Reconstruction of dynamic systems as applied to secure communications

V. S. Anishchenko, A. N. Pavlov, and N. B. Yanson

Saratov State University, 410026 Saratov, Russia
(Submitted June 16, 1997)
Zh. Tekh. Fiz. **68**, 1–8 (December 1998)

An analysis is made of a method of reconstructing signals which parametrically modulate a chaotic generator using a one-dimensional realization of its oscillation process. Test examples are used to demonstrate the efficiency of this method for simultaneous independent transmission of several information signals over a single communication channel. © *1998 American Institute of Physics.* [S1063-7842(98)00112-3]

INTRODUCTION

One line of research in modern nonlinear dynamics is the reconstruction of dynamic systems. Interest in this problem over the last seventeen years was stimulated by the appearance of Refs. 1 and 2. Packard *et al.*[1] showed that the phase portrait of the attractor of a dynamic system can be reconstructed from the scalar time series $a_i = a(i\Delta t)$ if data from the same series a_i, taken after some delay, are used as the missing coordinates of the state vector. In Ref. 2 the possibility of reconstructing the phase portrait of an attractor using a one-dimensional time series was given theoretical foundation in the form of the Takens theorem.

The appearance of Ref. 2 was the precursor to using this theorem to predict system behavior,[3,4] and to calculate the metric[5] and dynamic[6] characteristics of an attractor using a time series. In a study published in 1987, Cremers and Hübler[7] proposed a method of reconstructing the equation for a dynamic system using its one-dimensional realization. The idea of this method (global reconstruction method) has been developed in a wide range of studies (for instance, Refs. 8–10).

Despite some differences in the global reconstruction algorithms used by different researchers, they all propose to solve the modeling problem in two stages. The first involves calculation of the embedding space dimension n. After the value of n has been selected, the phase portrait of the dynamic system is reconstructed using the scalar time series. In addition to the delay method,[1,2] any of the methods proposed in Ref. 11 can be used for this purpose. At the second stage of the algorithm, the general form of the mathematical model is defined and the evolution equations specified.

When series differentiation of the initial time series is used to reconstruct the phase portrait,[7–11] the mathematical model has the form

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} = x_3, \ldots,$$

$$\frac{dx_n}{dt} = f(x_1, x_2, \ldots, x_n, \boldsymbol{\mu}), \tag{1}$$

where f is the nonlinear function which needs to be determined and $\boldsymbol{\mu}$ is the vector of the parameters.

This method of reconstruction envisages that the general form of the function $f(x_1, x_2, \ldots, x_n, \boldsymbol{\mu})$ must be defined *a priori*. In general, even an approximate form of this function is impossible to determine *a priori* and thus, this function is usually represented as an expansion in terms of a certain basis and the expansion coefficients determined numerically. In addition to the standard basis $1, x_1, x_1^2, x_1x_2, \ldots$, the nonlinearity can also be approximated by an expansion in terms of any set of orthogonal basis functions (such as Legendre polynomials). In any case, however, the need to define the form of the function f is a serious disadvantage of this method of reconstructing the dynamic model of a system.

A fundamentally different situation arises when the explicit form of the nonlinear function f is known *a priori* and only the expansion coefficients are unknowns. In this case, the global modeling problem can be solved to a certain accuracy, which is determined by the number of points in the initial time series available for observation, the smallness of the discretization step, and the accuracy of writing the data. In Ref. 12 we showed that a consequence of solving the problem of reconstructing the mathematical model may be interesting applications, one of which is the use of the reconstruction technique to obtain secure communications.

It should be noted that the appearance of wide range of applied problems in modern nonlinear dynamics (including secure communications) in recent years has been stimulated by the development of concepts of dynamic chaos, especially the phenomenon of synchronized chaos. This effect was used by the authors of pioneering work on the confidentiality of transmitted information, using wide-band oscillations of a chaotic generator as a masking[13–15] or carrier signal.[16–18] The first case was based on the principle of the synchronization of chaotic systems proposed in Ref. 19. An alternative method of securing transmitted information involved the experimental control of chaos.[20,21]

Here we propose a new method of solving the problem of secure communications based on the global reconstruction of a dynamic system. Whereas in Ref. 12 this method was only illustrated for a single example, that is a modified generator with inertial nonlinearity,[22,23] we shall now present

results of its application using the Lorenz and Rössler models. We shall also report results obtained for this type of generator which were not completely reflected in our previous study.

METHOD

We shall analyze a certain generator of dynamic chaos whose mathematical model is known

$$\frac{d\mathbf{x}}{dt}=\mathbf{F}(\mathbf{x},\boldsymbol{\mu}^0),\quad \mathbf{x}\in R^n,\quad \boldsymbol{\mu}^0\in R^m, \tag{2}$$

where $\mathbf{x}$ is the vector of state, $\mathbf{F}$ is the vector of the right-hand sides of the model system, and $\boldsymbol{\mu}^0$ is the vector of the constant values of the parameters.

We shall implement relatively slow modulation of an arbitrary number of parameters of the information signals $\mu_i(t)$, i.e., we shall introduce the variables

$$\mu_i^*=\mu_i^0+\mu_i(t), \tag{3}$$

which allows several communications to be transmitted simultaneously. In this case, the signal transmitted by the communication channel, comprising a one-dimensional realization of the oscillatory process of the chaotic generator, is generated by a nonautonomous dynamic system

$$\begin{aligned}&\frac{d\mathbf{x}}{dt}=\mathbf{F}(\mathbf{x},\boldsymbol{\mu}^0+\boldsymbol{\mu}(t)),\\ &\boldsymbol{\mu}^0=(\mu_1^0,\mu_2^0,\ldots,\mu_m^0),\\ &\boldsymbol{\mu}(t)=(\mu_1(t),\mu_2(t),\ldots,\mu_m(t)).\end{aligned} \tag{4}$$

In order to solve the problem of synthesizing dynamic systems using an observable one-dimensional realization, the system (4) should be reduced to the form (1). This implies that by replacing the variables, we need to transform the right-hand sides such that instead of the vector function $\mathbf{F}$ there remains only a single scalar nonlinear function f, which may have a very complex form, such as

$$\begin{aligned}&f=\frac{P}{Q},\quad P(\mathbf{x})=\sum_{l_1,l_2,\ldots,l_n=0}^{\nu_1} C_{l_1,l_2,\ldots,l_n}\prod_{k=1}^{n} x_k^{l_k},\\ &\sum_{k=1}^{n} l_k\leqslant\nu_1,\\ &Q(\mathbf{x})=\sum_{l_1,l_2,\ldots,l_n=0}^{\nu_2} D_{l_1,l_2,\ldots,l_n}\prod_{k=1}^{n} x_k^{l_k},\\ &\sum_{k=1}^{n} l_k\leqslant\nu_2,\quad \nu_2<\nu_1,\end{aligned} \tag{5}$$

where $C_{l_1,l_2,\ldots,l_n}$, $D_{l_1,l_2,\ldots,l_n}$ generally depends on time and is uniquely related to the parameters μ_i^* of the system (4).

For the model systems considered here, i.e., Lorenz, Rössler, and a generator with inertial nonlinearity, these transformations will be made in the following section. Since it is assumed that the average rate of change in the parameters is small compared with the base frequency of the generator oscillations (2), i.e., $\overline{d\mu_i^*/dt}\ll\overline{dx_j/dt}$ for any i and j, we can introduce the time interval t_0 during which the values of the parameters can permissibly be assumed to be almost constant, i.e., for times of the order of t_0 we can neglect the nonautonomy of the system (4). This means that it is possible to reconstruct instantaneous values of the system parameters using short sections of its one-dimensional realization, i.e., to reproduce the information signals $\mu_i(t)$ which parametrically modulate the chaotic generator.

If the condition $\overline{d\mu_i^*/dt}\ll\overline{dx_j/dt}$ is not satisfied, we must take into account the time derivatives of the parameters when transforming the system (4) to the form (1), which makes the nonlinear function f substantially more complex and impedes the solution of the global reconstruction problem.

By applying the reconstruction technique to the one-dimensional realization $x_1(t)$ of the chaotic generator which can be measured at the exit of the transmitting device, the information receiver, knowing the general form of the mathematical model (2), isolates the useful signals $\mu_i(t)$. To do this, the receiver must differentiate the realization $x_1(t)$ n times, thus determining the left-hand sides of the model system (1). As a result, the problem of determining the values of the parameters at a given time reduces to the need to solve an algebraic equation with a certain number of unknowns (when f can be represented in the form (5), these unknowns are the coefficients $C_{l_1,l_2,\ldots,l_n}$ and $D_{l_1,l_2,\ldots,l_n}$, which are uniquely related to the parameters μ_i^* of the system (4)). Since a discretized time dependence $x_1(i\Delta t)$ rather than an analog signal is required for computer processing, the derivatives are clearly determined at discrete times $i\Delta t$ using approximate numerical differentiation formulas.

We can approximate the unknown coefficients by writing a system of K algebraic equations ($K=[t_0/\Delta t]$) for L unknowns ($L\ll K$) and solving this by the least squares method. Quite clearly, the approximation error decreases with decreasing L. Thus, in order to implement the proposed method of secure communication in practice, the model system must be simplified as far as possible. If the series of parameters of the system (4) remains unchanged, it is advisable to assume that these are predefined and approximate for fewer unknowns.

MODELS INVESTIGATED

We selected the Lorenz, Rössler, and generator with inertial nonlinearity models as chaotic oscillators.

a) *Lorenz system*. We shall analyze the equations from the well-known Lorenz model

$$\frac{dx}{dt}=\sigma(y-x),\quad \frac{dy}{dt}=rx-y-xz,\quad \frac{dz}{dt}=-bz+xy. \tag{6}$$

We shall assume that the x coordinate of the system (6) is selected as the carrier signal. It was shown in Ref. 8 that for this case, the transformation of (6) to the form (1) can give a simpler form of the function f than that for the other two coordinates of the Lorenz equations. As a result of this transformation, which is made by the change of variables

$$X=x, \quad Y=\sigma(y-x), \quad Z=\sigma[(r+\sigma)x-(\sigma+1)y-xz], \tag{7}$$

we obtain the system

$$\frac{dX}{dt}=Y, \quad \frac{dY}{dt}=Z, \ldots, \quad \frac{dZ}{dt}=f(X,Y,Z,\boldsymbol{\mu}),$$

$$\boldsymbol{\mu}=(\sigma,r,b), \tag{8}$$

$$f=b\sigma(r-1)X-b(\sigma+1)Y-(b+\sigma+1)Z$$

$$-X^2Y-\sigma X^3+\frac{Y[(\sigma+1)Y+Z]}{X}, \tag{9}$$

which generally contains three unknown parameters. Using the Lorenz model as an example we demonstrate that secure transmission of information can be achieved by modulating only one of its parameters, for example b. In this case, formula (9) is best rewritten as follows:

$$f+(\sigma+1)Z+X^2Y+\sigma X^3-\frac{Y[(\sigma+1)Y+Z]}{X}$$

$$=b[\sigma(r-1)X-(\sigma+1)Y-Z]. \tag{10}$$

Since σ and r are assumed to be known, and Y, Z, and f can be determined by numerical differentiation of the time series $x(i\Delta t)$ obtained by integrating the system (6), where Δt is the discretization step, which is taken to be 0.025 for all the models studied, Eq. (10) is merely a linear algebraic equation with a single unknown.

Theoretically, in order to determine the instantaneous value of b, we need to know the phase coordinates and their derivatives only at one moment of time. In practice, we need to analyze a short section of the scalar time series and approximate the value of b using the results of calculations at different times in the interval t_0 during which the value of the parameter can be considered to be almost constant in order to improve the accuracy of calculating the parameter.

b) *Generator with inertial nonlinearity*. A modified generator with inertial nonlinearity[22,23] was taken as the second model system

$$\frac{dx}{dt}=m_0x+y-xz, \quad \frac{dy}{dt}=-x,$$

$$\frac{dz}{dt}=g_0z+0.5g_0(x+|x|)x. \tag{11}$$

We shall assume that the signal emitted by the generator is a one-dimensional realization $y(t)$. The system (11) is transformed to the form (1) by changing the variables

$$Y=y, \quad Z=-x, \quad X=-m_0x-y+xz, \tag{12}$$

as a result of which the generator equations have the form

$$\frac{dY}{dt}=Z, \quad \frac{dZ}{dt}=X, \quad \frac{dX}{dt}=f(X,Y,Z,\boldsymbol{\mu}), \quad \boldsymbol{\mu}=(m_0,g_0), \tag{13}$$

$$f=\frac{X(X+Y)}{Z}+(m_0g_0-1)Z$$

$$-g_0(X+Y)+0.5g_0(|Z|-Z)Z^2. \tag{14}$$

The model of a generator with inertial nonlinearity was used to illustrate the possibility of simultaneously transmitting two independent information signals along a single communication channel. For this purpose the parameter m_0 was modulated by a wide-band chaotic signal obtained by integrating the Rössler equation and the parameter g_0 was modulated by the harmonic signal

$$\frac{dx}{dt}=(m_0+dx_1)x+y-xz, \quad \frac{dy}{dt}=-x,$$

$$\frac{dz}{dt}=-g_0(1+k_1\sin(\omega t))[z-0.5(x+|x|)x],$$

$$\frac{dx_1}{dt}=k(-y_1-z_1), \quad \frac{dy_1}{dt}=k(x_1+ay_1),$$

$$\frac{dz_1}{dt}=k(b+z_1(x_1-c)), \tag{15}$$

where k is a constant which renormalizes the time in the Rössler model such that the process of variation of the parameter $m^*=m_0+dx_1$ is slower than the oscillations of the generator.

A transformation of the equation for a generator with inertial nonlinearity to the form (13) assuming slowly varying parameters can give the function f in the form (14) provided that m_0 and g_0 are replaced by $m^*=m_0+dx_1$ and $g^*=g_0(1+k_1\sin(\omega t))$, respectively, i.e.,

$$f-\frac{X(X+Y)}{Z}+Z=m^*g^*Z-g^*[X+Y+0.5(|Z|-Z)Z^2]. \tag{16}$$

If we introduce the notation $s^*=m^*g^*$, Eq. (16) can be considered to be a linear algebraic equation with two unknowns s^* and g^*, which are found using a short scalar time series by the least squares method. The instantaneous value of the parameter m^* can then be uniquely determined from a knowledge of s^* and g^*.

c) Rössler system.

