the region of maximum conversion efficiency of the piezoelectric corresponds to the portion of the lattice which provides for excitation of an acoustic beam in the direction corresponding to the Bragg condition. The present paper examines the influence of nonparallelism of the planar faces of a transducer on its working frequency band.

Figure 1 shows a model of the wedge-shaped transducer analyzed, which consists of piezoelectric layer *1*, which is confined between metallic electrodes of finite thickness, viz., superlayer *2* of thickness p and underlayer *3* of thickness g, which is in acoustic contact with isotropic sound conductor *4*. It is assumed that the length L of the transducer and its width b (in the direction perpendicular to the wedge-shaped cross section; not shown in the figure) significantly exceed its thickness h, so that the model can be considered one-dimensional, i.e., only the component of the electric field **E** directed perpendicularly to the surface of the crystal may be taken into account. Such an assumption also permits treatment of the wedge-shaped transducer as a set of n plane-parallel elements, which are arranged in succession along the length L with a gradual increase in the thickness of the piezoelectric layer from the minimum value h_1 to the maximum value h_2 and are connected electrically in parallel. The impedance Z of such a transducer can be found as the reciprocal of the sum of the complex conductances Y_i of all n elements

$$Z=\frac{1}{\Sigma_{i=1}^{n}Y_i}=\frac{1}{\sum\limits_{i=1}^{n}\frac{1}{Z_i}}=\frac{1}{\sum\limits_{i=1}^{n}\frac{1}{R_i+jX_i}}=R+jX, \tag{1}$$

where R and X are the active and reactive components of the impedance of the transducer and Z_i is the complex impedance of the ith element.

The question of finding the radiation impedance of a plane-parallel transducer was considered in Ref. 2. The frequency properties of the transducer are determined mainly by the active component of the conversion impedance. Following the results of Ref. 2, we write the expression for the

-

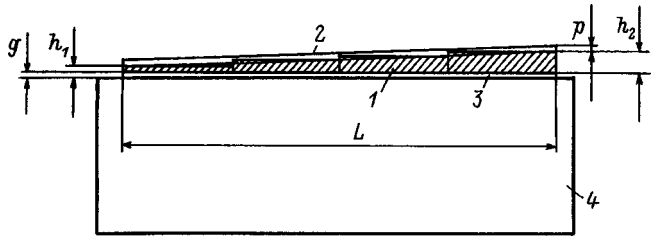

FIG. 1. Geometry of a wedge-shaped transducer.

dimensionless active radiation resistance R/X for the ith element

$$\frac{R_i}{X_i}=\frac{4k^2}{\beta h_i}F_{ai}. \tag{2}$$

Here $\beta=(2\pi f/\nu_2)$ is the propagation constant in the piezoelectric, f is the frequency, ν_2 is the velocity of sound in the piezoelectric, k is the electromechanical coupling coefficient, and F_{ai} is defined by the relation

$$F_{ai}=\frac{\sin^2\left(\frac{\beta h_i}{2}\right)\left[\sin\left(\frac{\beta h_i}{2}\right)\cos(a_i\beta h_i)+\frac{Z_{01}}{Z_{02}}\sin(a\beta h_i)\cos\left(\frac{\beta h_i}{2}\right)\right]^2}{(Z_{02}/Z_{04})M_i^2+(Z_{04}/Z_{02})N_i^2}, \tag{3}$$

where

$$M_i=(Z_{03}/Z_{02})A_i\sin(b_i\beta h_i)+B_i\cos(b_i\beta h_i),$$

$$N_i=A_i\cos(b_i\beta h_i)-(Z_{02}/Z_{03})B_i\sin(b_i\beta h_i),$$

$$A_i=\cos(\beta h_i)\cos(a_i\beta h_i)-(Z_{01}/Z_{02})\times\sin(\beta h_i)\sin(a_i\beta h_i),$$

$$a_i=(\nu_2/\nu_1)(p/h_i),$$

$$B_i=\sin(\beta h_i)\cos(a_i\beta h_i)+(Z_{01}/Z_{02})\times\cos(\beta h_i)\sin(a_i\beta h_i),$$

$$b_i=(\nu_2/\nu_3)(g/h_i),$$

h_i is the thickness of the ith element; ν_1 and ν_3 are the velocities of sound in the superlayer and in the underlayer; and Z_{01}, Z_{02}, Z_{03}, and Z_{04} are the acoustic wave impedances of the superlayer, the piezoelectric, the underlayer, and the sound conductor.

The reactive component of the impedance of the transducer is determined mainly by its capacitive reactance: $X_i=1/(\omega C_i)$, where ω is the angular frequency, C_i is the electric capacitance of the ith element. Here $C_i=\varepsilon_0\varepsilon_r S_i/h_i$, where ε_0 and ε_r are the absolute permittivity of free space and the relative dielectric constant of the piezoelectric, and $S_i=bL/n$ is the area of the ith element of the transducer. Thus, the expression for the active radiation resistance of the ith element of the transducer can be written in the form

$$R_i=\frac{4k^2\nu_2 n h_i}{\omega^2\varepsilon_0\varepsilon_r bL}F_{ai}. \tag{4}$$

Using Eqs. (1)–(4), we can calculate the dependence of the active component of the impedance R of the transducer on the frequency f.

Figure 2 shows plots of $R(f)$ calculated for a wedge-shaped (tapering) ZnO transducer with an Al superlayer and underlayer on a $Y_3Al_5O_{12}$ crystal for various values of the ratio h_1/h_2. Curves *1–4* correspond to $h_1/h_2=1$, 0.66, 0.43, and 0.25. The ratios of the thicknesses of the superlayer and superlayer to the thickness of the piezoelectric were set equal to $p/h=0.2$ and $g/h=0.2$, respectively. The following physical constants of ZnO, $Y_3Al_5O_{12}$, and Al, respectively, were used in the calculations: speed of sound [10^3 m/s] — 6.1, 8.6, and 6.32; acoustic wave impedance [10^6 kg/m$^2\cdot$s] — 34.4, 39.2, and 17.1; dielectric constant and electromechanical coupling constant of ZnO — 8.84 and 0.28.

The plots presented demonstrate the possibility of significantly expanding the frequency band of a transducer by using a wedge-shaped (tapering) piezoelectric layer. For example, a comparison of curves *1* and *4* shows that if the ratio between the thicknesses of the piezoelectric at the beginning and end of the transducer $h_1/h_2=0.25$, the frequency band at the 3 dB level increases from 0.8 to 8.4 GHz, i.e., by roughly 10 fold. Expansion of the band leads to a decrease in the radiation resistance, which characterizes the conversion efficiency. However, in some cases, this circumstance can be utilized to optimize an acoustic device. In particular, in the problem discussed in Ref. 1, a multielement filter-type transducer with variable parameters can be used to expand the frequency band of an acousto-optic Bragg detector. In such a transducer the conversion of electromagnetic energy into

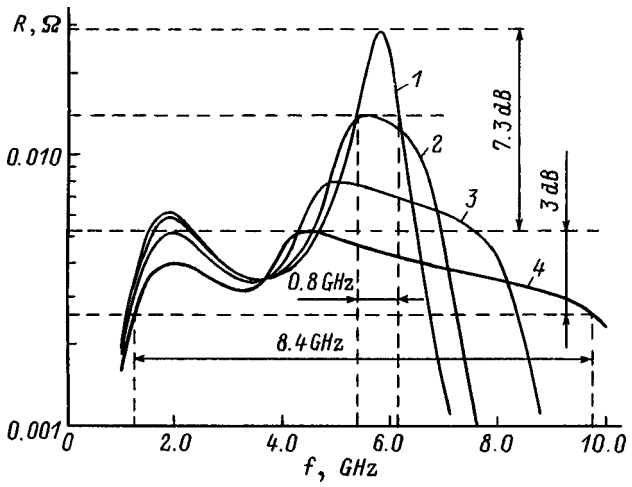

FIG. 2. Frequency dependence of the radiation for several values of the thickness drop along the wedge-shaped transducer: $h_1=0.5$ (*1*), 0.4 (*2*), 0.3 (*3*), and 0.2 μm (*4*) $h_2=0.5$ (*1*), 0.6 (*2*), 0.7 (*3*), and 0.8 μm (*4*).

acoustic energy takes place successively, i.e., from element to element. In this case, the optimal conversion coefficient of one cell, which generally differs from its maximum value[3] and must be reduced artificially, must be realized to achieve the highest diffraction efficiency.

The proposed method for expanding the frequency band can be used to develop high-frequency acousto-optic deflectors or for creating broad-band (about 8 Hz for the calculation presented) delay lines, as well as other information-processing devices.

[1] V. V. Petrov, Pis'ma Zh. Tekh. Fiz. **22**(22), 11 (1996) [Tech. Phys. Lett. **22**, 909 (1996)].
[2] M. A. Grigor'ev, V. V. Petrov, and A. V. Tolstikov, Radiotekh. Elektron. **35**, 1977 (1990).
[3] M. A. Grigor'ev, V. V. Petrov, and A. V. Tolstikov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **28**, 732 (1985).

Translated by P. Shelnitz

Estimation of the concentration of complex negative ions resulting from radioactive contamination of the troposphere

K. A. Boyarchuk

General Physics Institute, Russian Academy of Sciences, 117942 Moscow, Russia
(Submitted October 23, 1997; resubmitted August 21, 1998)
Zh. Tekh. Fiz. **69**, 74–76 (March 1999)

A model of the formation of charged particles and of the negative ions of nitrogen and carbon oxides is proposed, and the possible concentration of complex negative ions resulting from radioactive contamination of the troposphere is estimated. © *1999 American Institute of Physics.* [S1063-7842(99)01403-8]

Changes in the concentration of charged particles (ions and electrons) occur in the atmosphere under the action of ionizing radiations, for example, cosmic rays, or in cases of radioactive contamination. In Refs. 1 and 2 we estimated the changes in the concentrations of elementary ions during $\sim 10^{-7}$ s, but the procedure used was not complete, since the influence of the principal atmospheric gases was not taken into account. These variations in the composition of the atmosphere naturally depend on the initial concentrations of the atmospheric gases, i.e., the oxides of nitrogen and carbon, as well as water vapor. The purpose of the present work was to simulate the variation of the concentrations of ions of the type NO_x^- and CO_2^- in the troposphere under the effects of ionizing radiation.

As the radiation of radioactive elements (electromagnetic radiation, viz., γ and x rays, or fluxes of fast particles, viz., α particles, β particles, protons, and neutrons) passes through an air medium, it expends its energy on the ionization and excitation of molecules, and the secondary electrons appearing as a result of such ionization have the main effect on the medium. The energy of these electrons lies in the range from thermal energies to the energies of primary fast particles or γ quanta. A large portion of the secondary electrons have an energy smaller than the ionization energy, but they are capable of exciting molecules and thereby facilitating the dissociation of molecules or their ionization by other particles and electrons. The ionization of air molecules clearly begins at an electron energy equal to the ionization energy of the principal components: N_2 (15.576 eV) and O_2 (12.2 eV). Dissociative ionization can occur as the energy of the ionizing particles increases. As a result, at the very beginning of the ionization of air we have the following set of charged particles: N_2^+, N^+, O_2^+, O^+, and electrons. The yield of a specific species of ions depends on the energy of the secondary electrons.[3]

Thus, positive ions and free electrons form in the atmosphere under the action of radiation; the subsequent development of the electron-ion system depends on the composition, density, humidity, and temperature of the air and the character of the ionizing radiation. Electrons having an energy close to that of a thermal electron are trapped by positive ions, as well as by molecules and atoms having an electron affinity.[4] The negative ions thus formed then recombine with the positive ions. An important role in the outcome of these reactions is played by the detachment of electrons from negative ions during their collisions with excited air molecules.[5] Therefore, ultraviolet radiation from the sun or other sources of excited molecules has an influence here.

The negative ions of the oxides of nitrogen and carbon form along three main pathways:[6,7]

$$\left.\begin{matrix} e+O_3 \\ O^-+O_2 \end{matrix}\right\} \rightarrow O_3^- \rightarrow NO_2^- \rightarrow NO_3^-, \tag{1}$$

$$e+O_2 \rightarrow O_2^- \rightarrow O_3^- \rightarrow CO_3^- \rightarrow NO_2^- \rightarrow NO_3^-, \tag{2}$$

$$e+O_2 \rightarrow O_2^- \rightarrow O_4^- \rightarrow CO_4^- \rightarrow NO_3^-. \tag{3}$$

The O_4^- ion can be regarded as the ionic complex $O_2^- \cdot O_2$ [similarly, the CO_4^- ion can be regarded as the complex $O_2^- \cdot CO_2$ (Ref. 8)]. The reactions leading to the formation and dissociation of the O_4^- ion, i.e., $O_2^- + O_2 + M \leftrightarrow O_2^- \cdot O_2 + M$, are more characteristic of the upper stratosphere, the rate of the reverse reaction ($\sim 10^{-6}$ cm$^3 \cdot$ s^{-1}) being higher than the rate of the forward three-particle reaction ($\sim 10^{-30}$ cm$^6 \cdot$ s^{-1}). Therefore, we shall consider a simplified model of the ionized troposphere, which takes into account only the first two pathways as the most significant.[9]

The reactions between electrons and ions listed in Table I are the most probable processes. We shall assume that the collision of a fast electron with a neutral molecule results in the appearance of a secondary electron and either a molecular or a monatomic positive ion and that these processes take place with equal probabilities both on oxygen molecules and on nitrogen molecules. Let W be the probability of the appearance of a molecular ion in each ionization act, and let Q [cm$^{-3} \cdot$ s^{-1}] be the rate of the generation of electrons. Direct ionization with the formation of a molecular ion of nitrogen or oxygen takes place predominantly in the lower layers of the atmosphere; therefore, we shall estimate the probability of the appearance of a molecular ion in an ionization act as $W=0.75$.[3] The generation rates are $A_i = N_i Q W / N_L$ for molecular ions and $A_i = N_i Q(1-W)/N_L$ for monatomic ions, where N_i is the concentration of the respective molecules [cm^{-3}] and $N_L = 2.687 \times 10^{19}$ cm^{-3} is the Loschmidt number.

TABLE I. Principal Ion-Molecule Reactions in the Lower Atmosphere.

Reaction No.	Reaction	a_n, $cm^3 \cdot s^{-1}$	Reference
1	$e+O^+ \rightarrow O$	3.0×10^{-12}	Ref. 1
2	$e+N^+ \rightarrow N$	3.0×10^{-12}	Ref. 1
3	$e+O_2^+ \rightarrow O+O$	2.2×10^{-7}	Ref. 1
4	$e+N_2^+ \rightarrow N+N$	2.9×10^{-7}	Ref. 1
5	$e+NO^+ \rightarrow N+O$	4.1×10^{-7}	Ref. 1
6-1	$e+O_2+X \rightarrow O_2^- + X$	2.5×10^{-30} $[cm^6 \cdot s^{-1}]$	Ref. 1
6-2	$e+O_2 \rightarrow O^- + O$	4.8×10^{-14}	Ref. 1
7-1	$e+O_3 \rightarrow O_3^-$	$\sim10^{-13}$	Ref. 4
7-2	$e+O_3+O_2 \rightarrow O_3^- + O_2$	10^{-13}	Ref. 4
8	$A^- + B^+ \rightarrow A+B^*$	$\sim10^{-6}$	Ref. 5
9	$O^- + O_2^* \rightarrow O_3 + e$	3.0×10^{-10}	Ref. 6
10	$O_2^- + O_2^* \rightarrow O_2 + O_2 + e$	2.0×10^{-10}	Ref. 6
11	$O^+ + O_2 \rightarrow O_2^+ + O$	4.0×10^{-11}	Ref. 1
12	$O^+ + N_2 \rightarrow NO^+ + N$	4.0×10^{-12}	Ref. 1
13	$N^+ + O_2 \rightarrow O_2^+ + N$	5.0×10^{-10}	Ref. 1
14	$N^+ + O_2 \rightarrow NO^+ + O$	5.0×10^{-10}	Ref. 1
15	$N_2^+ + O_2 \rightarrow O_2^+ + N_2$	$\sim10^{-11}$	Ref. 1
16	$O^- + O_2 + M \rightarrow O_3^- + M$	$\sim1.1\times10^{-30}$ $(M=N_2, O_2)$	Ref. 4
17	$O_2^- + O_3 \rightarrow O_3^- + O_2$	6×10^{-10}	Ref. 7
18-1	$O_2^- + NO_2 \rightarrow NO_2^- + O_2$	7×10^{-10}	Ref. 3
18-2	$O_2^- + N_2O \rightarrow NO_2^- + NO$	2×10^{-14}	Ref. 3
19	$O_3^- + CO_2 \rightarrow CO_3^- + O_2$	5.5×10^{-10}	Ref. 3
20-1	$O_3^- + NO \rightarrow NO_2^- + O_2 \rightarrow NO_3^- + O$	8×10^{-10}	Ref. 3
20-2		$\sim10^{-11}$	Ref. 3
21-1	$O_3^- + NO_2 \rightarrow NO_2^- + O_3 \rightarrow NO_3^- + O_2$	7×10^{-10}	Ref. 3
21-2		2.8×10^{-10}	Ref. 3
22	$NO_2^- + O_3 \rightarrow NO_3^- + O_2$	1.2×10^{-10}	Ref. 3
23	$NO_2^- + NO_2 \rightarrow NO_3^- + NO$	2×10^{-13}	Ref. 7
24	$CO_3^- + NO_2 \rightarrow NO_3^- + CO_2$	2.0×10^{-10}	Ref. 3

The kinetic equations of the charged particles can be written, according to Refs. 1 and 2, in the following form:

$$\frac{d[e]}{dt} = Q - [e]\left(\sum a_i^{\text{at}}[N_i^n] + \sum a_i^{\text{rec}}[N_i^+] - \sum a_i^{\text{det}}[N_m^{*n}]\right),$$

$$\frac{d[N_x^+]_i}{dt} = A_i Q - \sum a_i^{\text{rec}}[N_i^+][N_j^-] - \sum a_i^n[N_i^+][N_j^n],$$

$$\frac{d[N_x^-]_i}{dt} = \sum a_i^{\text{at}}[e][N_i^n] + \sum a_i^n[N_i^-][N_j^n] - \sum a_k^n[N_l^-][N_m^n] - \sum a_i^{\text{rec}}[N_i^+][N_j^-] - \sum a_i^{\text{det}}[e][N_m^{*n}]. \quad (4)$$

Here $[N_i^{\pm}]$ are the concentrations of positive and negative ions; $[N_i^n]$ is the concentration of neutral molecules; $[N_i^*]$ is the concentration of excited molecules; and a^{at}, a^{rec}, a^{det}, and a^n are the rates of attachment, recombination, electron detachment with the destruction of negative ions, and the interaction with neutral molecules, respectively. It is not difficult to see that this system satisfies the conservation law of electric charge in the system: $[e]+\Sigma[N_i^-]=\Sigma[N_j^+]$.