We take the y coordinates of the Rössler model as the carrier signal

$$\frac{dx}{dt}=-y-z, \quad \frac{dy}{dt}=x+ay, \quad \frac{dz}{dt}=b+z(x-c). \tag{17}$$

By analogy with the previous model (11), we transform (17) to the form (13) by changing the variables[8]

$$Y=y, \quad Z=x+ay, \quad X=ax+(a^2-1)y-z. \tag{18}$$

We then obtain the following form of the nonlinear function f:

$$f=-b+(a-c)X-cY+(ac-1)Z$$

$$-aY^2-aZ^2-aXY+XZ+(a^2+1)YZ. \tag{19}$$

We assume that the vector $\boldsymbol{\mu}=(a,b,c)$ contains two unknown parameters b and c which are modulated by the information signals. Equation (19) can then be rewritten as follows:

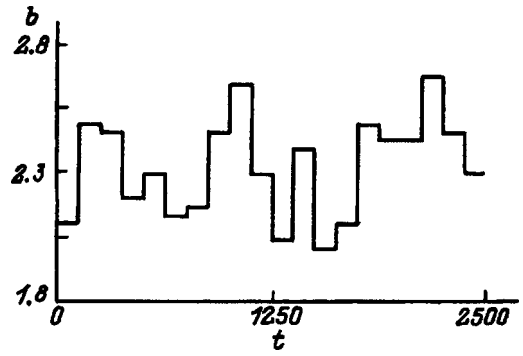


FIG. 1. Law of variation of the parameter b in a Lorenz system.

$$f-aX+Z+aY^2+aZ^2+aXY-XZ-(a^2+1)YZ = -b-c(X+Y-aZ). \quad (20)$$

The unknowns b and c of the linear algebraic equation (20) can again be obtained by applying the least squares method to the results of calculating the parameters at the times $i\Delta t$ of the discretized signal $y_i=y(i\Delta t)$ within the short time t_o. The Rössler model is used to illustrate the possibility of transmitting graphical information by modulating two of its parameters with the information signals.

RESULTS

Before giving some specific results which confirm the efficiency of the proposed method of secure communications, we shall make a few observations. A characteristic feature of the global reconstruction method is that it can be applied not only to steady-state signals but also to transient processes. Although the Takens theorem,[2] to which reference is usually made when reconstructing phase portraits, was demonstrated for the case where the signal is a one-dimensional projection of a phase trajectory assigned to the attractor of a dynamic system, the assignment of the phase trajectory to the attractor is not a necessary condition for the modeling problem. We have already observed that because of the relatively slow variation of the parameters, we can introduce the time interval t_0 during which it is permissible to assume that the values of μ_i^* in Eq. (3) are almost constant and the system (4) is autonomous. Since for the same values of the parameters the motion along the attractor and the transient processes are described by the same equations, the transience of the signal in the communications channel during the time t_0 is not of fundamental importance for determining the instantaneous values of μ_i^*.

We also note that any method of transferring information should be analyzed in terms of its performance under noise conditions of varying origin. Thus, in all the examples examined below, a normally distributed random quantity with a variance of 10^{-4} was added to the information signals performing the parametric modulation.

a) *Lorenz system.* We shall take the law of variation of the parameter b to be a stepped time dependence similar to that plotted in Fig. 1. This type of signal can be used to transmit graphical information. To illustrate this we shall scan the well-known Raphael painting of the Sistine Madonna (Fig. 2a) with 200×300 resolution. We shall divide the range of variation of the parameter $b\in[2-3]$ into 256 subranges, each corresponding to a shading gradation of the black and white image. The useful signal is a dependence such as that shown in Fig. 1, where each subrange of variation of the parameter corresponds to a particular step height. The carrier signal (the time dependence $x(t)$ of the Lorenz system), divided into the same number of subranges as the information signal, is shown in Fig. 2b. By applying the global reconstruction method, we isolated the modulation signal (Fig. 2c). Similar results were obtained by modulating the parameter r in system (6).

We note that having selected a stepped time dependence

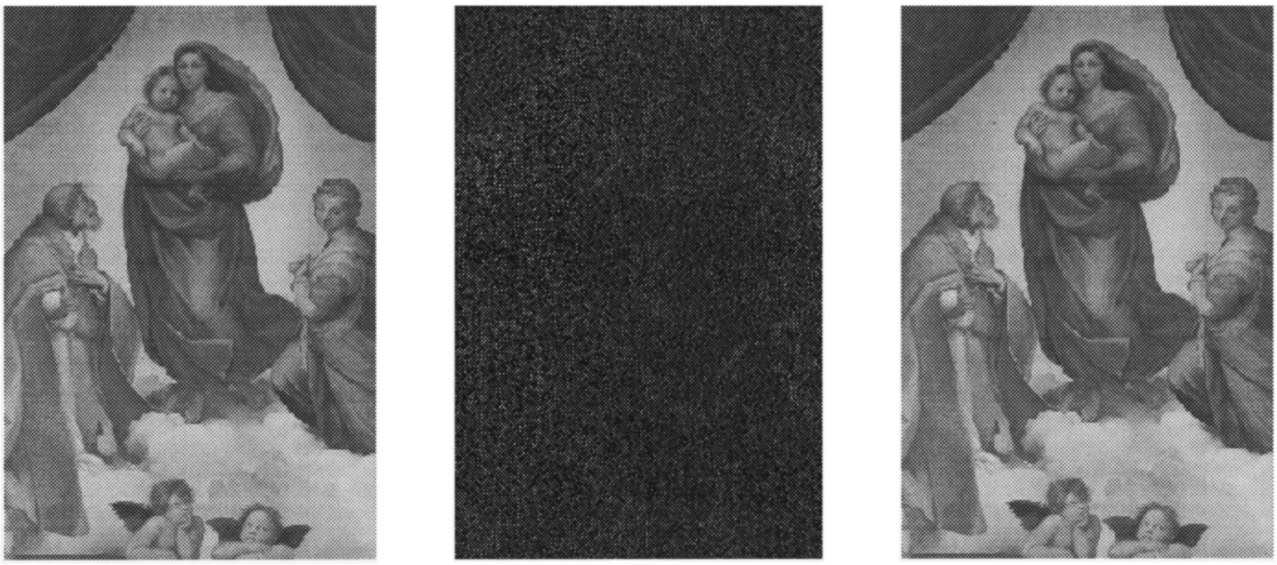


FIG. 2.

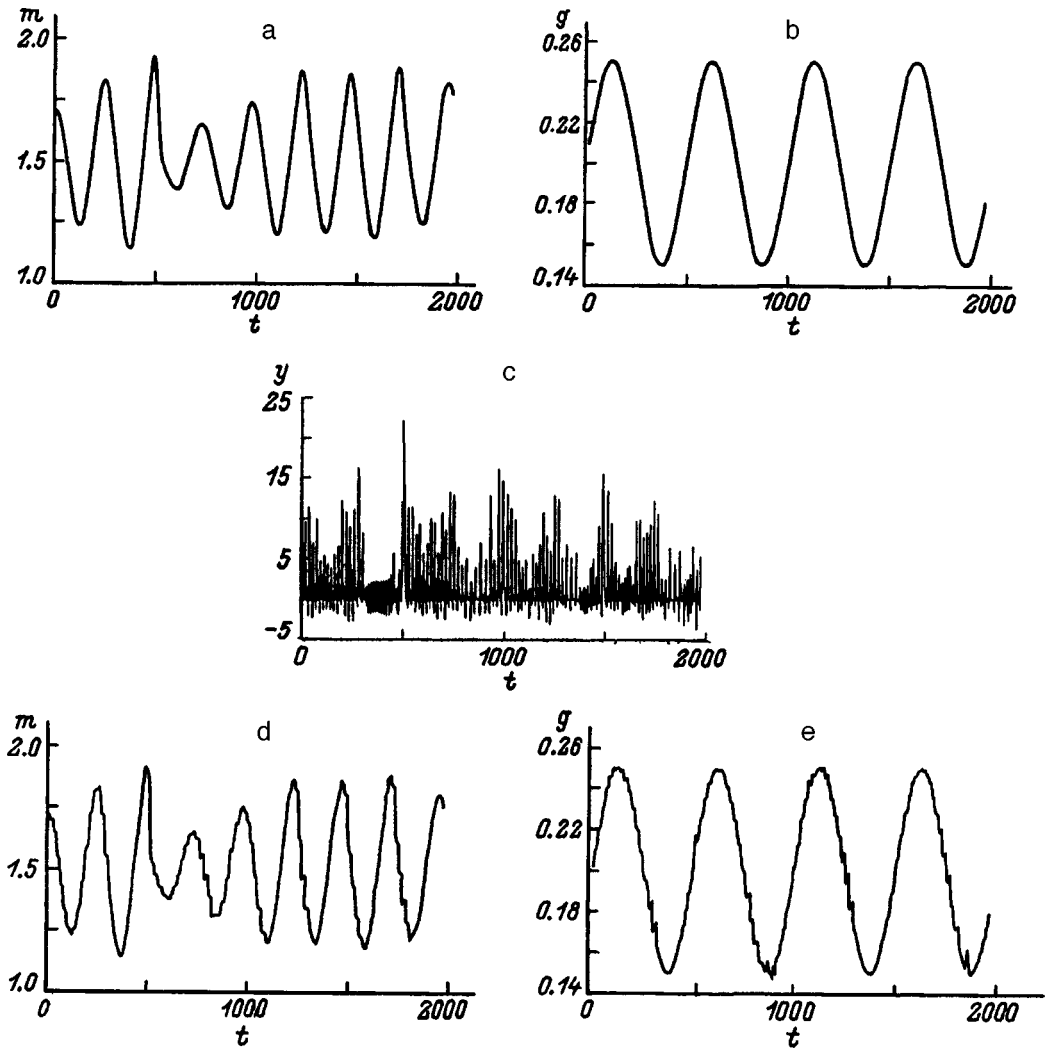


FIG. 3.

as the law of parametric modulation, we must recognize that we cannot use formula (10) for the switching time between subranges since it is abundantly clear that in the immediate vicinity of the switching time the system parameter cannot be considered to be slowly varying. Thus, the derivatives $d\mu_i/dt$ must be taken into account when transforming the system (4) to the form (1). This can be avoided by selecting fairly large widths for each step and disregarding the small sections near the switching times.

b) *Generator with inertial nonlinearity.* We shall demonstrate the possibility of transmitting two signals simultaneously: a chaotic and a regular one. We turn our attention to the system (15) in which we select the following values for the constants: $d=0.025$, $g_0=0.2$, $m_0=1.5$, $\omega=0.006$, $k_1=0.05$, $k=0.025$, $a=0.15$, $b=0.2$, and $c=10.0$. The time dependences of the parameters m^* and g^* of this system are illustrated in Figs. 3a and 3b. The information receiver, knowing the form of the nonlinear function (16), receives the signal (Fig. 3c) and reconstructs the modulation signals (Figs. 3d and 3e).

c) *Rössler system.* By analogy with the Lorenz model, we take a stepped time dependence as the parametric modulation law which allows us to transmit graphical information. However, two of its parameters, b and c, will now be modulated by the useful signals. We take two fragments of Leonardo da Vinci's painting ''Madonna of the Rocks'' (Figs. 4a and 4b) and scan them with 200×250 resolution. The range of variation of each parameter $b\in[0.1-0.3]$ and $c\in[8-12]$ is again divided into 256 subranges. Figure 4 shows the initial fragments of the painting (Figs. 4a and 4b) reconstructed using the global reconstruction technique (Figs. 4d and 4e), and also the signal in the communication channel (Fig. 4c).

A disadvantage of this method of transmitting a graphical image is that the information receiver must know the resolution which was used for scanning since the isolated signal is a single realization. This can be circumvented as follows. Let us assume that one of the parameters is modulated by a signal carrying information on the scanned graphical image. Simultaneous modulation of a second parameter can transmit information on the resolution. If we introduce the x and y coordinates of points on the transmitted fragment of the painting, one method of transmitting the graphical image is shown in Fig. 5. The number of steps (Fig. 5b) will carry information on the resolution along the y axis while the resolution along the x axis can be determined from the ratio

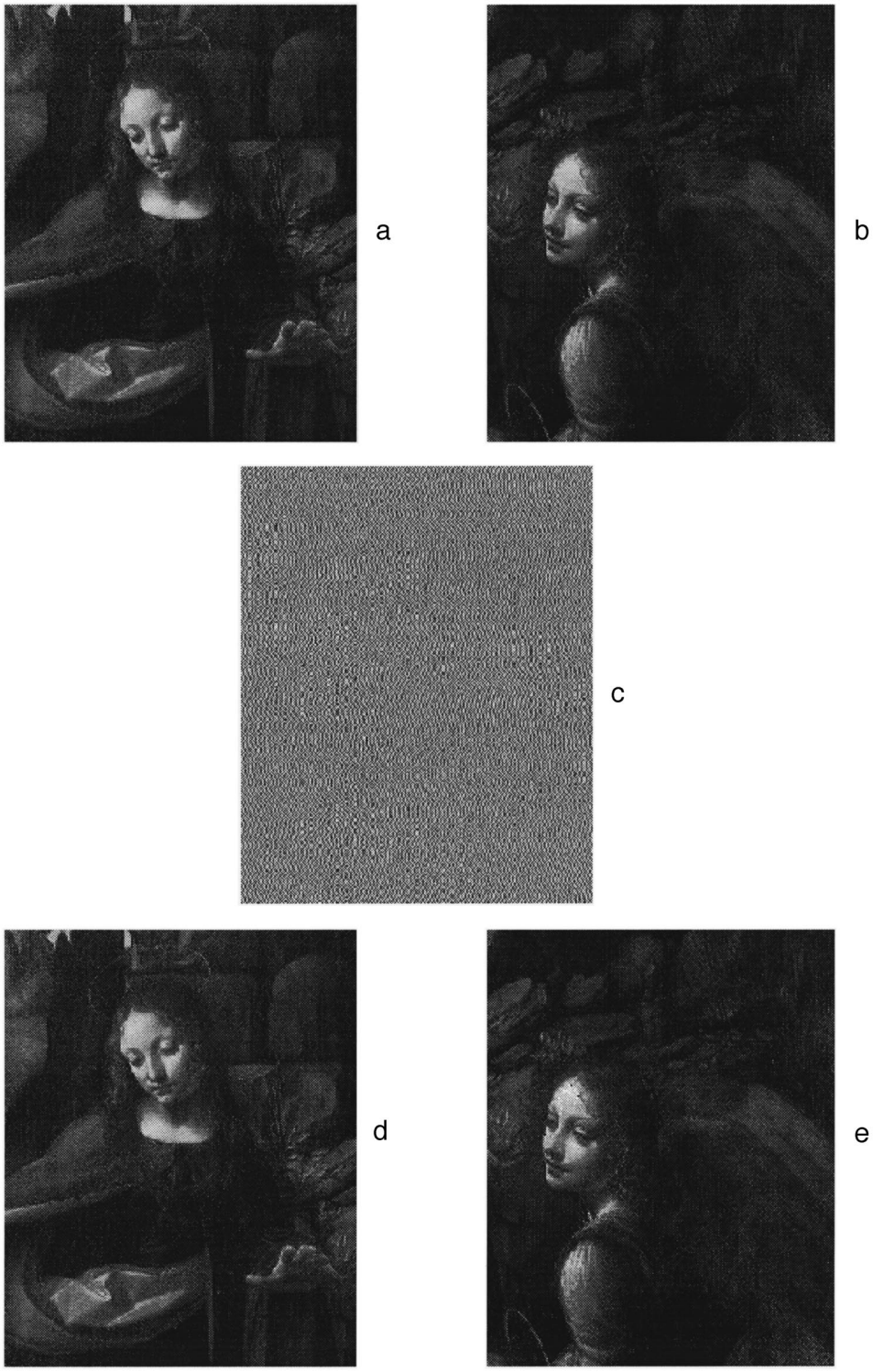


FIG. 4.

of the step width to the corresponding value for Fig. 5a.

The possibility of simultaneously modulating different parameters of a dynamic system by information signals means that the methods of transmitting a graphical image can be varied. In the simplest case when some drawing is transmitted using only two colors (black and white), i.e., the useful signal is a binary sequence of symbols (0 or 1), the rate of transmission of the information by modulating two parameters of the dynamic system can be enhanced by transmitting only the coordinates of points corresponding to one particular color. Quite clearly, other methods of transmitting graphical information can also be suggested.

CONCLUSIONS

Here we have illustrated a new method for secure communication based on the global reconstruction of dynamic

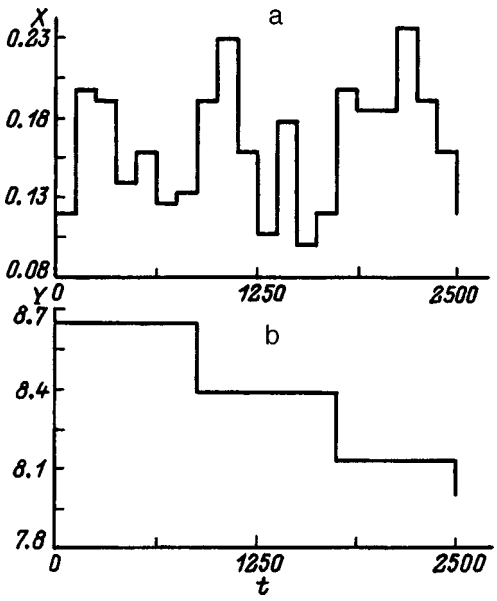


FIG. 5. One method of transmitting graphical information by simultaneously modulating two parameters.

systems. The selected model systems were Lorenz and Rössler systems and a generator with inertial nonlinearity. The results confirm the efficiency of the proposed method and also its reliability in the presence of noise considerably exceeding the thermal and shot noise of real electronic devices.

Theoretically, the proposed method imposes no constraints on the number of parameters which can be varied simultaneously. In practice such constraints arise from the finite accuracy of the calculations of μ_i^*. In particular, the instantaneous values of μ_i^* were not determined with a high degree of accuracy when three parameters were modulated simultaneously for the same discretization step ($\Delta t = 0.025$) and 256 subranges. Thus, when more than two parameters are modulated, fewer subranges must be used or the step Δt must be reduced, thereby increasing the number of carrier signal points contained in the range t_0. In this last case, the parameters of the numerical system must be optimized and these may differ for different model systems.

These constraints can be ascribed to the purely technical aspect of implementing the secure transmission of information, on which we decided not to focus attention in the present study, confining ourselves to demonstrating the fundamental possibility of simultaneously transmitting several information signals independently in a single communication channel using the global reconstruction technique.

This work was partly financed by the Russian State Committee on Higher Education (Grant No. 95-0-8.3-66).

[1] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. **45**, 712 (1980).
[2] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. A. Rang and L. S. Young, Lecture Notes in Mathematics, Vol. 898 (Springer-Verlag, Berlin, 1981), p. 366.
[3] J. D. Farmer and J. J. Sidorowich, Phys. Rev. Lett. **59**, 845 (1987).
[4] M. Casdagli, Physica D **35**, 335 (1989).
[5] P. Grassberger and J. Procaccia, Phys. Rev. Lett. **50**, 346 (1983).
[6] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Physica D **16**, 285 (1985).
[7] J. Cremers and A. Hübler, Z. Naturforsch., A: Phys. Sci. **42**, 797 (1987).
[8] G. Gouesbet and C. Letellier, Phys. Rev. E **49**, 4955 (1994).
[9] J. Kadtke and M. Kremliovsky, AIP Conf. Proc. **375**, 189 (1995).
[10] O. L. Anosov, O. Ya. Butkovskii, Yu. A. Kravtsov, and E. D. Surovyatkina, AIP Conf. Proc. **375**, 71 (1995).
[11] J. L. Breeden and N. H. Packard, Int. J. Bifurcation Chaos Appl. Sci. Eng. **4**, 311 (1994).
[12] V. S. Anishshenko and A. N. Pavlov, Phys. Rev. Lett. (in press).
[13] K. M. Coumo and A. V. Oppenheim, Phys. Rev. Lett. **71**, 65 (1993).
[14] K. M. Kocarev, K. S. Halle, K. Eckert *et al.*, Int. J. Bifurcation Chaos Appl. Sci. Eng. **2**, 709 (1992).
[15] C. W. Wu and L. O. Chua, Int. J. Bifurcation Chaos Appl. Sci. Eng. **3**, 1619 (1992).
[16] U. Parlitz, L. O. Chua, L. Kocarev *et al.*, Int. J. Bifurcation Chaos Appl. Sci. Eng. **2**, 973 (1992).
[17] K. M. Coumo, A. V. Oppenheim, and S. H. Strogatz, IEEE Trans. Circuits Syst. **40**, 626 (1993).
[18] H. Dedieu, M. P. Kennedy, and M. Hasler, IEEE Trans. Circuits Syst. **40**, 634 (1993).
[19] L. M. Pecora and T. L. Carroll, Phys. Rev. Lett. **64**, 821 (1990).
[20] S. Hayes, C. Grebogi, E. Ott, and A. Mark, Phys. Rev. Lett. **3**, 1781 (1993).
[21] H. D. I. Abarbanel and P. S. Linsay, IEEE Trans. Circuits Syst. **40**, 643 (1993).
[22] V. S. Anishchenko, *Complex Oscillations in Simple Systems* [in Russian], Nauka, Moscow (1990), 312 pp.
[23] V. S. Anishshenko, *Dynamical Chaos — Models and Experiments* (World Scientific, Singapore, 1995), p. 383.

Translated by R. M. Durham

Nonlocal hydrodynamics approach to describe a classical collisionless plasma

O. Yu. Dinariev

O. Yu. Shmidt Institute of Earth Physics, Russian Academy of Sciences, 123810 Moscow, Russia
(Submitted December 18, 1995)
Zh. Tekh. Fiz. **68**, 9–14 (December 1998)

By analogy with an approach developed earlier for a neutral gas, it is shown that for a certain class of external sources the linear kinetic theory of a nondegenerate collisionless plasma is equivalent to a hydrodynamic description in which the constitutive relations are nonlocal in space and time. The general algebraic properties of the nonlocal kernels were investigated. A scheme for calculation of their explicit expression in terms of the error function is indicated. An analysis is made of the dynamics of weak perturbations of the rest state and it is shown that the results of the nonlocal hydrodynamic model agree with the effects obtained using other theories. © *1998 American Institute of Physics.* [S1063-7842(98)00212-8]

INTRODUCTION

It is known that the kinetic description of a gas for a certain class of external influences is equivalent to the hydrodynamic description in which the constitutive relations are nonlocal in space and time.[1–4] The corresponding nonlocal kernels are calculated using the collision integral. The method of converting from the kinetic theory to nonlocal hydrodynamics can also be applied to plasma physics. As for a gas, for a certain class of external sources this conversion is accurate since the solution of the hydrodynamic problem can be used to reconstruct the solution of the initial rate equation. The case of a collisionless plasma is distinguished by the fact that the Fourier transforms of the nonlocal kernels can be calculated by using special functions. Thus, for a collisionless plasma the nonlocal hydrodynamic description is as constructive as it is possible to get. In most other cases, the theory can only establish certain properties of the kernels and in specific calculations approximate or model expressions must be used for these.

A system of units of measurement is used in which the velocity of light in vacuum c, Planck's constant $\hbar$, and the Boltzmann constant k are unity. The units of measurement of the electromagnetic quantities are determined in accordance with the Gauss approach. The Greek indices have values of 0, 1, 2, and 3, corresponding to some inertial system of measurement x^α (x^α is the time). The Latin indices a, b, and c have values of 1, 2, 3 and correspond to the spatial coordinates. The Latin indices A, B, C, and D have values of $0, \ldots, (k+3)$, where k is the number of plasma components. Unless stated otherwise, summation is performed over recurrent indices. For an arbitrary function $g=g(x^\alpha)$, $g(k)$ will denote its Fourier transform

$$g_F = g_F(k_\alpha) = \int \exp(-ik_\alpha x^\alpha) g(x^\alpha) dx^\alpha.$$

NONLOCAL HYDRODYNAMIC EQUATIONS FOR A COLLISIONLESS PLASMA

In the kinetic description, the state of the system is characterized by the single-particle distribution function $f = f(x^\alpha, v^b, r)$, where v^a is the particle velocity and r is a collective parameter corresponding to the internal degrees of freedom (for instance, the component number, rotational and vibrational motion, and so on). In the parameter space (v^a, r) we introduce the measure $d\zeta = dv^1 dv^2 dv^3 d\mu(r)$ which can be used to average microscopic quantities. For example, if the function of the microscopic parameters $W = W(v^a, r)$ exists, it is possible to calculate the macroscopic field

$$w = w(x^\alpha) = \langle W \rangle = \int W(v^a, r)\, f(x^\alpha, v^i, r)\, d\zeta.$$

Let us consider a K-component collisionless plasma. In this case, the parameter r has the composite form (i, r'), where i has values of $1, \ldots, K$, corresponding to the numbers of components and the parameter r' is related to the other internal degrees of freedom. Integration $d\mu(r)$ is broken down into summation over the parameter i and integration over a certain $d\zeta'(r')$. We further assume that the indices i and j have values of $1, \ldots, K$ and the indices I and J have values of $(3+i)$, $i = 1, \ldots, K$. If the indices i, j and I, J are used in the same formula, their values are related by $I = i+3$, $J = j+3$.

We shall assume that m_i and e_i are the mass and charge of a particle of the ith component and $U(r)$ is the potential energy of the possible particle states. We define the sets of functions $J_A(v^a, r)$, $\Sigma^\alpha(v^a, r)$, and $\Phi^\alpha(v^a, r)$ by the following formulas in which no summation is performed over i

$$J_0(v^a, r) = \frac{1}{2} m_i v^b v^b + U(r),$$

$$J_a(v^b, r) = m_i v^a, \quad J_J(v^b, r) = \delta_{ij},$$

$$\Sigma^0(v^a, r) = e_i m_i^{-1}, \quad \Sigma^a(v^b, r) = e_i m_i^{-1} v^a,$$

$$\Phi^0(v^a, r) = e_i, \quad \Phi^a(v^b, r) = e_i v^a, \quad r = (i, r').$$

The plasma dynamics are described by the Vlasov equation

$$\partial_0 f + v^a \partial_a f + \Psi^a D_a f = S \tag{1}$$

and the system of Maxwell equations

$$\varepsilon_{abc}\partial_b E_c = -\partial_0 B_a, \quad \partial_a B_a = 0,$$

$$\varepsilon_{abc}\partial_b B_c = 4\pi j^a + \partial_0 E_a, \quad \partial_a E_a = 4\pi j^0. \tag{2}$$

The notation $\partial = \partial/\partial x^\alpha$ and $D = \partial/\partial v^a$ is used in Eq. (1), $S = S(x^\alpha, v^b, r)$ is the source function which describes the interaction between the plasma and the surrounding medium, and Ψ^a is the Lorentz force which is determined using the vectors of the electric E_a and magnetic B_a fields

$$\Psi^a = \Sigma^0 E_a + \varepsilon_{abc}\Sigma^b B_c.$$

The Vlasov equation is conventionally written with a right-hand side of zero, implying a Cauchy problem formulation. Equation (1) which contains sources is more general and embraces the Cauchy problem as a particular case.

The electric four-current j^α is broken down into the sum of the external current j^α_{ex} and the induced current j^α

$$j^\alpha = j^\alpha_{\text{ex}} + j^\alpha_{\text{in}}, \quad j^\alpha_{\text{in}} = \langle \Phi^\alpha \rangle. \tag{3}$$

The condition of total current conservation is required for compatibility with the system of Maxwell equations (2)

$$\partial_\alpha j^\alpha = 0. \tag{4}$$

The condition (3) will be interpreted as a constraint on the external four-current j^α_{ex}. Equation (1) yields the hydrodynamic equations

$$\partial_\alpha Q^\alpha_A = s_A + E_a L^a_A + B_a M^a_A,$$

$$Q^0_A = Q^0_A(x^\beta) = \langle J_A \rangle, \quad Q^a_A = Q^a_A(x^\beta) = \langle v^a J_A \rangle,$$

$$s_A = s_A(x^\alpha) = \int J_A(v^b, r) S(x^\alpha, v^b, r)\, d\zeta,$$

$$L^a_A = \langle \Sigma^0 D_a J_A \rangle, \quad M^a_A = -\varepsilon_{abc}\langle \Sigma^b D_c J_A \rangle. \tag{5}$$

The values of L^a_1, M^a_1, and M^a_0 are obviously zero. In the absence of an electromagnetic field and sources, the rate equation (1) is satisfied in the class of equilibrium states

$$f_e = f_e(v^a, r) = \exp(F^A J_A), \tag{6}$$

$$F^0 = -\beta, \quad F^a = \beta u^a, \quad F^I = \beta \mu_i, \tag{7}$$

where β is the reciprocal temperature, u^a is the velocity of the medium, and $\mu_i = (\mu_{i0} - (1/2) m_i u^a u^a)$ is the dynamic chemical potential.