Let us consider the solution of this system for the case of $Q=10^8$ $cm^{-3}\cdot s^{-1}$ (which corresponds to the shutdown emission level from the ventilation tube of an atomic electric power plant[10]). We assume that $[N^n] \gg [N^{\pm}]$. Figure 1 presents the result of the numerical solution of the system (4) by the Runge–Kutta method in a time interval up to 0.1 s. We set the natural concentrations of nitrates in the troposphere equal to the following values: O_3 — 5×10^{11} cm^{-3}, NO — 10^9 cm^{-3}, NO_2 — 2×10^{11} cm^{-3}, CO_2 — 10^{10} cm^{-3}.[3]

Thus, on the basis of our simplified model, NO_3^- ions should accumulate in the atmosphere under the action of

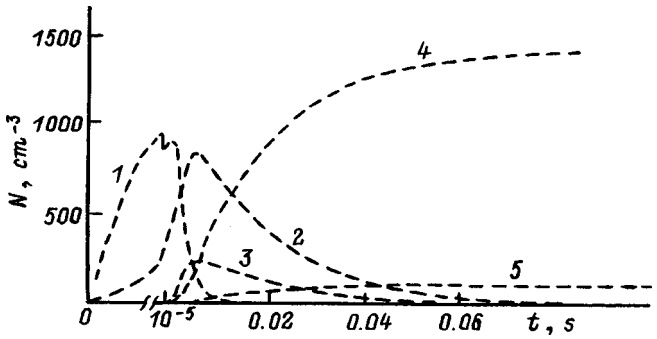

FIG. 1. Temporal variation of the concentration of ions N $[cm^{-3}]$ (*1* — O_2^-, *2* — O_3^-, *3* — NO_2^-, *4* — NO_3^-, *5* — CO_3^-) for the case of the natural concentration of nitrates in the troposphere ($[NO]\sim10^9$ cm^{-3}, $[NO_2]\sim2\times10^{11}$ cm^{-3}).

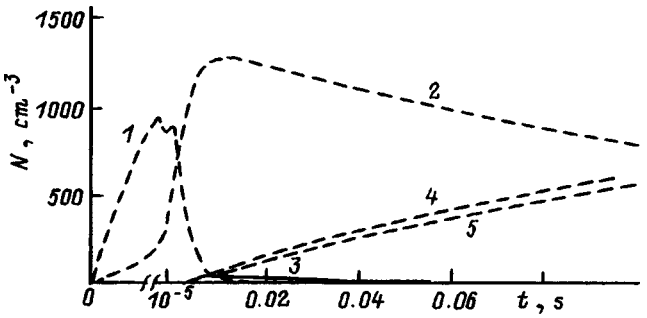

FIG. 2. Temporal variation of the concentration of ions N [cm^{-3}] (*1* — O_2^-, *2* — O_3^-, *3* — NO_2^-, *4* — NO_3^-, *5* — CO_3^-) for the case of the natural concentration of nitrates in the troposphere ($[NO]\sim 10^8$ cm^{-3}, $[NO_2]\sim 2\times 10^{10}$ cm^{-3}).

ionizing radiation. It should be noted that the formation of complex ions based on NO_3^- in the earth's troposphere can occur as a result of the oxidation of nitrogen when it reacts with CO_2 and O_2 [pathway (2)]. Generally speaking, carbon dioxide plays a lesser role than oxygen in an air medium, since a considerable number of O_2^- and O_3^- ions appears in the atmosphere under the action of ionizing radiation and the formation of NO_x^- takes place mainly along pathway (1).

Since the earth's troposphere contains an enormous number of water-vapor molecules ($\sim 10^{17}$ cm^{-3}), which have an appreciable dipole moment $p_{H_2O}=1.87$ D, the hydration of elementary ions and the formation of ionic complexes of the type $NO_3^-\cdot(H_2O)_n$ are fairly fast processes, the values $n=2-3$ being characteristic of the troposphere.[3] At the levels of ionization of the atmosphere considered, the concentration of elementary ions is small in comparison to the concentration of water-vapor molecules; therefore, it can be assumed to within sufficient accuracy that all negative elementary ions are hydrated. Consequently, the results of the calculation of the concentrations of NO_3^- and CO_3^- ions shown in Figs. 1 and 2 correspond to the concentration of the ionic complexes of the types $NO_3^-\cdot(H_2O)_n$ and $CO_3^-\cdot(H_2O)_n$.

The proposed simplified model is sensitive to the initial concentrations of CO_2 and NO_x in the atmosphere. If the concentration of the compounds NO_x in air is an order of magnitude smaller than their natural content and if the concentration of CO_2 is higher, the $CO_3^-\cdot(H_2O)_n$ ions begin to dominate, and the formation of $NO_3^-\cdot(H_2O)_n$ slows. This result is consistent with the calculation in Ref. 11, but the production of the basic ions is slower in our model. Figure 2 presents the result of the numerical solution of the system (4) for concentrations of nitrates an order of magnitude smaller: NO — 10^8 cm^{-3}, NO_2 — 2×10^{10} cm^{-3}. A slower decrease in the concentration of ozone ions is also noticeable here, and the concentration of CO_3^- ions becomes comparable to the concentration of NO_3^- ions. Hence it follows that under certain conditions this mechanism can have an influence on the total concentration of ozone in the atmosphere.

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-16540).

[1] K. A. Boyarchuk and Yu. P. Svirko, Pis'ma Zh. Tekh. Fiz. **22**(14), 47 (1996) [Tech. Phys. Lett. **22**, 575 (1996)].
[2] K. A. Boyarchuk, Izv. Ross. Akad. Nauk, Fiz. Atmos. Okeana No. 2 (1997).
[3] V. V. Smirnov, *Ionization in the Troposphere* [in Russian], Gidrometeoizdat, St. Petersburg (1992), 312 pp.
[4] Yu. P. Raizer, *Gas Discharge Physics*, 2nd ed., Springer, Berlin–New York (1997), 536 pp.
[5] S. Ya. Pshezhetskiĭ and M. T. Dmitriev, *Radiation-Induced Physicochemical Processes in an Air Medium* [in Russian], Atomizdat, Moscow (1978), 182 pp.
[6] H. Kawamoto and T. Ogawa, Planet. Space Sci. **34**, 1229 (1986).
[7] E. E. Ferguson, in *Kinetics of Ion-Molecule Reactions (NATO Advanced Study Institutes Series, Ser. B: Physics, Vol. 40)*, P. Ausloos (ed.), Plenum Press, New York (1979), pp. 377–401.
[8] H. S. W. Massey, *Negative Ions*, Cambridge University Press (1976); Mir, Moscow (1979)], 754 pp.
[9] E. Arijs, Planet. Space Sci. **40**, 255 (1992).
[10] K. A. Boyarchuk, G. A. Lyakhov, Yu. P. Svirko *et al.*, Bull. Russ. Acad. Sci. Phys. Suppl. Phys. Vib. **59**, 222 (1995).
[11] Ya. Ĭ. Sal'm and A. M. Luts, Uch. Zap. Tartu. Gos. Univ. **809**, 64 (1988).

Translated by P. Shelnitz

Asymmetric passage of current through a laser-produced plasma plume

A. N. Panchenko and V. F. Tarasenko

Institute of High-Current Electronics, Russian Academy of Sciences, Siberian Branch, 639055 Tomsk, Russia

S. I. Yakovlenko

General Physics Institute, Russian Academy of Sciences, 117942 Moscow, Russia

(Submitted November 11, 1996; resubmitted January 20, 1998)
Zh. Tekh. Fiz. **69**, 77–79 (March 1999)

The observation of the asymmetric passage of current in a plasma interrupter under certain conditions is reported. The effect can be attributed to different temperatures of the electrodes for the laser production of plasma in the interrupter. © *1999 American Institute of Physics.* [S1063-7842(99)01503-2]

1. The current cutoffs in a plasma of metal vapors formed by laser radiation were investigated in Refs. 1–3, and it was shown that the stable cutoff of current with an amplitude up to 10 kA is achieved under certain conditions, including a periodic pulsed regime. The interest in the study and utilization of the conductivity decreases in plasmas created by different methods is due to the employment of plasma current interrupters in the development of high-power pulse generators and in inductive energy storage.[4,5]

This paper reports the observation of the asymmetric passage of current pulses through a laser plasma with a spark-gap current interrupter connected to it in parallel and offers a qualitative interpretation of this effect.

2. The experimental setup is shown in Fig. 1. The laser plasma was created precisely as in Refs. 1–3. Two stainless steel electrodes *2* and *3* were placed in vacuum chamber *1*. Potential electrode *2* has a cylindrical depression containing an aluminum insert at its base. The output of a XeCl laser (308 nm, 12 mJ) was focused through grounded electrode *3* onto the aluminum insert. The laser radiation led to the formation of a plasma, which propagated systematically from electrode *2* to electrode *3*.