In order to ensure that the equilibrium state satisfies the Maxwell equations (2) in the absence of external currents, the additional constraint of electrical neutrality must be imposed on the parameters (7)

$$\langle \Phi^\alpha \rangle_e = 0. \tag{8}$$

The equilibrium rest state $f_0 = f_0(v^a, r)$ is characterized by the fact that the velocity of the medium u^a in Eqs. (6) and (7) is zero. Note that in the rest state we have

$$L^a_A = \langle \Sigma^0 D_a J_A \rangle_0 = 0, \quad M^a_A = -\varepsilon_{abc}\langle \Sigma^b D_c J_A \rangle_0 = 0. \tag{9}$$

The hydrodynamic equations (5) may be given a more conventional form if the hydrodynamic four-currents Q^α_A are redefined in terms of different quantities. We introduce the particle densities $n_i = Q^0_I$, the mass density $\rho = \Sigma m_i n_i$, the mass-averaged velocity $u^a = \rho^{-1} Q^0_a$, the diffusion fluxes $d^a_i = (Q^a_1 - n_i u^a)$, the kinetic energy of the medium $K_* = (1/2)\rho u^a u^a$, the internal energy of the medium $U_* = Q^0_0 - (1/2)\rho u^a u^a$, the stress tensor $p^{ab} = (\rho u^a u^b - Q^a_b)$, and the heat flux vector $q^a = Q^a_0 + p^{ab}u^b - (U_* + K_*)u^a$. We stress that the new definition of the velocity of the medium u^a is consistent with the old one for the equilibrium states (6) and (7). The system (5) can then be rewritten as:

$$\partial_0(U_* + K_*) + \partial_a(q^a - p^{ab}u^b + (U_* + K_*)u^a) = s_0 + E_a L^a_0,$$

$$\partial_0(\rho u^a) + \partial_a(-p^{ab} + \rho u^a u^b) = s_b + E_a L^a_b + B_b M^a_A,$$

$$\partial_0 n_i + \partial_a(d^a_i + n_i u^a) = s_1.$$

We shall now discuss the definition of the viscous stress tensor. For the equilibrium states (6) and (7) the stress tensor is reduced to the spherical form $p^{ab} = -p\,\delta^{ab}$. In the class of equilibrium states the pressure p can then be expressed as a function of the internal energy ε and the densities n_i. This functional relationship is used to extend the definition of pressure to nonequilibrium states. The viscous stress tensor is then calculated using the formula

$$\tau^{ab} = p^{ab} + p\,\delta^{ab}.$$

We shall analyze the dynamics of weak perturbations of the rest state. Let us assume that $g^\alpha_A = \Delta Q^\alpha_A$ are linear perturbations of the hydrodynamic four-currents. Using Eq. (9), the equations (5) are reduced to the form

$$\partial_\alpha g^\alpha_A = s_A. \tag{10}$$

In order to close the problem (2) and (10), we need to have constitutive relations, i.e., expressions for the three-currents g^a_A (or, which amounts to the same thing, expressions for the viscous stresses τ^{ab} and heat flux q^a), and also an expression for the induced four-current j^α_{in} in terms of the components g^0_A and the electromagnetic field. These expressions can be obtained by investigating the kinetic problem (1).

Let us assume that the equilibrium rest distribution f_0 is perturbed by weak sources S. We shall take the usual representation for a perturbed distribution $f = f_0(1 + \varphi)$. Linearizing Eq. (1), we obtain

$$(\partial_0 + v^a \partial_a)\varphi = \beta E_a \Phi^a + s, \quad s = f_0^{-1} S. \tag{11}$$

The functions φ considered in terms of the dependence on the arguments v^a, r, belong to the Hilbert space H with the scalar product

$$(\varphi_1, \varphi_2) = \int f_0 \varphi_1^* \varphi_2\, d\zeta.$$

Let us assume that H_h is a subspace extended to the family of vectors J_A and H_a is the orthogonal complementary minor to H_h: $H = H_h \otimes H_a$. It is easily established that as a result of conditions (7), the equalities $(J_A, \Phi^a) = 0$ are satisfied and thus the functions Φ^a lie in H_a. On the other hand, $\Phi^0 = \Sigma e_i J_I$ is satisfied and thus the function Φ^0 lies in H_h.

The metric tensor $\gamma_{AB}=(J_A,J_B)$ is defined in the subspace H_h and can be used to increment and decrement the indices A and B. The following metric components are nonzero: $\gamma_{00},\gamma_{ab}=\sigma\delta_{ab},\xi_I=\gamma_{0I}=\gamma_{I0},\ \xi_I=\gamma_{II}$ (no summation is performed over I!).

It is easily established that as a result of conditions (7), the equalities $(J_A,\Phi^a)=0$ are satisfied and thus the functions Φ^a lie in H_a. On the other hand, $\Phi^0=\Sigma e_i J_I$ is satisfied and thus the function Φ^0 lies in H_h. [sic]

In Eqs. (2) and (11) we now convert to the Fourier transforms

$$G\varphi_F=s_F+\beta E_{aF}\Phi^a,\quad G=ik_0+ik_a v^a, \tag{12}$$

$$\varepsilon_{abc}ik_b E_{cF}=-ik_0 B_{aF},\quad ik_a B_{aF}=0, \tag{13}$$

$$\varepsilon_{abc}ik_b B_{cF}=4\pi j^a_F+ik_0 E_{aF},\quad ik_a E_{aF}=4\pi j^0_F. \tag{14}$$

We introduce the additional operators $P_h:H\to H_h,P_a:H\to H$ — projection operators, $I_h:H_h\to H$, $I_a:H_a\to H$ — telescoping, $G_{hh}=P_hGI_h$,

$$G_{ah}=P_aGI_h,\quad G_{ha}=P_hGI_a,\quad G_{aa}=P_aGI_a.$$

The unknown function φ is conveniently divided into ''hydrodynamic'' $h=P_h\varphi$ and ''nonhydrodynamic'' $a=P_a\varphi$ components. By expanding in terms of the basis in H_h, we obtain the components $h_A=(J_A,h)$. We can easily calculate the perturbations of the hydrodynamic four-currents

$$g^0_A=(J_A,\varphi)=h_A,\quad g^a_A=(J_A v^a,\varphi). \tag{15}$$

We shall now assume by analogy with Refs. 1–4 that the sources s in Eq. (10), being functions of the parameters v^a and r, belong to the space H_h and thus are completely characterized by the components s_A. This is a key assumption in the method of converting from kinetics to hydrodynamics proposed in Refs. 1–4. It is easy to see that Eq. (11) yields the system

$$G_{hh}h_F+G_{ha}a_F=s_F,$$

$$G_{ah}h_F+G_{aa}a_F=\beta E_{aF}\Phi^a.$$

This last equation can be used to express a as a function of h and the electric field

$$a_F=E_{aF}\beta G^{-1}_{aa}\Phi^a-G^{-1}_{aa}G_{ah}h_F. \tag{16}$$

For an accurate determination of the operator G^{-1}_{aa} we need to make the substitution $k_0\to(k_0-i\varepsilon)$, where ε is an infinitely small positive quantity. We introduce the matrices

$$Z^a_{AB}=(v^aJ_A,J_B),\quad R^{ab}_{AB}=(P_av^aJ_A,\quad G^{-1}_{aa}P_av^bJ_B).$$

Now, using expressions (3), (15), and (16), we can find the constitutive relations

$$g^a_{AF}=(Z^a_{AB}-ik_bR^{ab}_{AB})h^B_F+\beta\sum_j e_jR^{ab}_{AJ}E_{bF}, \tag{17}$$

$$j^0_{\mathrm{in}F}=\sum_i e_ih_{IF}, \tag{18}$$

$$j^a_{\mathrm{in}F}=\sum_i e_i\left(\beta\sum_j e_jR^{ab}_{IJ}E_{bF}-ik_bR^{ab}_{IB}h^B_F\right). \tag{19}$$

Since the coefficients of the functions R^{ab}_{AB} depend nonpolynomially on the wave four-vector k_α, the hydrodynamic model is nonlocal in space and time.

The constitutive relations (17)–(19) close the problem (2) and (10). The conversion to the hydrodynamic problem is accurate. If some solution of this problem is known, Eqs. (15) and (16) can be used to find the components h and the distribution functions, and to reconstruct the solution of the problem in the kinetic formulation.

The matrix R^{ab}_{AB}, which characterizes the nonlocality of the theory, satisfies various general conditions. Quite clearly, for the real four-vector k_α we have

$$R^{ab}_{AB}(k_\beta)^*=R^{ab}_{AB}(-k_\beta). \tag{20}$$

Although the functions J_A are by definition linearly independent, the set of functions J_A,v^aJ is linearly dependent. Generally speaking, each identity of the type

$$\lambda^AJ_A+\Lambda^A_av^aJ_A=0, \tag{21}$$

where λ^A and Λ^A_a are constant coefficients, yields relations

$$\Lambda^A_aR^{ab}_{AB}=0,\quad \Lambda^B_bR^{ab}_{AB}=0. \tag{22}$$

It is easy to show that one identity of the type (21) is

$$\sum_i m_iv^aJ_I-J_a=0. \tag{23}$$

If there are no internal degrees of freedom r', another identity also exists

$$v^aJ_a-2J_0=0. \tag{24}$$

Introducing the notation $\nu=k_ak_a$ and using the Sokhotskii–Plemel formula, we can easily obtain

$$R^{ab}_{AB}+R^{ba}_{BA}{}^*=(P_av^aJ_A,(G^{-1}_{aa}+G^{-1+}_{aa})P_av^aJ_B)$$
$$=2\pi\nu^{-1/2}(P_av^aJ_A,\delta(k_0+k_cv^c)P_av^bJ_B).$$

From this it follows that for any nonzero set of complex quantities C^A_a the following inequality is satisfied

$$(R^{ab}_{AB}+R^{ba}_{BA}{}^*)C^A_a{}^*C^B_b>0. \tag{25}$$

It follows from the definitions that the coefficients of the function R^{ab}_{AB} are continued analytically in terms of the parameter k_0 in the lower complex half-plane. We divide k_0 into complex and imaginary parts: $k_0=\omega_1+i\omega_2$, $\omega_2\leqslant 0$. We then have

$$R^{ab}_{AB}+R^{ab}_{BA}{}^*=(P_av^aJ_A,\Lambda(k_\alpha,v^c)P_av^bJ_B),$$

$$\Lambda=-2\omega_2(\omega_2^2+(\omega_1+k_cv^c)^2)^{-1}.$$

From this it can be seen that inequality (25) is satisfied in the entire half-plane $\omega_2\leqslant 0$.

We shall discuss the consequences of the reversibility of the processes at the microlevel. Let us assume that I is the time reversal operator in the space H (Refs. 1–4). The operator I reverses the sign of the particle velocities $Iv^aI=-v^a$. The integrals J_A and the state f_0 are eigenfunctions of the operator I

$$IJ_a=\varepsilon_AJ_A,\quad \varepsilon_A=\pm 1,\quad If_0=f_0,$$

where $\varepsilon_0=\varepsilon_I=1$, $\varepsilon_a=-1$.

This gives the reciprocity conditions (analog of the Onsager relations)

$$R^{ab}_{AB}(k_0,k_c)=\varepsilon_A\varepsilon_B R^{ba}_{BA}(k_0,-k_c). \tag{26}$$

Equation (17) directly yields a representation for the viscous stresses, the heat flux, and the diffusion fluxes

$$\tau^{ab}_F=ik_c R^{ac}_{bB}h^B_F-\beta\sum_j e_j R^{ac}_{bJ}E_{cF},$$

$$q^a_F=-ik_b R_{0B}h_F+\beta\sum_j e_j R_{0J}E_{bF},$$

$$d^a_{iF}=-ik_b R^{ab}_{IB}h_F+\beta\sum_j e_j R^{ab}_{IJ}E_{bF}.$$

The identities (23) and (22) for the diffusion fluxes yield the natural result $\Sigma_i d^a_{IF}=0$. In the absence of internal degrees of freedom, it follows from the identities (22) and (24) that there is no second viscosity $\tau^{aa}_F=0$.

CALCULATION OF NONLOCAL KERNELS

It has been noted that for a collisionless plasma we can calculate the dependence of the coefficients of the functions R^{ab}_{AB} on the four-vector k_α. Note that the following equations are satisfied

$$P_a v^a J_A=v^a J_A-Z^{aB}_A J_B,\quad Z^{aB}_A=Z^a_{AC}\gamma^{BC}. \tag{27}$$

We shall analyze the functions of the set of parameters ξ^A

$$\varphi=\varphi(\xi^A)=\exp(\xi^A J_A),$$

$$U^{ab}=U^{ab}(\xi^A)=(\varphi,(k_0-i\varepsilon+k_c v^c)^{-1}v^a v^b),$$

$$V^a=V^a(\xi^A)=(\varphi,(k_0-i\varepsilon+k_c v^c)^{-1}v^a),$$

$$W=W(\xi^A)=(\varphi,(k_0-i\varepsilon+k_c v^c)^{-1}),$$

$$H^{ab}_{AB}=H^{ab}_{AB}(\xi^A)=\frac{\partial^2 U^{ab}}{\partial\xi^A\partial\xi^B}-Z^{bC}_B\frac{\partial^2 V^a}{\partial\xi^A\partial\xi^C}-Z^{aC}_A\frac{\partial^2 V^b}{\partial\xi^B\partial\xi^C}+Z^{aC}_A Z^{bD}_B\frac{\partial^2 W}{\partial\xi^C\partial\xi^D}.$$

Using the Eq. (27) and the definition of the coefficients R^{ab}_{AB}, it is easily established that the values of R^{ab}_{AB} are calculated in terms of the function

$$H^{ab}_{AB}=H^{ab}_{AB}(\xi^A):R^{ab}_{AB}=i^{-1}H^{ab}_{AB}|_{\xi^A=0}.$$

Thus, the problem has been reduced to calculating the functions U^{ab}, V^a, and W. The solution of this last purely technical problem is given in Appendix A.

It is easy to see that for $k_a=0$ a simple expression is obtained for the coefficients of the functions

$$R^{ab}_{AB}=(ik_0)^{-1}(P_a v^a J_A,P_a v^b J_B). \tag{28}$$

The set of coefficients $R^{ab}_{AB}(k_\alpha)$ can be conveniently expanded in terms of the basis of the quantities $I^{abn}_{AB}(k_c)$, which have a polynomial dependence on the wave vector k_a and satisfy the condition $I^{abn}_{AB}(k_c)^*=I^{abn}_{AB}(-k_c)$

$$R^{ab}_{AB}(k_\alpha)=I^{abn}_{AB}(k_c)X_n(k_\alpha). \tag{29}$$

Here $X_n(k_\alpha)$ are scalar functions which by virtue of Eq. (20) satisfy the condition $X_n(k_\alpha)^*=X_n(-k_\alpha)$. Thus, they are Fourier transforms of certain real kernels $Y_{nF}=X_n$. The explicit form of the expansion (29) is given in Appendix B.

DYNAMICS OF FREE OSCILLATIONS

We shall consider the problem of free plasma oscillations. From Eqs. (10), (13), and (14) we obtain the system

$$ik_\alpha g^\alpha_{AF}=0, \tag{30}$$

$$\varepsilon_{abc}ik_b E_{cF}=-ik_0 B_{aF},\quad ik_a B_{aF}=0, \tag{31}$$

$$\varepsilon_{abc}ik_b B_{cF}=4\pi j^a_{\text{in}\,F}+ik_0 E_{aF},\quad ik_a E_{aF}=4\pi j^0_{\text{in}\,F}. \tag{32}$$

We note that because of the electrical neutrality condition (8), the following equality is satisfied

$$\sum_i e_i Z^a_{IA}=0.$$

Thus we have $j^\alpha_{\text{in}}=\Sigma_i e_i g^\alpha_I$ and the compatibility condition (4) of the Maxwell equations (31) and (32) follows automatically from Eq. (30). Since the constitutive relations (17)–(19) contain no magnetic field, it is sufficient to investigate the system (29) and the equation for the electric field deduced from Eqs. (31) and (32)

$$i(4\pi k_0)^{-1}(k_0^2E_{aF}-\nu E_{af}+k_a k_b E_{bF})+j^a_{\text{in}\,F}=0. \tag{33}$$

We shall interpret Eqs. (30) and (33) as a system of linear equations for a set of unknowns h^A_F, E_{aF}. Let us assume that $A=(A_{nm})$ is the matrix of the system. As the condition for the existence of a nontrivial solution we obtain the dispersion equation

$$\det A=0. \tag{34}$$

The analytic expressions for the functions R^{ab}_{AB} obtained in Appendix A can be used to seek complex solutions of k_α. We shall assume that the real wave vector k_a is defined. Equation (34) then defines the set of frequencies k_0 of the plasma natural oscillations as a function of the wave number. The matrix A as a function of the frequency k_0 is continued analytically into the lower complex half-plane $\operatorname{Im}k_0\leqslant 0$ with the isolated point $k_0=0$. Let us assume that C^n is an arbitrary nonzero set of complex quantities. Using Eqs. (25), (17), and (19), it is easily established that in the lower complex half-plane with the isolated point, the following inequality is satisfied

$$\operatorname{Re}(A_{nm}C^{n*}C^m)>0.$$

Thus for $\operatorname{Im}k_0\leqslant 0$ Eq. (34) has no solutions or, in other words, in the nonlocal hydrodynamic description all the natural oscillations of a collisionless plasma are damped.

From the right-hand side of Eq. (34) the factor $Q_t(k_\alpha)^2$, corresponding to the transverse waves is factorized. For these waves we have

$$h^0=h^1=0, \quad k_a h_F^a=0, \quad k_a E_{aF}=0.$$

The function $Q_t(k_\alpha)$ can be expressed in terms of the scalar functions X_n, using the representation (A2)

$$Q_t=q_1q_4-q_2q_3,$$

$$q_1=ik_0\sigma+\nu X_8, \quad q_2=-\nu\beta\sum_j e_j X_{10\,J},$$

$$q_3=\nu\beta\sum_i e_i X_{11\,I},$$

$$q_4=i(4\pi k_0)^{-1}(k_0^2-\nu)+\beta\sum_i e_i e_j X_{13\,I\,J}.$$

We shall now analyze the dynamics of long-wavelength oscillations caused by a longitudinal electric field. In this case, we find $E_{aF}=-ik_a\varphi(k_\alpha)$. Expressions (30) and (32) yield the system of equations

$$0=ik_0\gamma_{AB}h_F^B+ik_aZ_{AB}^a h_F^B+k_ak_bR_{AB}^{ab}h_F^B + k_ak_b\beta\sum_j e_jR_{AJ}^{ab}\varphi,$$

$$\nu\varphi=4\pi\Sigma_i e_i h_{IF}.$$

In the limit $k_a\to 0$, formula (28) can be used to obtain an equation for φ

$$0=(-k_0^2+\omega_L^2)\varphi, \quad \omega_L^2=4\pi\beta(\Phi^1,\Phi^1).$$

This gives the usual spectrum of Langmuir oscillations $k_0=\pm\omega_L+0(\nu)$.

Thus, the proposed nonlocal hydrodynamic description reproduces the results obtained previously by different approaches.[5,6]

CONCLUSIONS

It has been shown that in a certain class of sources the nonlocal hydrodynamic description accurately corresponds to the kinetic description. This distinguishes the proposed approach from other hydrodynamic schemes. When solving a nonlocal hydrodynamic problem, we can be certain that each hydrodynamic process corresponds exactly to a certain kinetic process. The investigation was confined to linear theory but in fact this constraint is not significant. The method can be simply generalized to the nonlinear case when the sources in Eq. (1) are not necessarily weak. Naturally, in this last case the functional form of the constitutive relations is substantially more complex.

Note that the representation (29) and (A2) does not depend on the fact that a collisionless plasma was analyzed. It will also be valid for a collisional plasma for instance. For long waves ($k_a\to 0$) and slow processes ($k_0\to 0$) terms corresponding to the normal viscosity, heat conduction, and diffusion can be isolated in the expansion (A2). However, it can be seen from formula (28) that when taking the next limit $k_a\to 0$, $k_0\to 0$, the expressions for the corresponding transport coefficients are unbounded.

APPENDIX A:

Here, we shall calculate the functions $U^{ab}=U^{ab}(\xi^A)$, $V^a=V^a(\xi^A)$, and $W=W(\xi^A)$. We recall that the integral in the sense of the principal value

$$N=V.p.\int_{-\infty}^{+\infty} x^{-1}\exp(-\alpha x^2+\beta x)dx, \quad \alpha>0 \qquad (A1)$$

is calculated in terms of a special function, i.e., the error function

$$\Phi(x)=2\pi^{-1/2}\int_0^x \exp(-t^2)\,dt.$$

The necessary procedures are described in Ref. 5. The result of the calculations is: $N=\pi\Theta(\beta 2^{-1}\alpha^{-1/2})$, where $\Theta(z)=i^{-1}\Phi(iz)$.

We shall make a detailed analysis of the procedure for calculating the function $W(\xi^A)$ for $\nu=k_ak_a\neq 0$. We set $e_1^a=\nu^{-1/2}k_a$, and assume that e_2^a and e_3^a are two other unit vectors such that all three vectors e_b^a form the orthogonal basis $e_b^a e_c^a=\delta_{bc}$. Performing the substitution $v^a=e_b^a w^b$, we reduce the expression for the function $W(\xi^A)$ to the form

$$W(\xi^A)=\nu^{-1/2}\sum_i (\Omega-i\varepsilon+w^1)^{-1} \times\exp\left(-\frac{1}{2}A_1w^bw^b+b_i^bw^b+C_i\right) \times dw^1dw^2dw^3d\zeta'(r'),$$

$$A_i=(-\beta+\xi^0)m_i, \quad b_i^a=B_i^be_b^a, \quad B_i^a=\xi^am_i,$$

$$C_i=(-\beta+\xi^0)U(i,r')+\beta\mu_{i0}+\xi^1, \quad \Omega=k_0\nu^{-1/2}.$$

Integration with respect to w^2 and w^3 is reduced to taking the Gaussian integrals. The integral over w^1 is reduced to the integral (A1) by using the Sokhotskii–Plemel formula and shifting the argument. We finally obtain

$$W=\sum_i\int\Psi(i,r')d\zeta'(r'),$$

$$\Psi(i,r')=\nu^{-1/2}\pi(2\pi/A_i) \times\exp(C_i+B_i^a(\delta_{ab}-\nu^{-1}k_ak_b)B_i^b/(2A_i)) \times\exp\left(-\left(\frac{1}{2}A_ik_0^2+B_i^ak_ak_0\right)\nu^{-1}\right) \times(\Theta((B_i^ak_a+A_ik_0)\nu^{-1/2}/(2A_i)^{1/2})+i).$$

The functions $U^{ab}=U^{ab}(\xi^A)$ and $V^a=V^a(\xi^A)$ can be calculated similarly. They can be expressed in the compact form

$$U^{ab}=\sum_i\int\Psi^{ab}(i,r')d\zeta'(r'),$$

$$V^a = \sum_i \int \Psi^a(i,r')\,d\xi'(r'),$$

$$\Psi^{ab}(i,r') = \frac{\partial^2 \Psi}{\partial B_i^a \partial B_i^b}, \quad \Psi^a(i,r') = \frac{\partial \Psi}{\partial B_i^a}.$$

APPENDIX B:

The purely algebraic structure of the coefficients R_{AB}^{ab} allows us to write an expansion of the form (28) which is invariant with respect to the rotation group. For this we need to bear in mind the symmetry conditions

$$R_{cA}^{ab} = R_{aA}^{cb}, \quad R_{Ac}^{ab} = R_{Ab}^{ac}$$

and the reciprocity conditions (26). It is sufficient to write the expression (28) for the components R_{00}^{ab}, R_{0c}^{ab}, $R_{0\,I}^{ab}$, R_{cd}^{ab}, $R_{c\,I}^{ab}$, and $R_{I\,J}^{ab}$

$$R_{00}^{ab} = \delta_{ab} X_0 + k_a k_b X_1,$$

$$R_{0c}^{ab} = R_{0b}^{ac} = ik_a \delta_{bc} X_2 + (ik_b \delta_{ac} + ik_c \delta_{ab}) X_3 + ik_a k_b k_c X_4,$$

$$R_{0\,I}^{ab} = \delta_{ab} X_{5\,I} + k_a k_b X_{6\,I},$$

$$R_{cd}^{ab} = \delta_{ac}\delta_{bd} X_7 + (\delta_{ab}\delta_{cd} + \delta_{cb}\delta_{ad}) X_8 + \delta_{ac} k_b k_d X_7 + (k_a k_b \delta_{cd} + k_c k_b \delta_{ad} + k_a k_d \delta_{cb} + k_c k_d \delta_{ab}) X_8 + k_a k_b k_c k_d X_9,$$

$$R_{c\,I}^{ab} = (ik_a \delta_{bc} + ik_c \delta_{ab}) X_{10\,I} + ik_b \delta_{ac} X_{11\,I} + ik_a k_b k_c X_{12\,I},$$

$$R_{I\,J}^{ab} = \delta_{ab} X_{13\,I\,J} + k_a k_b X_{14\,I\,J}. \qquad \text{(B1)}$$

[1] O. Yu. Dinariev, Zh. Éksp. Teor. Fiz. **106**, 161 (1994) [JETP **79**, 88 (1994)].
[2] O. Yu. Dinariev, Izv. Vyssh. Uchebn. Zaved. Fiz. No. 2, 95 (1995).
[3] O. Yu. Dinariev, Akust. Zhurn. **41**, 415 (1995) [Acoust. Phys. **41**, 364 (1995)].
[4] O. Yu. Dinariev, Zh. Éksp. Teor. Fiz. **107**, 1877 (1995) [JETP **80**, 1041 (1995)].
[5] *Plasma Electrodynamics*, edited by A. I. Akhiezer *et al.* (Pergamon Press, Oxford, 1975; Nauka, Moscow 1974, 720 pp.).
[6] A. F. Aleksandrov, L. S. Bogdankevich, A. A. Rukhadze, *Oscillations and Waves in Plasmas* [in Russian], Moscow State University Press, Moscow (1990), 272 pp.

Translated by R. M. Durham

Initiated subcritical microwave streamer discharge and the problem of global elimination of Freons from the Earth's atmosphere

L. P. Grachev, I. I. Esakov, and K. V. Khodataev

Moscow Radio Engineering Institute, Russian Academy of Sciences, 113519 Moscow, Russia
(Submitted January 29, 1997; resubmitted April 9, 1998)
Zh. Tekh. Fiz. **68**, 15–24 (December 1998)

An analysis is made of an initiated streamer discharge in a microwave wave beam at subcritical field from the point of view of using this type of discharge for global elimination of Freons, which are damaging to the ozone layer, from the Earth's atmosphere. The various stages of discharge evolution are illustrated with photographs and the physical factors determining its properties at these stages are indicated. © *1998 American Institute of Physics.* [S1063-7842(98)00312-2]

INTRODUCTION

The main anthropogenic factor responsible for reducing the ozone content in the Earth's atmosphere and the formation of holes in the ozone layer is the buildup of Freons,[1] the most active of which are CF_2Cl_2 and $CFCl_3$. The methods of using Freons are such that sooner or later they enter the atmosphere. Freons are chemically inert and persist in the atmosphere for many decades, gradually accumulating. After rising to the stratosphere, they decompose under the action of the sun's hard ultraviolet radiation. The product, atomic chlorine Cl, acts as a catalyst for ozone decomposition reactions. The influence of Freons on theO_3 balance in the ozonosphere can only be reduced substantially by restricting their productions and eliminating from the atmosphere the Freons already accumulated .

Tsand *et al.*[2] proposed a method of purification using microwave radiation which heats the ionospheric electrons and ultimately converts Cl atoms into Cl^- ions which do not participate in catalytic reactions with O_3.