Unlike the experiments in Refs. 1–3, where plasma current cutoff was investigated, in the present work an air spark gap (*SG*) with pointed electrodes was connected in parallel to the electrodes of the vacuum chamber. This allowed us to investigate the regime for switching the current from the plasma interrupter to the spark gap upon breakdown of the latter due to an overvoltage pulse appearing when the discharge current through the laser plasma decreases.

When discharger *D* was tripped, the voltage pulse from capacitor *C* was fed to electrode *2* and to the spark gap after the plasma short-circuited electrodes *2* and *3*. The residual pressure in the vacuum chamber was $\sim 10^{-5}$ Torr. The distance between electrodes *2* and *3* was 8.5 mm, and the gap between the electrodes of air spark gap *SG* was 10 mm. The total current I and the current through the spark gap I_2 were recorded. The current through the plasma interrupter (electrodes *2* and *3*) was calculated using the formula $I_1 = I - I_2$.

3. The main results of the experiments performed are as follows. When the voltage was small or the spark gap was large, no breakdown of air occurred in the spark gap during cutoff of the current and the formation of an overvoltage pulse. Under these conditions all the energy is dissipated in the plasma current interrupter, and the current and voltage oscillograms correspond to the regimes considered in Refs. 1–3.

When the charging voltage of capacitor C was sufficiently high ($U_0 > 8$ kV) or the spark gap was diminished, breakdown took place, and up to 90% of the total current passed through the spark gap. In these regimes the discharging has an oscillatory character similar to oscillogram *1* in Fig. 2, and no current flows through the plasma of the plasma interrupter during the second half-period.

The most interesting result in the case of an oscillatory regime is that at certain values of U_0 (in the present experiments, at $U_0 = 8$ kV, Fig. 2) the current in the plasma interrupter depends strongly on its direction of flow. It can even be stated that the effective conductivity of the laser plasma is strongly asymmetric with respect to that the direction. Asymmetric current oscillations through the plasma interrupter take place under a definite choice of conditions (the parameters of the discharge circuit, the width of spark gap, the

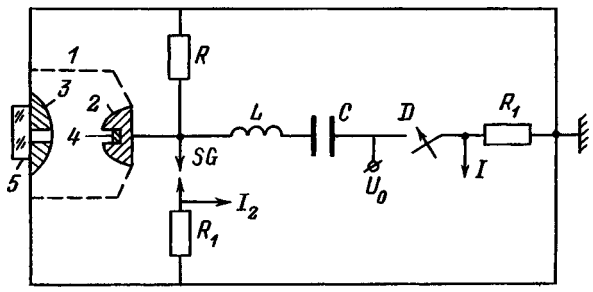

FIG. 1. Diagram of the experimental setup: *1* — vacuum chamber; *2*, *3* — electrodes of the plasma current interrupter; *4* — aluminum insert; *5* — quartz window; C — storage capacitor; L — inductance; *SG* — air spark gap; U_0 — charging voltage; R_1, R — resistances; D — controllable discharger; I — total current; I_2 — current through the spark gap.

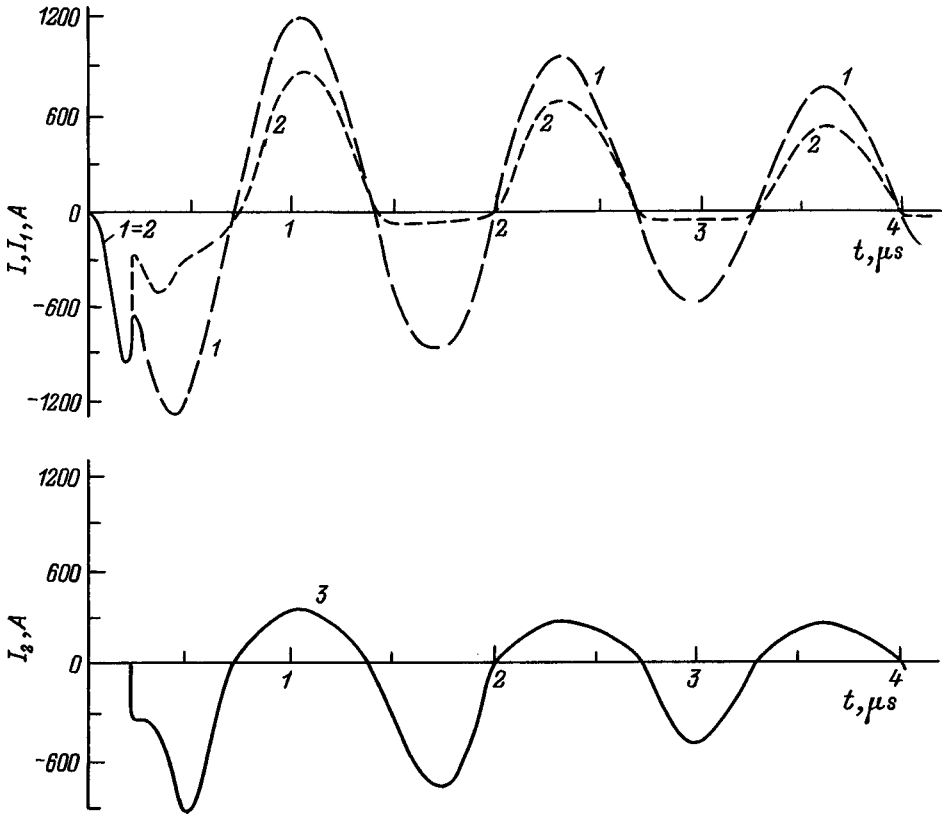

FIG. 2. Oscillograms of the total current I (*1*), the current through the plasma interrupter I_1 (*2*), and the current through the spark gap I_2 (*3*).

laser output energy, the distance between electrodes *2* and *3*, U_0, etc.). For example, all other conditions being equal, when the initial voltage was increased slightly (by ~ 1 kV), asymmetric oscillations occurred over the course of one or two periods, and then the bulk of the current flowed either through the air gap or through the plasma. When the change in U_0 was large, no nonmonotonic flow of current was observed. On the other hand, when $U_0 = 8$ kV and the laser output energy was varied by 10%, current could flow through the spark gap only during the negative half-periods of the total current and through the plasma interrupter only during the positive half-periods.

4. We attribute the asymmetric conductivity of the laser plasma to the following circumstances. The asymmetry in the current appears mainly because of the different temperatures of the left-hand (*3*) and right-hand (*2*) electrodes (Fig. 1). When the polarity corresponds to the flow of electrons from the plasma into electrode *3*, a layer, in which a separation of charges occurs, forms near that electrode. Ions are retarded by the external field, and an appreciable portion of the electrons can pass through the vacuum layer and reach electrode *3*. When the current changes direction, the external field squeezes out electrons, but cold electrode *3* cannot ensure sufficient emission. Near the surface of electrode *2*, which is heated by the laser pulse and is in contact with the denser hot plasma, such blocking of the electron current does not occur.

Let us make some estimates. If we assume that about 20% of the energy of a laser pulse (12 mJ) is expended on ionization, 4×10^{15} electrons and ions are generated during a pulse. Setting the length of the plasma plume approximately equal to the distance between the electrodes, i.e., 1 cm, and its diameter to a value of the order of half of the length, i.e., $d \approx 5$ mm, (since the plasma expands from the depression), we obtain an estimate of the electron density $N_e \approx 2 \times 10^{16}$ cm^{-3}. Next, taking the electron temperature $T_e \approx 5$ eV, for the Debye radius we obtain $r_D = \sqrt{T_e/4\pi e^2 N_e} \approx 10^{-5}$ cm and for the field which ensures discontinuity of the plasma at the Debye radius we obtain $E_D = T_e/er_D \approx 4 \times 10^5$ V/cm. Hence it follows for a voltage $U \approx 8$ kV (the overvoltage pulse is usually significantly greater than the charging voltage[1–3]) that the external field can heat the plasma over a thickness $a \approx U/E_D \approx 0.2$ mm.

According to the familiar three-halves rule, such a voltage and distance between the electrodes correspond to a current

$$J = \frac{1}{9\pi}\frac{1}{a^2}\sqrt{\frac{2e}{m_e}}\frac{\pi d^2}{4}U^{3/2} \approx 900 \text{ A}.$$

This is consistent with the observed value of the peak current (Fig. 2).

When the polarity from electrode *3* is reversed, electrons depart, and the passage of current through the near-electrode layer could be mediated by ions. However, according to the three-halves rule, the ion current should be $\sqrt{m_i/m_e} \approx 200$ times smaller (m_i is the mass of aluminum), whereas a reverse current, which is only 1.5 times smaller, is observed experimentally for the conditions in Fig. 2. Apparently, the observed current is also provided by some electron emission from electrode *3*, which can be small if the plasma concen-

tration in the interrupter varies in response to variation of the laser output energy.

The time for formation of the blocking layer is fairly short. It corresponds to the time τ_i for departure of the ions from the electrode to a distance a under the action of the field $E \sim U/2a$, i.e, $\tau_i \approx 2a\sqrt{m_i/eU} \sim 5$ ns. This means that the width of the layer follows the variation of the voltage on the electrodes.