The authors of Ref. 3 pointed out that the Freon concentration in the ionosphere is low and suggested using a freely localized discharge in a microwave beam at lower altitudes H where the concentration is substantially higher. Freons dissociate in the discharge to form Cl^- as a result of the ''attachment'' of free electrons. The cross section of this process for $CFCl_3$ andCF_2Cl_2 is highest at electron energies $0 \leqslant \varepsilon_q \leqslant 1$ eV. This reaction has almost no threshold and the role of the discharge essentially involves generating comparatively ''cold'' electrons with the required concentration n_e. Undesirable compounds such as oxides of nitrogen may also form in a microwave discharge[4] and if these enter the stratosphere where the O_3 concentration is highest, the situation may be even worse. With this in mind, Askaryan *et al.*[3] suggested igniting the discharge in the troposphere ($H = 10-15$ km). At these altitudes, the chemically active Freon fragments and other discharge products will eventually fall to Earth through precipitation and will not enter the stratosphere.

A necessary condition for the establishment of a freely localized discharge is that the electrical component of the electromagnetic wave E_0 should exceed the breakdown level $E_{\rm br}$. This condition determines the energy cost of the dissociation of Freon molecules $\Delta\varepsilon$. An estimate made in Ref. 3 for $H = 15$ km gives $\Delta\varepsilon = 20$ keV/mol. At lower altitude the ''cost'' merely increases. The power of the electromagnetic radiation capable of igniting a discharge at these altitudes was also estimated[3] and was found to be considerable which, in the view of the authors,[3] casts doubts on the feasibility of this method for global purification of the atmosphere.

However, at comparatively high air pressures p, i.e., at low altitudes, the microwave discharge is a streamer and propagates into the region of the electromagnetic beam where $E_0 \ll E_{\rm br}$ (Ref. 5). This property means that a discharge can be ignited with breakdown conditions only achieved locally.[6–8] In a uniform field with cw radiation, $E_{\rm br}$ for a freely localized discharge is the same as the so-called critical field[9] $E_{\rm cr}$ and a discharge with $E_0 < E_{\rm cr}$ may be described as subcritical.

In Ref. 6 we investigated a discharge initiated by a metal sphere. In this case, when $p \geqslant 100$ Torr ($H = 15$ km) a discharge can be ignited at $E_0 \cong (1/3)E_{\rm cr}$, i.e., the subcriticality factor $\Psi = 3$ can be achieved. In the presence of a sphere, the energy density flux of the electromagnetic radiation Π_0, which is proportional to E_0^2, required to ignite the discharge may be reduced by almost an order of magnitude. breakdown, Microwave breakdown can be achieved with considerably weaker fields if a ''vibrator'' is used to initiate it. For instance, the authors of Ref. 10 succeeded in achieving $\Psi = 60$ by placing a resonant vibrator parallel to $\mathbf{E}_0$.

When an initiated discharge is used as an electron supplier for the dissociation of Freons, the cost $\Delta\varepsilon$ is no longer determined by the inequality $E_0 > E_{\rm cr}$ and may be lower. In this case, all the ideas and work proposed in Refs. 2 and 3 are now being continued and the possibility of implementing these is becoming a realistic proposition. It is extremely expensive to carry out experiments in the atmosphere at any appreciable altitudes which makes laboratory studies of this process quite topical. There is a particular need for studies of a microwave discharge with $\Psi \gg 1$.

Here we generalize known information on this type of discharge and report new experimental results which refine various properties of this discharge.

EXPERIMENTAL CONDITIONS

The experiments were carried out in air using the apparatus described in Ref. 11. Pulsed, linearly polarized $\lambda = 8.5$ cm microwave radiation with a rectangular envelope was focused into an anechoic hermetically sealed chamber. At the focus in the plane perpendicular to the Poynting vector $\mathbf{\Pi}_0$, the field E_0 had an approximately Gaussian distribution with a characteristic scale of several centimeters. Parallel to $\mathbf{\Pi}_0$ the field was almost uniform over a length of approximately 10 cm.

To initiate the discharge a lead sphere of diameter $2a = 2.5$ mm was positioned at the focus of the electromagnetic beam using 0.1 mm diameter Kapron thread, perpendicular to $\mathbf{E}_0$ or a microwave vibrator made of copper wire with hemispherically rounded ends, having a diameter $2a = 0.8$ mm and different lengths $2L$, was placed parallel to $\mathbf{E}_0$.

These initiators were exposed to ultraviolet radiation from a special source to ensure that seed photoemission electrons were present at the surface.

The microwave pulse duration t_{pul} was varied between 1 and 40 μs and the intervals between pulses were at least 1 min. The field could be varied in different pulses. The amplitude of the microwave pulse envelope was measured in experiments using a calibrated field detector. The measured value was set to a specific value of E_0 at the beam focus. The value of E_0 at the focus was determined with an accuracy better than 10% in the range between 6.5 and 2 kV/cm. For weaker fields the accuracy dropped substantially because of ''induction'' to the detector. For this reason, in this series of experiments no vibrator experiments were carried out for $E_0 < 2$ kV/cm. In the sphere experiments it was possible to make additional measurements of the field using a technique described in Ref. 12. These experiments were carried out using weaker fields.

The discharge was photographed. In the photographs shown below, the vector $\mathbf{E}_0$ is vertical and $\mathbf{\Pi}_0$ is directed from right to left. Unless stated otherwise, the photographs were obtained for an exposure time $t_{\text{ex}} > t_{\text{pul}}$. The discharge region was probed by $\lambda = 1$ cm radiation along a line passing through the axis of the main beam and perpendicular to $\mathbf{E}_0$ and $\mathbf{\Pi}_0$. The probing line was 1.5 cm from the focus (closer to the emitter). The apertures of the receiving and transmitting antennae were 3×3 cm. Measurements were made of the attenuation of this radiation as it propagates through the discharge and this was used to estimate the average electron concentration n_e.

PRELIMINARY OBSERVATIONS ON THE PHYSICS OF A MICROWAVE DISCHARGE

A microwave streamer discharge is only established if at the initial stage[6] we have

$$\nu_c \gg \omega, \tag{1}$$

where ν_c is the electron–molecule collision frequency and ω is the angular frequency of the field.

In a microwave discharge with $E \cong E_{\text{cr}}$, the average electron energy $\overline{\varepsilon_q}$ is a few electronvolts.[9] For the estimates we take[13]

$$\nu_c \cong 5 \times 10^9 p, \ s^{-1}. \tag{2}$$

In formula (2) and the following formulas the pressure p is in torr and the gas is taken to be at room temperature T. In fact, the kinetic coefficients are functions of the molecular concentration n. Their known dependences on the pressure p can be converted to dependences on the concentrations n and vice versa by using[13]

$$n = 2.7 \times 10^{19} \frac{p}{760} \frac{273}{T}, \ \text{cm}^{-3}. \tag{3}$$

In the troposphere, p varies between 760 and 100 Torr and relation (1) is valid for $\lambda \gg 0.4$ cm. Under the experimental conditions described, where $\lambda = 8.5$ cm it is satisfied for $p \gg 4$ Torr.

In analyses of the processes in a microwave plasma it is important to estimate the relation between $\delta\nu_c$ and ω, where δ is the average relative electron energy lost in a collision with molecules. In air where $E \cong E_{\text{cr}}$, we find $\delta = 10^{-3} - 10^{-2}$ (Ref. 14). When a microwave discharge is ignited in the proposed natural experiment, the wavelength λ of the electromagnetic radiation will barely exceed a few decimeters because of size constraints. If this wavelength is in the centimeter range, it follows from formula (2) that at tropospheric pressures p the following condition will almost always be satisfied in a microwave plasma

$$\delta\nu_c \ll \omega. \tag{4}$$

In the decimeter range and under the experimental conditions described where $\omega = 2.2 \times 10^{10} \ \text{s}^{-1}$, inequality (4) is satisfied for pressures p of tens of torr and only very approximately for hundreds of Torr. Nevertheless, for our estimates we shall assume that under the experimental conditions this inequality is satisfied up to atmospheric pressure p. In this case, i.e., when inequality (4) is valid, it is conventional to use the concept of an effective field[13]

$$E_{\text{ef}} = E_{\text{a}} / \left(\sqrt{2} \ \sqrt{(1 + (\omega/\nu_c)^2)} \right), \tag{5}$$

where E_{a} is the amplitude of the microwave field, and to use the values of the kinetic coefficients measured in a dc plasma by equating E in this plasma to E_{ef} in the microwave field. Taking this observation into account, the equation for the electron balance in an air plasma may be written as[13]

$$\frac{\partial}{\partial} \frac{n_e}{t} = (\nu_i - \nu_a) n_e + D \nabla^2 n_e - \alpha_r n_e^2, \tag{6}$$

where ν_i and ν_a are the ionization and attachment frequencies, D and α_r are the diffusion and recombination coefficients, and t is the time.

It is usually assumed[9] that only ν_i depends on E, while ν_a, D, and α_r are constant at the chosen pressure p. Near E_{cr} we find[15]

$$\nu_i = \nu_a (E_{\text{ef}}/E_{\text{cr}})^{\beta}, \tag{7}$$

where $\beta = 5.34$,

$$\nu_a = 2\times 10^4 p,\ s^{-1}; \tag{8}$$

$$E_{\text{cr}} = 40p,\ \text{V/cm}. \tag{9}$$

It can be seen from Eq. (7) that E_{cr} is the effective microwave field for which $\nu_i = \nu_a$. In Eq. (7), E_{ef} should generally be understood as the local value of the field which is made up of the ''unperturbed'' field E_0 of the electromagnetic wave and the induced fields in the resulting discharge plasma. In the microwave range, a complete system of Maxwell equations allowing for a ''delay'' effect must generally be solved to find E_{ef} (we shall subsequently use the symbol E to denote the effective value of the field, dropping the ''ef'' subscript).

In a plasma, the ionization process described by Eq. (6) is accompanied by the dissipation of electromagnetic energy. Under isobaric conditions

$$p = nT = \text{const} \tag{10}$$

the heating of the air may be described by

$$\frac{7}{2} n \frac{\partial}{\partial} \frac{T}{t} = \eta \sigma E^2, \tag{11}$$

where η is the relative fraction of the electron energy dissipated in heating the gas within the characteristic times t of the discharge process, and $\sigma(n_e)$ is the plasma conductivity.

The coefficient η in Eq. (11) depends strongly on the parameter E/n (Ref. 14). For instance, for $E_{\text{cr}}/n = 1.3 \times 10^{-15}\,\text{V}\cdot\text{cm}^2$, around 40% of the energy acquired by the electrons in the electromagnetic field is initially transferred to vibrations of N_2 molecules and 60% is dissipated in the electronic excitation of N_2 and O_2. The vibrational reservoir remains almost ''frozen'' over the characteristic times of most of the discharge stages and only the electronic excitation energy relaxes into heat within $t \cong 10^{-8}$ s. This initial energy distribution is a prerequisite for a highly nonequilibrium microwave plasma and stimulates plasma-chemical processes in the plasma.

We shall subsequently use cylindrical (z, ρ, φ) and matched spherical (r, θ, φ) coordinates with the origin at the focus of the electromagnetic beam and the Z axis parallel to $\mathbf{E}_0$.

The properties of the plasma relative to the electromagnetic field can be characterized by its complex permittivity

$$\varepsilon = 1 - i\frac{\sigma}{\varepsilon_0 \omega} = 1 - i\frac{n_e}{n_{e0}}, \tag{12}$$

where $\varepsilon_0 = 10^{-9}/(36\pi)$, F/m; m_e and q_e are the electron mass and charge, and the concentration

$$n_{e0} = \frac{\varepsilon_0 m_e}{q_e^2} \nu_c \omega,\ \text{m}^{-3} \tag{13}$$

is the ''scale'' of n_e. For $n_e = n_{e0}$ the conductivity is $\sigma = \varepsilon_0 \omega$. As the pressure p increases from 100 to 760 Torr and $\omega = 2.2\times 10^{10}\,\text{s}^{-1}$ the concentration n_{e0} increases from 3×10^{12} to $2\times 10^{13}\,\text{cm}^{-3}$ (for information we note that when $n_e = n_{e0}$ the skin layer depth is $\Delta = 0.35\lambda$).

In order to assess the influence of a plasma with a specific value of n_e on the electromagnetic field, we insert a homogeneous plasma sphere of radius $a \ll 1/k = \lambda/(2\pi)$ and $a < \Delta$ in this field. For a uniform field $\mathbf{E}_0$ the field inside the sphere $\mathbf{E}_{\text{in}}$ remains uniform and parallel to $\mathbf{E}_0$ (Ref. 16). When $n_e = n_{e0}$, this field is $E_{\text{in}} = 0.95E_0$. At the poles of this sphere ($\theta = 0$ and $r = a$) the external field is $E_{\text{out}} = 1.3E_0$. It can be seen that a small drop in E_{in} was accompanied by an appreciable increase in E_{out}. If an ionization process takes place in the field E_{out}, it follows from Eq. (7) that the ionization frequency ν_i will increase by more than a factor of 1.5 in this field [Eq. (7) indicates that when E shows a small variation of χ%, the frequency ν_i varies by $(\beta\chi)$%]. For a plasma ellipsoid elongated along $\mathbf{E}_0$, this property shows up even more strongly.[16] As the eccentricity increases, we find $E_{\text{in}} \to E_0$, and the field at the poles begins to exceed E_0 quite significantly. This characteristic of elongated plasmoids to some extent determines the properties of a streamer microwave discharge.[17]

MICROWAVE BREAKDOWN IN AIR

The main problem involved in implementing a subcritical discharge is its ignition. The discharge must be initiated. This has been achieved experimentally by using dielectric and metal-dielectric plates,[7] a laser jet at a solid target,[7] a brush of thin wires,[8] a ring source of ionizing radiation,[18] a laser spark,[19] a metal sphere,[12] or a microwave vibrator.[10] Most of these methods produced discharge ignition with $\Psi \leqslant 10$. The use of a microwave vibrator was unique[10] in that atmospheric air was successfully broken down with $\Psi = 60$ and this is the method which we shall analyze below.