When $U_0 > 8$ kV, the flow of current through the spark gap does not depend on the polarity of the electrodes; therefore, the conditions on the electrodes and their temperatures do not differ significantly. The differences in the amplitude and form of the current for $U_0 = 8$ kV (Fig. 2) are caused by the influence of the plasma interrupter, which is connected in parallel to spark gap.

5. Thus, the observation of asymmetric current flow in a plasma interrupter under certain conditions has been reported in this paper. This effect can be attributed to the different temperatures of the electrodes for the laser production of plasma in the interrupter.

[1] A. N. Panchenko and V. F. Tarasenko, Zh. Tekh. Fiz. **58**, 1551 (1988) [Sov. Phys. Tech. Phys. **33**, 930 (1988)].
[2] A. N. Panchenko and V. F. Tarasenko, Fiz. Plazmy **16**, 1061 (1990) [Sov. J. Plasma Phys. **16**, 616 (1990)].
[3] A. N. Panchenko, V. F. Tarasenko, and S. I. Yakovlenko, Zh. Tekh. Fiz. **60**(10), 42 (1990) [Sov. Phys. Tech. Phys. **35**, 1136 (1990)].
[4] V. V. Kremnev and G. A. Mesyats, *Methods for Multiplying and Transforming Pulses in High-Current Electronics* [in Russian], Nauka, Novosibirsk (1987), 226 pp.
[5] A. Quenther, M. Kristiansen, and T. Martin, *Opening Switches, Vol. 1*, Plenum Press, New York–London (1987), 280 pp.

Translated by P. Shelnitz

Employment of absorbing *a*-C:H films in reflective liquid-crystal light modulators

E. A. Konshina and A. P. Onokhov

S. I. Vavilov State Optical Institute, All-Russian Scientific Center, 199034 St. Petersburg, Russia
(Submitted May 15, 1998)
Zh. Tekh. Fiz. **69**, 80–81 (March 1999)

The effectiveness of employing *a*-C:H films as light-blocking layers in reflective liquid-crystal light modulators with *a*-Si:H and *a*-Si:C:H photosemiconductor layers is investigated. The possibility of reducing the light flux penetrating into a photosemiconductor by 100 fold using an *a*-C:H film with a thickness of 1 μm and an extinction coefficient equal to 5×10^4 cm^{-1} at a wavelength of 632.8 nm is demonstrated. © *1999 American Institute of Physics.* [S1063-7842(99)01603-7]

One complex technical problem in developing reflective liquid-crystal light-controlled modulators is ensuring optical separation between the writing and reading radiations. The dielectric and metallic mosaic mirrors used for this purpose cannot really provide for complete reflection of the reading light. A light field in the visible wavelength range penetrating into the semiconductor layer adjoining the mirror is absorbed in it and has an effect on light modulation. This reduces the dynamic intensity range of the images recorded, the signal-to-noise ratio, and the sensitivity of a reflective liquid-crystal modulator.[1] More effective separation of the writing and reading light fluxes can be attained when a layer which absorbs the reading radiation is inserted between the mirror and the photosemiconductor. The layer employed for this purpose in a liquid-crystal modulator with a CdS photosemiconductor is a narrower-band CdTe semiconductor with a thickness of 2 μm, an extinction coefficient no less than 10^5 cm^{-1} at a wavelength of 525 nm, and a surface resistance of about 10^{11} Ω (Ref. 2). At the same time, it has been reported that a ''black'' diamond-like carbon monolayer of thickness 1 μm with transmission in the visible region of the spectrum equal to less than 2% can be used to enhance the contrast of a matrix display by depositing it on the areas between the matrix elements.[3]

The present work is the first investigation of the efficiency of the blocking of light at a wavelength of 632.8 nm from *a*-Si:H and *a*-Si:C:H photosemiconductor layers using an absorbing film of hydrogenated amorphous carbon (*a*-C:H).

The *a*-C:H films were obtained using the chemical deposition of hydrocarbon vapors in a dc glow-discharge plasma. Previously performed investigations of the optical constants of films obtained from various hydrocarbons showed that the films obtained by this method from acetylene vapor have the highest extinction coefficient at a wavelength of 632.8 nm.[4,5] Therefore, in this work we used the latter to obtain absorbing films. The films were deposited at ambient temperature on glass substrates with preliminarily deposited layers of a transparent conducting electrode based on indium and tin oxides (In_2SnO_3) and the photosemiconductor *a*-Si:H or *a*-Si:C:H. The latter were obtained by chemical deposition from silane vapor or a mixture of silane and methane in an rf discharge.[6]

The optical constants of the *a*-C:H films depend on the kinetics of the deposition process in the glow-discharge plasma, and their absorption coefficient can be raised by lowering the deposition rate.[4,5] In this work the deposition rate was varied by varying the acetylene pressure in the vacuum chamber in the range from 0.1 to 0.01 Pa or the acetylene concentration in an acetylene-argon mixture at a constant interelectrode voltage of 900 V. A He-Ne laser was used to measure light absorption at 632.8 nm. The thickness of the *a*-C:H films was determined using an MII-4M microinterferometer to within a relative measurement error less than 10%. The extinction coefficient was calculated according to the Lambert–Beer law. The experimental values of the extinction coefficient (α) are plotted as a function of the deposition rate (v) of *a*-C:H films on the surface of a glass/In_2SnO_3/*a*-Si:H (or *a*-Si:C:H) structure in Fig. 1. As

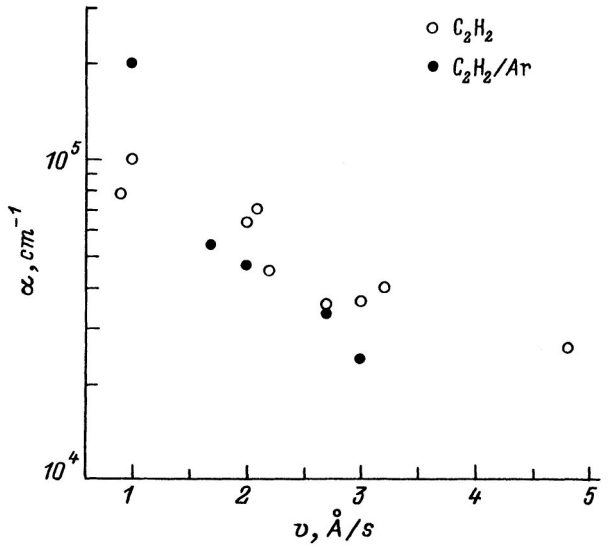

FIG. 1. Dependence of the extinction coefficient (α) at 632.8 nm on the deposition rate (v) of *a*-C:H films on glass substrates with layers of a transparent conducting In_2SnO_3 electrode and a photosemiconductor deposited on them.

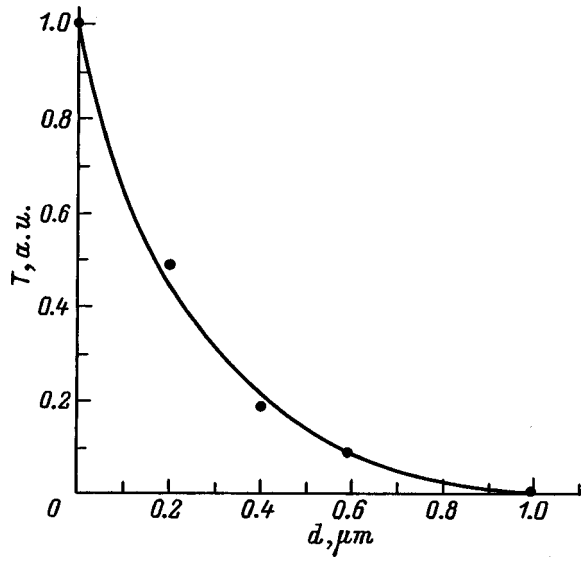

FIG. 2. Dependence of the relative transmission (T) of a thin-film In_2SnO_3/a-Si:H/a-C:H structure on the thickness (d) of an absorbing film with $\alpha = 5\times10^4$ cm^{-1}.

the rate was varied in the range from 1 to 5 Å/s, the values of α at 632.8 nm varied by an order of magnitude in the range from 2×10^5 to 2.5×10^4 cm^{-1}.

The efficiency of reducing the transmission (T) of an a-Si:H layer using an absorbing a-C:H film as a function of its thickness (d) is illustrated by Fig. 2. When the thickness of an a-C:H film with $\alpha = 5\times10^4$ cm^{-1} was increased to 1 μm, the transmission of the structure at 632.8 nm decreased by more than two orders of magnitude relative to its initial value. A resistivity of the order of $\sim10^{11}-10^{12}$ $\Omega\cdot$cm is characteristic of a-C:H films with $\alpha \leqslant 5\times10^4$ cm^{-1} at 632.8 nm. The use of a-C:H films with $\alpha > 5\times10^4$ cm^{-1} permits more efficient weakening of the intensity of the impinging light. For example, a film with $\alpha = 1\times10^5$ cm^{-1} obtained with a deposition rate of 1 Å/s weakened the transmission of an a-Si:H layer by a factor of $\sim$500 already at a thickness of $\sim0.5\mu$m. However, when an In_2SnO_3/a-Si:H/a-C:H/Al structure (a mosaic mirror) with such a film was employed in a liquid-crystal modulator, we were unable to obtain spatial light modulation. As a comparison of the current-voltage characteristics of an In_2SnO_3/a-SiH structure before and after the deposition of an a-C:H film showed, this can be due to a significant increase in the value of the dark current and a loss of resolving power. The changes in absorption in the visible portion of the spectrum are accompanied by changes in the electronic structure and the optical gap width of a-C:H films.[7] It was established in that study that the increase in the extinction coefficient at 632.8 nm ($\alpha > 5\times10^4$ cm^{-1}) was accompanied by a decrease in the optical gap width to $\sim$1 eV and a decrease in the resistivity of a-C:H films to values of the order of $\sim10^7$ $\Omega\cdot$cm.

It was shown as a result of the experiments performed that the extinction coefficient of a-C:H films at 632.8 nm can be varied in the range from 2×10^5 to 2.5×10^4 cm^{-1} by varying their deposition rate from an acetylene or acetylene-argon dc glow-discharge plasma from 1 to 5 Å/s. A 100-fold reduction in the intensity of the light passing through the glass/In_2SnO_3/a-Si:H structure used in liquid-crystal modulators can be obtained using a film of thickness 1 μm with an extinction coefficient $\sim5\times10^4$ cm^{-1}.

The employment of absorbing a-C:H films to block light in the visible wavelength range from a-Si:H and a-Si:C:H photosemiconductor layers is a new technical solution for the problem of optical separation between the writing and reading light and opens up prospects for developing new reflective light-controlled liquid-crystal modulators. At the same time, the achievement of high spatial resolution in liquid-crystal modulators based on the thin-film structures investigated along with optimization of the optical absorption and thickness of the light-blocking a-C:H layer calls for optimization of its electrical properties and matching them to the properties of the photosemiconductor.

[1] A. A. Vasil'ev, D. Kasasent, I. N. Kompanets, and A. V. Parfenov, *Spatial Light Modulators* [in Russian], Radio i Svyaz', Moscow (1987), 320 pp.
[2] W. P. Bleha, L. P. Lipton, and E. Wiener, Opt. Eng. (Bellingham) **17**, 371 (1978).
[3] B. Singh, S. McClelland, F. Tams *et al.*, Appl. Phys. Lett. **57**, 2288 (1990).
[4] E. A. Konshina and V. A. Tolmachev, Zh. Tekh. Fiz. **65**, 175 (1995) [Tech. Phys. **40**, 97 (1995)].
[5] V. A. Tolmachev and E. A. Konshina, Diamond Relat. Mater. **5**, 1397 (1996).
[6] *Amorphous Semiconductor Technologies and Devices*, edited by Y. Hamakawa, North-Holland, Ohmsha–Amsterdam (1981).
[7] E. A. Konshina, Fiz. Tverd. Tela (St. Petersburg) **37**, 1120 (1995) [Phys. Solid State **37**, 610 (1995)].

Translated by P. Shelnitz

Pulsed method for measuring the capacitance of semiconductor structures using a ballast capacitor

V. V. Monakhov and A. B. Utkin

St. Petersburg State University, 198904 St. Petersburg, Russia
(Submitted May 19, 1998)
Zh. Tekh. Fiz. **69**, 82–84 (March 1999)

An easily implemented method for measuring the capacitance of semiconductor structures using small-amplitude current pulses is described. It is shown that the accuracy of the method described can be improved significantly when a ballast capacitor is employed and, in addition, a calibration procedure is performed. The influence of the ballast capacitor on the measurement process is analyzed in detail. © *1999 American Institute of Physics.* [S1063-7842(99)01703-1]

Capacitance-voltage characteristics are widely employed in the investigation of semiconductors, and, therefore, perfecting the experimental techniques for measuring them is a crucial concern of researchers.[1–3] This paper describes a method which permits improvement of the accuracy of capacitance-voltage measurements performed according to the two-pulse scheme[4] in electrolyte-semiconductor (ES) and electrolyte-insulator-semiconductor (EIS) systems and allows measurement of the capacitance-voltage characteristics of metal-insulator-semiconductor (MIS) structures.

The method involves charging of the structure under investigation and a specially selected ballast capacitor connected to it in parallel by a small current followed by their discharging by a current pulse of the same amplitude and duration, but of opposite polarity. When measurements are performed in systems with an electrolyte and probe pulses with durations less than 10 μs are employed, we find that the influence of the electrochemical processes can be neglected in most cases. Under such conditions the high-frequency impedance of the interface is determined in the absence of a dielectric layer on the surface by the capacitance of the space-charge layer of the semiconductor C_{sc} and the impedances of the surface states connected to it in parallel, each of which consists of a capacitance C_{ss} and a trapping resistance R_{ss} connected in series. In measurements with selected durations of the cycling pulses and of the time intervals between them, usually only one surface state, whose time constant $\tau_{ss}=C_{ss}R_{ss}$ is commensurate with the duration of the probe current pulses, participates in space-charge relaxation. The semiconductor structure under investigation can be represented in the form of the equivalent circuit shown in Fig. 1. We shall assume that the resistances of the electrolyte and of the bulk of the semiconductor are negligibly small. We use C_0 to denote the combined capacitance of the cables, the ballast capacitance, and the high-frequency surface states. When a constant current pulse flows in the circuit, the equation describing the variation of the surface potential ΔV_s is written in the following form:

$$R_{ss}C_h\frac{d\Delta V_s}{dt}(t)+\Delta V_s(t)\left(1+\frac{C_h}{C_{ss}}\right)=j_0\frac{R_{ss}C_{ss}+t}{C_{ss}}, \qquad (1)$$

where t is the time measured from the onset of the pulse, j_0 is the amplitude of the probe current pulse, and $C_h=C_0+C_{sc}$.

After solving Eq. (1) with the boundary condition $\Delta V_s(0)=0$, we obtain the expression for $\Delta V_s(t)$ during charging of the structure

$$\Delta V_s(t)=\frac{j_0}{C_h+C_{ss}}(t+t_0(1-\exp(-t/\tau))). \qquad (2)$$

Here $t_0=\tau_{ss}(1+C_h/C_{ss})$, and $\tau=\tau_{ss}/(1+C_{ss}/C_h)$. After completion of the current pulse, redistribution of the charge within the structure takes place during the time t_p. In this case the total current $j=0$ and, therefore, in the interval between the charging and discharging pulses

$$j_1=-j_2=-C_h\frac{d\Delta V_s}{dt},$$

$$\Delta V_s(t)=-R_{ss}C_{ss}\frac{d\Delta V_s}{dt}+\frac{j_0t_p-C_h\Delta V_s}{C_{ss}}. \qquad (3)$$

The solution of this equation obtained with matching to the previously obtained result (2) at the point $t=t_p$ is written as

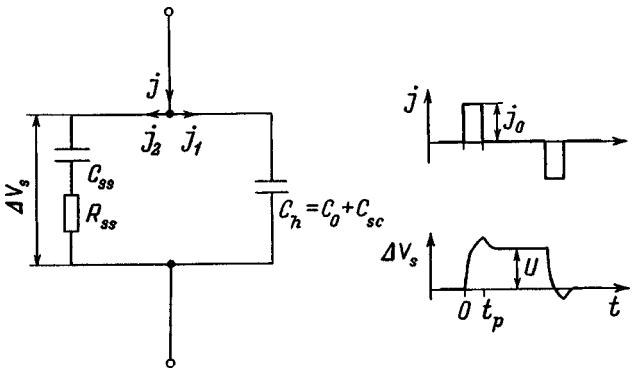

FIG. 1. Equivalent circuit of an electrolyte-semiconductor interface.

$$\Delta V_s(t)=\frac{j_0 t_p}{C_p+C_{ss}}+\frac{j_0 t_0}{C_h+C_{ss}}(1-\exp(-t_p/\tau))\times\exp(-(t-t_p)/\tau). \quad (4)$$

If the expression (4) is explored at $t-t_p\gg\tau$, $\Delta V_s(t)$ reaches its limiting value U and ceases to depend on time (Fig. 1), permitting determination of the total capacitance of the system

$$C=C_h+C_{ss}=C_0+(C_{sc}+C_{ss})=j_0 t_p/U. \quad (5)$$

After a current pulse of the same duration and amplitude, but of opposite polarity, is supplied to the structure, the system returns to its original state.

In order to find the capacitance $C_{sc}+C_{ss}$ to a high accuracy from the measured value of U, we propose performing additional calibration measurements with replacement of the structures under investigation by standard capacitors. We can use a set of capacitors whose known ratings cover the interesting range of variation of the capacitance of the structure under investigation to construct a calibration plot of $U((C+C_0)^{-1})$ with linear interpolation of the function on the segments between points and extrapolation outside the calibrated interval. Utilization of the curve obtained permits the performance of measurements over a broader range of values in comparison to the ordinary method[4] and takes into account the presence of the parasitic and ballast capacitances C_0. In addition, the use of a calibration curve makes it possible to compensate for the deviation of $U((C+C_0)^{-1})$ from a linear proportionality to a considerable extent. The latter deviation appears both because of the nonlinearity of the amplifying circuit and because of the penetration of synchronous pulsed disturbances from the commutating digital circuits. For the reasons cited, the latter leads to displacement of the $U((C+C_0)^{-1})$ curve along the U axis by a certain constant amount U_∞, which corresponds to the measured pulsed voltage on the structure when $C=\infty$. At large capacitances the magnitude of U_∞ is always commensurate with or even exceeds the net increment of U, which is shown in Fig. 1 for an ideal system.