Figure 1 shows a discharge initiated at 760 Torr ($E_{\text{cr}} = 30\,\text{kV/cm}$), $E_0 = 3\,\text{kV/cm}$ ($\Psi = 10$), and $t_{\text{pul}} = 40\,\mu\text{s}$ using a vibrator of length $2L = 10$ mm and diameter $2a = 0.8$ mm. The vibrator is positioned on the left of the figure. The distance between its luminous rings gives a geometric scale. For $E_0 = 3$ kV/cm the length $2L = 10$ mm is the shortest which ensures breakdown. For these values of the pressure p, initial field E_0, and vibrator diameter $2a$ no discharge was ignited for lengths $2L < 10$ mm.

As E_0 decreases keeping the diameter $2a$ the same, the length $2L$ must be increased to ignite the discharge (vibrators with $2a < 0.8$ mm were not used in the experiments since the shape of their ends was difficult to monitor). Figure 2 shows a discharge initiated at 760 Torr, $E_0 = 2\,\text{kV/cm}$ ($\Psi = 15$), and $t_{\text{pul}} = 3\,\mu\text{s}$ using a vibrator of length $2L = 18$ mm. This especially shows a photograph of a discharge with short t_{pul}. It can be seen that the discharge is created at the ends of the vibrator where its near field is strongest.[20] We shall estimate the initiation conditions typical of Fig. 1. From Ref. 10 the field at the poles of the vibrator ($z = \pm L$), parallel to the vector $\mathbf{E}_0$, for $ka \ll \pi/2$ and $kL \leqslant \pi$ is given by

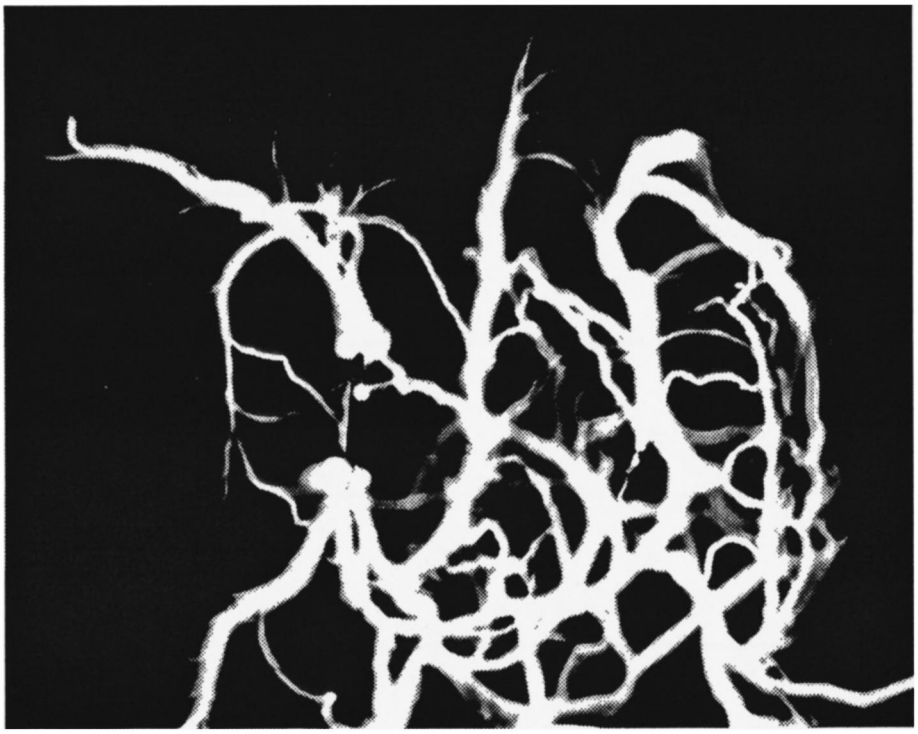


FIG. 1. Subcritical microwave discharge at $p=760$ Torr, $E_0=3$ kV/cm, and $t_{\text{pul}}=40$ μs initiated by a conductor with $2L=10$ mm and $2a=0.8$ mm, $\Psi=9$, and calculated $E_{\text{out}}=27$ kV/cm.

$$E_{\text{out}} \cong E_0 \sqrt{\left\{1+A\left[\frac{1}{kL}\cos(kL)+\sin(kL)\right]\right\}^2+\left\{A\left[\frac{1}{kL}\sin(kL)-\cos(kL)\right]\right\}^2}. \tag{14}$$

Here the coefficient is

$$A=\frac{2}{3\pi}\frac{Z_0}{Z}\frac{h_{\text{ef}}}{a}, \tag{15}$$

$Z_0=120\pi\ \Omega$ is the wave impedance of free space, $h_{\text{ef}}\cong(4/\pi)\,L$ is the effective vibrator length, $Z=\sqrt{R_\Sigma^2+X_0^2}$ is the modulus of its vector impedance

$$R_\Sigma \cong \frac{1}{6}\frac{Z_0}{\pi}(kh_{\text{ef}})^2$$

is the resistance of the vibrator radiation,

$$X_0=-\frac{Z_0}{\pi}[\ln(L/a)-1]\cot(k/L)$$

is its reactance for $(L/a)\gg 1$, and

$$X_0=-\frac{Z_0}{\pi}\frac{h_{\text{ef}}}{3ka^2}$$

is its capacitive reactance for $L=a$.

Equation (14) indicates that for fixed a, E_{puls} increases to a resonant value equal to $\lambda/2$ as $2L$ is approached. Equation (15) also indicates that the coefficient A which determines the field at the vibrator tips is inversely proportional to a. However, this does not mean that by reducing a, we can break down air at a specific pressure p with an arbitrarily low value of E_0. The near field of the vibrator only satisfies the condition $E\gg E_0$ in very small regions near its ends.[20]

In order to find E_{br} we need to solve Eq. (6), neglecting recombination but allowing for diffusion, for a specific field geometry.[13] For a comparatively long duration t_{pul} the condition $\partial n_e/\partial t=0$ is taken as the breakdown criterion and the corresponding value of E_0 is taken as E_{br}. It can be seen from Eq. (6) that in a uniform field, electron losses are only caused by attachment and the breakdown criterion is satisfied for $\nu_i=\nu_a$, or, taking account of Eq. (7) $E_{\text{br}}=E_{\text{cr}}$.

In Ref. 12 the problem of breakdown with $D\neq 0$ was solved for a sphere with $ka\ll\pi/2$. At its poles the field is $E_{\text{out}}=3E_0$, as is deduced from Eq. (14) for $L=a$. However, as the sphere is moved away from the surface, E decays rapidly,[16] approximately as $1/r^3$.

It was shown in Ref. 12 that to within 10% we obtain $E_0=E_{\text{br}}=(1/3)\,E_k$ in the presence of a sphere when

$$a\gg 10^2\,L_a, \tag{16}$$

where the attachment length is $L_a=\sqrt{D/\nu_a}$.

A comparison between theory and experiment reveals that D must be taken as the coefficient of ambipolar diffusion

$$D_a=\frac{1.4\times 10^4}{p},\ \text{cm}^2/\text{s} \qquad \text{(Ref. 9)}. \tag{17}$$

For example, at 760 Torr the attachment length is $L_a=10^{-3}$ cm and criterion (16) has the form $a\geq 0.1$ cm, i.e., under these conditions breakdown only takes place in the polar regions of the sphere, as in a uniform field E_0 of three times the value.

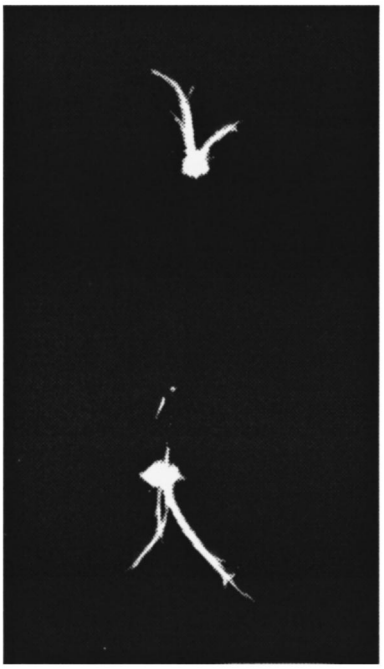


FIG. 2. Subcritical microwave discharge at $p=760$ Torr, $E_0=2$ kV/cm, and $t_{pul}=3$ μs initiated by a conductor with $2L=18$ mm and $2a=0.8$ mm, $\Psi=15$, and calculated $E_{out}=41$ kV/cm.

If condition (16) is not satisfied, as a decreases we find $E_{br}\rightarrow E_{cr}$. The theory developed in Ref. 12 cannot be applied when $a\leqslant 3L_a$, although the experiment showed that when $a\leqslant 2L_a$ the field is $E_{br}=E_{cr}$ to within 10%. In this case, breakdown takes place as if no sphere were present in the field E_0.

The problem of microwave breakdown in the presence of a vibrator with allowance for diffusion has not yet been resolved. Quite clearly the solution for a sphere will depend on the ratio between a and L_a. However, for a vibrator with $L>a$ the field near the ends decreases more slowly with increasing distance from the vibrator surface compared with a sphere.[20] Thus, it may be predicted that the vibrator radius at which diffusion does not influence the breakdown process will be smaller than that given by criterion (16).

For example, under the conditions obtaining in Fig. 1, Eq. (14) gives $E_{out}/E_0=9$. Thus, for $E_0=3$ kV/cm, we obtain $E_{out}=27$ kV/cm, which is the same as E_{cr} within experimental error, i.e., as is deduced experimentally, a discharge should not be ignited with a vibrator of given $2a$ but smaller $2L$. In this particular case, the dimension a is equal to $(40L_a)$. Thus, for this ratio of a and L_a the spatial nonuniformity of the breakdown process still need not be taken into account.

Nevertheless, for thin vibrators a criterion of the type (16) still needs to be determined. Ultimately, the depth of subcriticality at which air can be broken down in a microwave field and thus to some extent, the feasibility of carrying out natural experiments, depend on this criterion.

DISCHARGE EVOLUTION PROCESS

Let us now return to Fig. 1, which shows a typical streamer structure. On the integrated photographs ($t_{ex}>t_{pul}$) this structure is a chaotic web of plasma channels of varying thickness and brightness. Their diameter does not exceed 1 mm and the typical distance between the channels is tenths of λ. The direction of the channels is weakly correlated with the direction of $\mathbf{E}_0$. Their relative position is not reproduced in different pulses and the discharge process is accompanied by a sound shock.

In the range $1\leqslant\Psi\leqslant 15$ being studied experimentally, an important property of a microwave streamer discharge in terms of the particular application being considered is its capacity to separate from the initiator. This means that in a natural experiment a discharge may be initiated in a large volume having a few initiated vibrators. The maximum value of Ψ at which the discharge still becomes detached from the initiators has not yet been determined. This is one of the most important problems for further laboratory research in addition to determining the minimum value of E_{br} (note that in Ref. 10, the authors did not follow the discharge evolution and merely identified that breakdown had occurred from flashes of light at the ends of the vibrators).

The experiments showed that a discharge detached from the initiator develops according to a specific scheme. At each stage the plasma is generated under different conditions, which must be taken into account when planning a natural experiment. We shall examine these stages subsequently.

The experiments were mainly carried out using an initiating sphere. At the initial stage the discharge has the same geometry in different pulses which simplifies its observation. However, when comparatively long initiating vibrators are used, the axial symmetry of the discharge is impaired almost immediately and a factor appears which randomizes the discharge geometry.

The experiments were carried out at pressures below tropospheric. By varying the pressure p, we were able to investigate the formation of complexities in the discharge structure. As the pressure increases, the rate of many ionization processes increases and in the experiments some of these could only be resolved in time and space at comparatively low pressures.

DIFFUSE STAGE

Figure 3 shows a discharge at 20 Torr, $E_0=0.7$ kV/cm ($\Psi=1.1$), and $t_{pul}=1$ μs, i.e., at the very beginning of its evolution. This reaffirms that the discharge is generated in regions where $E>E_{cr}$ and where an avalanche-like increase in n_e is initiated in accordance with Eqs. (6) and (7). This is an ionization-field process, i.e., as n_e increases, the field undergoes rearrangement. As a result, regions with $E>E_{cr}$ begin to move along $\mathbf{E}_0$ which then leads to displacement of the discharge boundary along this vector. Its velocity is

$$V=2\sqrt{D(\nu_i-\nu_a)}. \tag{18}$$

We must not labor under the delusion of the simplicity of Eq. (18). In this equation, ν_i as given by Eq. (7) is a power function of E which depends on the curvature of the plasma boundary, the n_e profile at the boundary, and the entire configuration of the conducting channel with allowance for its relation to λ.

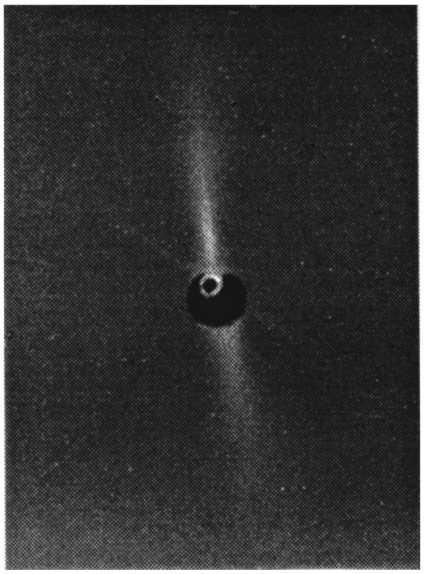


FIG. 3. Microwave discharge at $p=20$ Torr, $E_0=0.7$ kV/cm, and $t_{\text{pul}}=1$ μs initiated by a sphere with $2a=2.5$ mm.

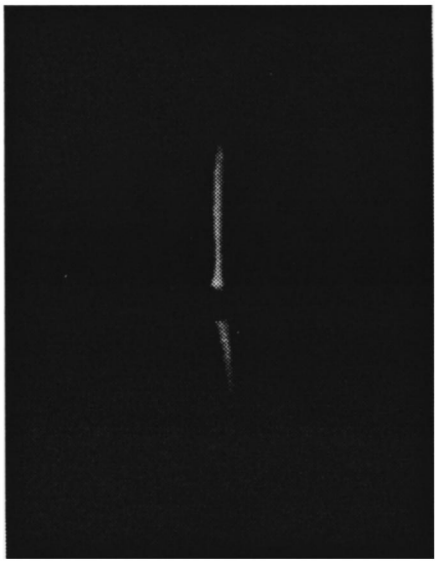


FIG. 4. Microwave discharge at $p=30$ Torr, $E_0=1$ kV/cm, and $t_{\text{pul}}=7$ μs initiated by a sphere with $2a=2.5$ mm.