Let us consider the influence of the presence of the capacitance C_0 on the accuracy of the capacitance-voltage measurements in greater detail. The presence of the $R_{ss}C_{ss}$ circuit increases the error in the value of C_{sc} obtained, since part of the charge of the space-charge layer participates in charge exchange with the surface states with the time constant $\tau=R_{ss}/(1/C_{ss}+1/(C_0/C_{sc}))$. It can be seen that τ increases with the capacitance of the ballast capacitor. Thus, by increasing C_0 we can slow the error-causing relaxation processes. However, as follows from formula (5), this effect will be accompanied by a decrease in the amplitude of the measured signal at fixed values of j_0 and t_p, which restricts the maximum attainable value of C_0 associated with the finite number of data bits in the analog-to-digital converter (ADC) used to input the signal. We note that increasing j_0 at high frequencies is associated with considerable difficulties and can compensate the decrease in the signal with increasing C_0 only to a certain limit.

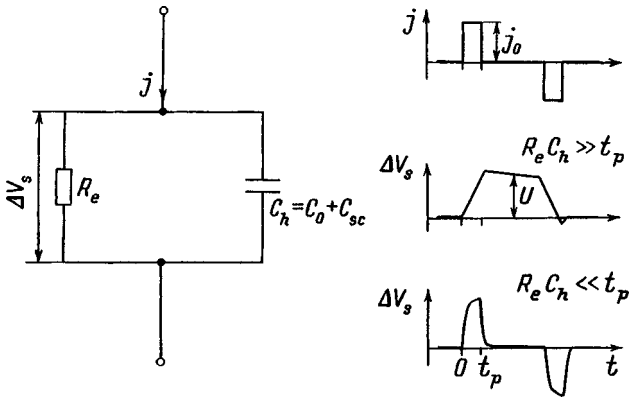

FIG. 2. Equivalent circuit of ES, EIS, and MIS structures with leakage currents in the absence of surface states.

When there are no surface states or their influence is small, the main errors in the measuring process will be associated with the presence of leakage currents. In this case a leakage resistance R_e connected in parallel to the capacitance $C_h=C_{sc}+C_0$ must be added to the equivalent circuit (Fig. 2). Its influence on the measured value of C_{sc} is more significant, the smaller is C_0. This is attributed to the fundamental difference between the charging-discharging process of a system with $R_eC_h>t_p$ [for which the function $U((C+C_0)^{-1})$ presented above remains linear] and a system with $R_eC_h\ll t_h$. The influence of the leakage currents can be reduced with consideration of the constraints indicated above by increasing the rating of the ballast capacitor. Another advantage of employing a ballast capacitor is expansion of the dynamic range of the measured capacitances in comparison to the conventional method.[4] The system used in Ref. 5 employs a 12-bit ADC, which provides for digitizing over three and a half orders of magnitude with respect to the voltage, while the capacitance of the structure can vary over six orders of magnitude. Since the measured value of U is inversely proportional to $C=C_0+C_{sc}+C_{ss}$, at fixed values of j_0 and t_p it is possible to select the rating of C_0 so that the maximum value of the voltage will not exceed the digitizing range and will satisfy the small-signal condition $U\ll kT/q$. In this case the unavoidable loss of accuracy due to an increase in the noise level relative to the net signal can be partially compensated by multiple repetition of the measurement cycle (from 100 to 1000 ADC measurements per point) followed by averaging. Such a method significantly lowers the noise level and raises the accuracy by at least one order of magnitude.

Thus, without calibration and the use of a ballast capacitor, measurements of $C<100$ pF are essentially impossible for the reasons cited. In the system described in Ref. 5 the charging and discharging pulse duration used equals 5 μs, the relaxation delay time is 5 μs, the interval between pulses is 20 μs, and the pulse repetition frequency is 2 kHz (500 μs). The optimal rating of the ballast capacitor for these parameters is 1 nF, which provides for the measurement of surface capacitances in the range from 10^{-12} to 10^{-7} F.

[1] M. J. Chudobiak, **66**, 3703 (1995).
[2] M. B. Moore, in *MICRO'95*, IREE Society, Milsons Point, New South Wales (1995), pp. 186–188.
[3] I. V. Shashkin, I. R. Karetnikova, A. V. Myurel' *et al.*, Fiz. Tekh. Poluprovodn. **31**, 926 (1997) [Semiconductors **31**, 789 (1997)].
[4] O. V. Romanov, M. A. Sokolov, and S. N. Sultanmagomedov, Élektrokhimiya **16**, 935 (1980).
[5] V. V. Monakhov and A. B. Utkin, Vestn. St.-Peterb. Univ., Ser. 4: Fiz. Khim. **3**(18), 84 (1997).

Translated by P. Shelnitz

Influence of a forming process and electric fields on the insulator-conductor transition in thin films of polyheteroarylenes

T. G. Zagurenko, V. M. Kornilov and A. N. Lachinov

Institute of the Physics of Molecules and Crystals, Russian Academy of Sciences, 450075 Ufa, Russia
(Submitted May 19, 1998)
Zh. Tekh. Fiz. **69**, 85–87 (March 1999)

The results of experiments on initiation of the high-conductivity state of a metal/polymer/metal system by varying the boundary conditions with allowance for the possible influence of diffusion of the electrode material into the polymer film are presented. It is established that forming does not occur in poly(phthalidylidenebiphenylylene) and that the transition to the high-conductivity state can occur without an external source of electric voltage. © *1999 American Institute of Physics.* [S1063-7842(99)01803-6]

The anomalously high conductivity in undoped polymers is generally observed in thin films. For this reason, great significance is attached to the conditions under which high conductivity appears, since the choice of the model for explaining this phenomenon depends on it. In particular, in Ref. 1 such a state was obtained in polyimide films of thickness 12 μm by ''soft breakdown,'' which was regarded as a forming process. Traces of ''breakdown'' in the form of microscopic holes were discovered visually. Experiments showed that one result of such forming is a reversible switching effect when the pressure is raised and lowered. It was confirmed in Ref. 1 that the application of pressure to a polymer film in the absence of an applied electric voltage does not lead to the appearance of the ''pressure sensor effect''[3] in this polymer. It was postulated that ''soft breakdown'' is accompanied by the formation of compounds of carbon with the metal, which constitute a high-conductivity phase near the holes.

In Ref. 4 the transition to the high-conductivity state was effected by another method (by varying the boundary conditions on the metal/polymer interfaces) and on another polymer, viz., a polyheteroarylene. This method calls for the employment of a metal which changes its state of aggregation during the measurements as one of the electrodes, since such changes can create favorable conditions for soft-breakdown forming of the sample.[2] In this context the purpose of the present work was to investigate the influence of diffusion of the electrode material into the polymer film and the need for the presence of an electric field for the transition to the high-conductivity state when the latter is initiated by varying the boundary conditions. The plan of the experiments was as follows.

1) The role of diffusion of the electrode material in the bulk of the polymer was investigated in experiments with a molten electrode. In this case the diffusion process should be facilitated by several factors: the state of aggregation of one of the electrodes, the applied electric field with an intensity up to 10^6 V/cm, and the relatively large current flowing through the polymer film when the system passes into the high-conductivity state. According to the hypothesis under discussion, a metallopolymer phase, which completely specifies the electrophysical properties of the system, particularly the temperature for the transition to the high-conductivity state, should form during the first few cycles of transitions between the low- and high-conductivity states. For this reason, we performed measurements of the conductivity of a metal/polymer/metal system in heating-cooling cycles with successive replacement of the low-melting electrodes (Wood's alloy or indium) by aluminum or copper electrodes on the same polymer film. It was theorized that if diffusion of the electrode material plays a significant role in the transition of the system from the low-conductivity state to the high-conductivity state, the plot of the temperature dependence of the current flowing through the sample will have similar features at the transition temperatures after replacement of the electrodes.

2) The influence of the electric field on initiation of the high-conductivity state in the polymer film was investigated in heating-cooling cycles in a low-melting-metal/polymer/metal system without using an electric field to probe the conductive state. In this case the transition to the high-conductivity state can be detected by observing the disappearance of the noise at the input to the electrometric voltmeter in analogy to the method used in Ref. 4.

The object of investigation was poly(phthalidylidenebiphenylylene),[5] in which phenomena associated with the generation of a high-conductivity state were previously observed.[6] Uniform films with thicknesses from 1 to 5 μm were obtained by centrifuging a solution of the polymer in cyclohexanone. The experimental cell was a metal/polymer/metal sandwich. The lower electrode (the electrode on which the polymer film was poured) was composed of copper or vanadium. The upper electrode was composed of the following metals: Wood's alloy, indium, copper, or aluminum. The method used to measure the current in the experimental circuit was similar to the one previously used in Ref. 4. The difference pertains to the measurements performed without the use of an external current source. In that case the VK2-16 electrometer was connected directly to the electrodes of the measuring cell. The measuring cell was placed in a heating

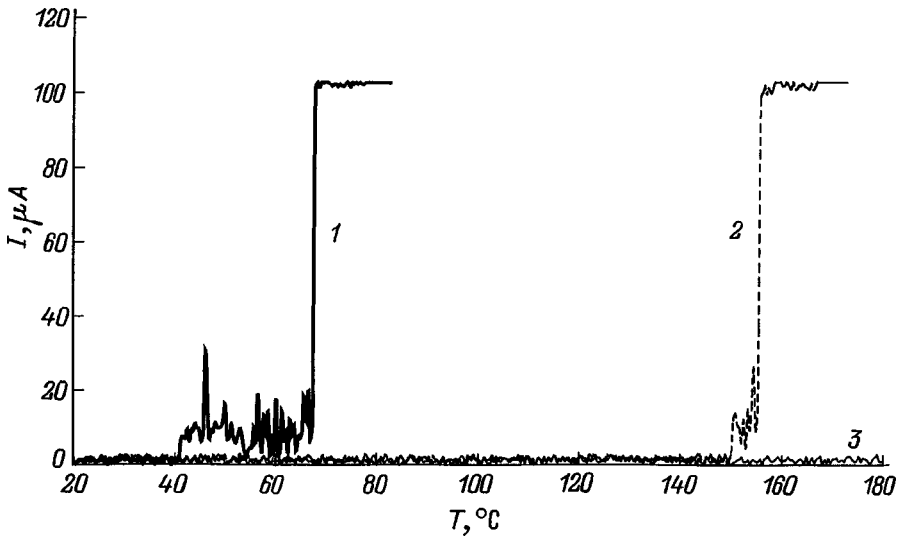

FIG. 1. Temperature dependence of the current through a poly(phthalidylidenebiphenylylene) film: *1* — upper electrode composed of Wood's alloy, *2* — upper electrode composed of indium, *3* — upper electrode composed of copper. The film thickness was 2 μm, the lower electrode was composed of vanadium, the applied voltage was 5 V, and the heating rate was 8 deg/min.

device, which allowed us to heat the cell to 250 °C with a controlled rate of variation of the temperature.

Figure 1 presents typical plots of the temperature dependence of the current through a film of thickness $\sim 2\,\mu$m with the successive use of Wood's alloy, indium, copper, or aluminum as electrodes. It should be noted that similar results were obtained for films of all the thicknesses indicated; therefore, here we present the most typical results for a thickness of 2 μm. At room temperature the polymer film has a conductivity equal to $\sim 10^{-14}\ (\Omega \cdot \mathrm{cm})^{-1}$, and the value of the current flowing through the sample is essentially equal to the instrumental zero.

In the temperature range corresponding to the premelting of Wood's alloy (40−45 °C), a sharp increase in the current fluctuations in the measuring circuit was observed. Achievement of the melting point of Wood's alloy led to an abrupt decrease in the resistance of the sample to 0.1−5 Ω, attesting to a transition of the polymer film to the high-conductivity state (curve *1* in Fig. 1).

Similar laws were observed in the course of the $I(T)$ curve in the case where an indium electrode was used with the exception of the temperature range, which was shifted into the region of the melting point of indium. In this case current fluctuations began to be observed near 140 °C, and the transition to the high-conductivity state took place at the melting point, i.e., at 156 °C (curve *2* in Fig. 1).

When the sample was cooled, an increase in the current fluctuations was observed near the crystallization temperature. A transition to the original low-conductivity state occurred below the crystallization temperature with a delay at 1−5 °C. This hysteresis can probably be attributed to the methodical features of the experiment.

At least 10 measurement cycles were performed. No differences in the character of the transitions from the low-conductivity to the high-conductivity state as a function of the number of cycles were discovered. Thus, after the measurements with the low-melting metals were performed, the total number of measurement cycles at the same site on the polymer film was no less than 20, apparently attesting to the weak influence of diffusion of the metals into the polymer on the shaping of the high-conductivity state.

The subsequent use of copper or aluminum electrodes, which do not melt in the experimental temperature range selected, as the probe electrodes led to the complete disappearance of any features in the flow of current over the entire temperature range measured (curve *3* in Fig. 1).

Let us consider the results of the second group of experiments. Figure 2 presents the temperature dependence of the amplitude of the electrical signal fed into electrometer, which is connected directly to the polymer sample. The same metals, viz., Wood's alloy and indium, were used in this experiment.

A potential difference U, which is probably caused by the space-charge field and the contact potential difference, was detected on the electrodes at room temperature. The value of U increased with increasing temperature up to a temperature close to the melting point of the electrode. When it was achieved, the potential difference dropped abruptly to the instrumental zero. In order to be convinced of the correct interpretation of the variation of the recorded signal, control measurements of the conductance of the sample were performed in several cases at a temperature exceeding the melting point of the electrode. All the control measurements showed that the samples were in the high-conductivity state. Thus, the following facts have been established.

1. The forming phenomenon characteristic of other polymers[1] does not occur in thin films of poly(phthalidylidenebiphenylylene), i.e., the characteristics of the first measurement cycle on a freshly prepared film do not differ fundamentally from those of subsequent cycles.

2. No influence of modification of the polymer as a result of the diffusion of metal into its bulk on the characteristics of the insulator-conductor transition was discovered.

3. The transition to the high-conductivity state can occur without a source of electric voltage in a circuit supplying power to the polymer sample. The space-charge field of the polymer, whose nature is still unclear, is sufficient for this transition.

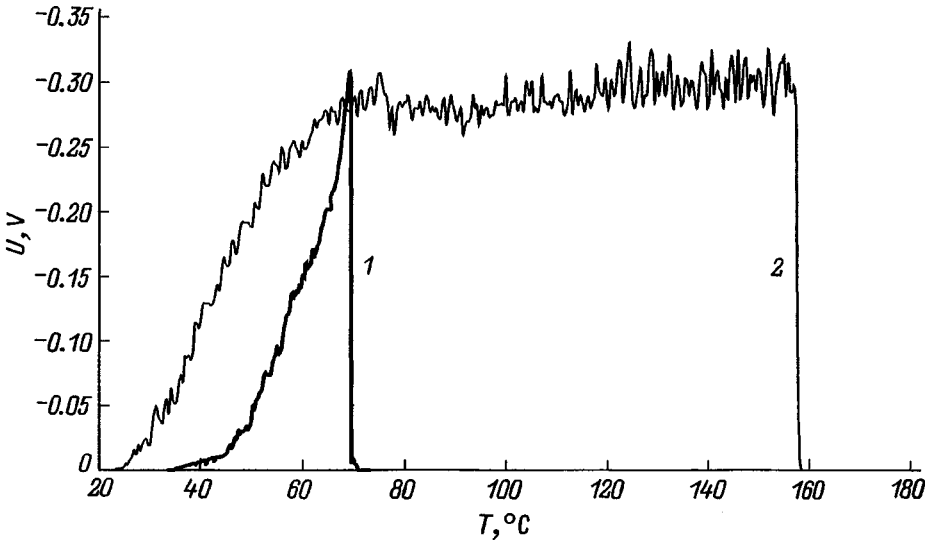

FIG. 2. Temperature dependence of the contact potential difference between the electrodes U: *1* — upper electrode composed of Wood's alloy, *2* — upper electrode composed of indium. The film thickness was 2 μm, the lower electrode was composed of vanadium, and the heating rate was 8 deg/min.

The disparity between the results obtained and the results previously published in Ref. 1 is attributable to the fact that a thickness of about $1-5$ μm is critical for polymer films, i.e., the transition to the high-conductivity state occurs according to an electronic mechanism in samples whose thickness is less than the values indicated, whereas it takes place according to an electronic-thermal mechanism in samples with larger thicknesses. Similar thickness constraints were previously noted in such objects as amorphous oxides,[7] chalcogenide glasses,[8] and polymers.[9] As a rule, an increase in sample thickness has a significant influence on the forming mechanism and leads to definite degradation of a film sample.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-19208).

[1] A. M. El'yashevich, A. N. Ionov, V. M. Tuchkevich *et al.*, Pis'ma Zh. Tekh. Fiz. **23**(14), 8 (1997) [Tech. Phys. Lett. **23**, 538 (1997)].
[2] A. M. El'yashevich, A. N. Ionov, M. M. Rivkin *et al.*, Fiz. Tverd. Tela (Leningrad) **34**, 3457 (1992) [Sov. Phys. Solid State **34**, 1850 (1992)].
[3] A. M. El'yashevich, A. N. Ionov, and V. V. Kudryavtsev, Vysokomol. Soedin. **35**, 50 (1993).
[4] V. M. Kornilov and A. N. Lachinov, Zh. Éksp. Teor. Fiz. **111**, 1513 (1997) [JETP **84**, 833 (1997)].
[5] M. G. Zolotukhin, V. A. Kovardakov, S. N. Salazkin *et al.*, Vysokomol. Soedin., Ser. A **26**, 6 (1984).
[6] A. N. Lachinov, A. Yu. Zherebov, and V. M. Kornilov, JETP Lett. **52**, 103 (1990).
[7] G. Dearnaley, A. M. Stoneham, and D. V. Morgan, Rep. Prog. Phys. **33**, 1129 (1970) [Usp. Fiz. Nauk **112**, 83 (1974)].
[8] *Electronic Phenomena in Chalcogenide Glassy Semiconductors* [in Russian], K. D. Tsendin (Ed.), Nauka, St. Petersburg (1996), 486 pp.
[9] L. F. Pender and R. J. Fleming, J. Appl. Phys. **46**, 3426 (1975).

Translated by P. Shelnitz