From Eq. (11) the rate of the dissipative processes in the plasma is characterized by its uniform heating time (the term ''uniform heating'' will become clear from the following analysis):

$$\frac{1}{\gamma_0}=\frac{(2\eta\sigma E^2)^2}{7n_0T_0},$$

where T_0 is the initial gas temperature, and n_0 is the unperturbed molecular concentration in the gas.

This value can be compared with the characteristic time of the ionization processes $1/(\nu_i\nu_a)$ obtained from Eq. (6). Estimates show that for the discharge shown in Fig. 3 we find $(\nu_i-\nu_a)\gg\gamma_0$ and the heating of the gas can be neglected.

The velocity of the plasma boundary given by formula (18) is usually called the diffusion velocity and this type of discharge evolution when heating of the gas in the plasma can be neglected is called the diffuse stage. Under the experimental conditions at pressure $p=20$ Torr, the diffuse stage is also the final stage.[6]

It can be seen from Fig. 3 that as the plasma ''columns'' grow along E_0, their transverse dimension increases. This combined with the electrodynamic characteristics of the microwave range ultimately limits the total length of the diffuse discharge along the Z axis to a length not exceeding $\lambda/2$ (Ref. 6). The discharge remains, as it were, ''locked'' to the initiator. Quite clearly the term ''streamer'' discharge is hardly suitable for this type of discharge.

IONIZATION OVERHEATING STAGE

In the experiments the discharge structure became complex at $p\geqslant 20$ Torr. Above these pressures p a thin bright channel, elongated along the vector $\mathbf{E}_0$, formed in the initial diffuse region with $E>E_{\text{cr}}$ (as an example Fig. 4 shows such a discharge at 30 Torr, $E_0=1$ kV/cm, $\Psi=1.2$, and $t_{\text{pul}}=7$ μs). For low Ψ the characteristic time for evolution of the channel $1/\gamma$ is a few microseconds and its diameter varied over the length. The channel appeared to taper toward the end (Fig. 4). Its maximum diameter decreased with increasing p but, as can be seen from Fig. 4, it did not exceed fractions of a millimeter, even at 30 Torr.

The formation of a channel is now interpreted as follows. As a result of random fluctuation effects, n may decrease in a certain thin cylindrical region of an initial diffuse plasma formation with $E>E_{\text{cr}}$ elongated along $\mathbf{E}_0$. In this case, the value of E remains almost constant and in accordance with Eq. (9), the critical field E_{cr} decreases, i.e., the ionization frequency ν_i as given by Eq. (7) increases. As ν_i increases, the electron concentration n_e and thus the gas temperature T increase in this region at faster rates. Under the conditions (10) this then leads to a further decrease in n. Thus, a positive feedback ''loop'' has been closed and in this sense, the effect is interpreted as ionization overheating instability.[21]

This last term is in a certain sense arbitrary because when $E>E_{\text{cr}}$, the system has no equilibrium state. In the real effect, the channel is formed against an essentially transient plasma background. Nevertheless, the system of equations (6), (7), (10), and (11) can be investigated for stability in the constant field approximation, using a conventional algorithm and assuming[22]

$$\gamma\gg(\nu_i-\nu_a) \quad \text{and} \quad \gamma\gg\gamma_0. \tag{19}$$

This approach in a linear local approximation yields a natural result: the processes of variation of n_e, n, and T which were previously almost unrelated, begin to develop at the same rate, having the increment

$$\gamma=\sqrt{\beta\,\xi^{\beta}\,\nu_a\,\gamma_0}, \tag{20}$$

where $\xi=E/E_{\text{cr}}$.

This interpretation explains most of the experimentally observed characteristics of channel formation (subsequently called ionization-overheating channels). For instance, the

fundamental fluctuation principle is responsible for the experimentally observed irreproducibility of the discharge geometry in different pulses under the same initial conditions.

It can be seen from Eqs. (5) and (7) that ionization overheating effects have a ''threshold'' in terms of p, which is defined by $\nu_c > \omega$ (when $\nu_c < \omega$, the ratio E/E_{cr}, which essentially determines ν_i, ceases to depend on n). Under the experimental conditions this corresponds to the inequality $p > 5$ Torr. Channel formation is only determined experimentally at $p = 20$ Torr.

An ionization-overheating channel should be formed such that it is immediately fairly elongated in the direction of $\mathbf{E}_0$. In this case, the field inside the channel is conserved as the electron concentration n_e increases which incidentally means that the constant field approximation can be used.

An analysis of the experiments reveals that an ionization-overheating channel only develops against a certain initial diffuse background. Formally the n_e threshold for the ionization overheating effect can be obtained, for example, from the first inequality (19). For the conditions in Fig. 4 with $\xi = 1.1$ inequality (19) yields a threshold degree of ionization of the order of 10^{-7} or $n_e \cong 10^{11}\ \mathrm{cm}^{-3}$.

Condition (10) allowing for Eq. (20) gives an ionization-overheating channel size $2a \leqslant V_s/\gamma$, where V_s is the velocity of sound in air, ($a \cong 10^{-2}$ cm), which does not contradict the experimental data. Nevertheless, a sufficiently complete picture of the ionization-overheating stage has not yet been obtained. Various model approaches to its nonlinear stage were analyzed in Refs. 23 and 24. These confirm the explosive nature of the process, i.e., the solution has a singularity within a finite time. The real increase in the electron concentration n_e and gas temperature T, and the drop in the gas density in the ionization-overheating channel are naturally limited. The main factors responsible for this limitation have not yet been determined.

STREAMER STAGE

The existence of an ionization overheating effect significantly alters the discharge pattern. A factor limiting the increase in its transverse dimension appears as the diffuse region with $E > E_{cr}$ undergoes ionization-field displacement along $\mathbf{E}_0$ (plasma head). In the presence of an ionization-overheating channel, the curvature of the boundary of the plasma head perpendicular to the field and ∇n_e at this boundary increase. This leads to an increase in the induced field at this boundary and under the experimental conditions where $p > 50$ Torr, results in the formation of a microwave streamer propagating into the region of the electromagnetic beam where $E_0 \ll E_{cr}$. Figure 5 shows a discharge at 120 Torr, $E_0 = 2.5$ kV/cm ($\Psi \cong 2$), and $t_{pul} = 7\ \mu$s. It can be seen that this is the initial stage of a streamer discharge (compare with Fig. 1). Each of the streamers forming the discharge has the structure of a bright channel which tapers toward the end and is surrounded by a diffuse ''fur coat.''

It has been noted that the main property of streamers is their growth. In Ref. 25 an empirical formula is given for the average growth rate of a microwave streamer

$$V_{st} = 2\sqrt{D_a \nu_a}\sqrt{0.81p} = 3 \times 10^4 \sqrt{p},\ \mathrm{cm/s}. \tag{21}$$

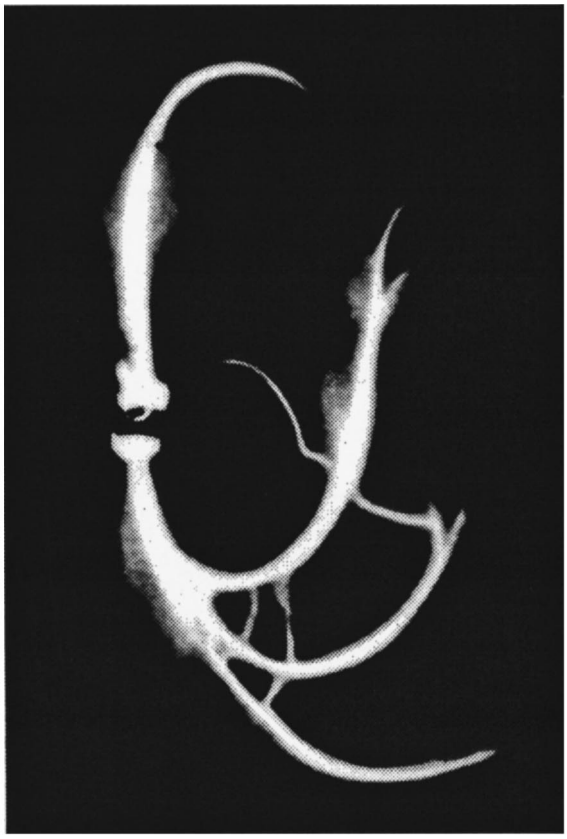


FIG. 5. Microwave discharge at $p = 120$ Torr, $E_0 = 2.5$ kV/cm, and $t_{pul} = 7\ \mu$s initiated by a sphere with $2a = 2.5$ mm.

From this it follows that at the pressures p under study, the velocity V_{st} has a scale of $10^5 - 10^6$ cm/s. This velocity does not depend on E_0 and is proportional to $\sqrt{p}$ (compare Figs. 4 and 5 for the same t_{pul}).

For an advanced streamer discharge (Fig. 1) we can introduce the concept of a front. Assuming that the streamers forming the discharge grow predominantly toward each other and perpendicular to $\mathbf{\Pi}_0$, its velocity is[25]

$$V_{fr} = \frac{V_{st}}{\sqrt{3}}. \tag{22}$$

Note that the ionization conditions differ significantly in different sections of the streamers. For instance, in the diffuse sections at their heads we find $E > E_{st} n_0$, whereas in the incipient ionization-overheating channels we find $n < n_0$ for constant E. Beyond the heads, E decreases to E_0. At the same time, the streamer channel remains highly conducting at least for a few microseconds. Otherwise, the field at the heads would not be sensitive to the structure of already quenched channels. The physics responsible for the lengthening of the plasma decay times in these channel has not yet been clarified.

A microwave streamer is a source of ultraviolet radiation which ionizes the surrounding air.[26] In the halo the photoelectron concentration n_e is of the order of $10^{12}\ \mathrm{cm}^{-3}$ and its decrease with increasing distance from the streamer channel has the characteristic dimension $L_{uv} = 114/p$, cm.

FIG. 6. Initiated microwave discharge at $p=330$ Torr and $E_0=4.4$ kV/cm; $t_{ex}=1.5\ \mu$s, and time after beginning of discharge 10 μs.

RESONANT STAGE

Although the ionization-overheating stage plays an important role in the formation of streamers, a short streamer ($L \ll 1/k$) absorbs no electromagnetic wave energy.[23] However, resonance conditions may be satisfied for comparatively long sections of streamers. This is accompanied by a light flash and the formation of shock waves. An analysis of the experiments reported by Khodataev[23] shows that almost all the electromagnetic wave energy incident on a streamer discharge is absorbed in these sections.

Figure 6 shows a discharge initiated by a sphere at 330 Torr ($E_0=4.4$ kV/cm ($\Psi=3$) and $t_{ex}=1.5\ \mu s \ll t_{pul}$). The photograph was taken 10 μs after the beginning of the microwave pulse. On the right we can see five ''flashed'' resonant sections with an average length of 3 cm, distributed over the discharge front. The initiator, positioned approximately 4 cm to the left of the front and the entire discharge between the initiator and the front do not luminesce at this time. Nonresonant streamer sections are visible near the front.

For the given value of E_0, the electromagnetic beam power is $P_{em} \cong 1.2\times 10^6$ W (the beam geometry is given in Ref. 13). Each section dissipates the power

$$P_{dis}=\frac{(E_0/\sqrt{3})^2 h_{ef}^2}{Z}. \tag{23}$$

Here $\sqrt{3}$ reflects the arbitrary nature of their direction relative to $\mathbf{E}_0$, $h_{ef}=\lambda/\pi$ is the effective length of the resonant vibrator, $Z=R_\Sigma+R_\sigma$ is its impedance, $R_\Sigma=73\ \Omega$ is its radiation resistance, and R_σ characterizes the Ohmic losses in the streamer channel. If $R_\Sigma=R_\sigma$, Eq. (24) gives $P_{dis}=3\times 10^5$ W or the five vibrators, which can be seen in Fig. 5, '''intercept'' the entire power P_{em} to within the accuracy of our estimates. This further confirms the conclusions reached in Ref. 23. Thus, at the discharge front the number of resonant streamer sections is sufficient to completely intercept the electromagnetic beam energy.

The assumed condition $R_\Sigma=R_\sigma$ presupposes a high degree of ionization in the streamer channel. In Ref. 18 n_e is estimated as 10^{16}–10^{17} cm^{-3}. At 330 Torr this gives $\sigma=2\times(10-2)\times 10^2\ (\Omega/\text{cm})$ or from

$$R_\sigma=(1/\sigma)[h_{ef}/(2a)^2]=R_\Sigma$$

the diameter of the highly conducting channel $2a=4-1.3$ mm. It can be seen from Fig. 6 that the experimental value is $2a\cong 1$ mm.

The branching property of microwave streamers is also associated with their resonant sections. At the ends where the induced field has an antinode, the size of the region with $E>E_{E_{cr}}$ increases. This region is naturally filled with diffuse discharge and may serve as the background for the formation of ionization-overheating channels which form the beginnings of new streamers (Fig. 5). We recall that the incipient ionization overheating channel should immediately have a specific length along $\mathbf{E}_0$ which imposes a constraint on the minimum size of the initial diffuse formation parallel to $\mathbf{E}_0$. Its entire length participates in the growth of the already formed channel and the diffuse ''precursor'' in front of the end of the growing ionization-overheating channel may be unextended.

The newly formed ionization-overheating channel may have contact with the initial channel, which resembles branching on an integrated photograph (Fig. 5) or it may not have contact.[27] In a structurally advanced discharge resonant sections may appear in the central part of the streamers, not adjacent to the ends, since the streamer system is a multimode electrodynamic system.

We shall estimate the time intervals characterizing the discharge. The quasistatic changes in its structure have a geometric scale no greater than 0.1λ. They are filled with plasma at the velocity V_{st}, i.e., under the experimental conditions the time taken for these changes is $0.1\lambda/V_{st}\cong 10^6$ s. The electrodynamic situation changes an order of magnitude more slowly — $\lambda/V_{st}\cong 10^{-5}$ s. The energy ''pumping'' of the resonant sections takes place over a few hundred periods of the microwave field, i.e., has a scale of 10^{-7} s. Finally, after this pumping the release of energy into the exploding plasma filament in the gasdynamic sense takes place over a few microseconds.[23]

PLASMA DECAY

In a subcritical discharge the active energy processes only take place at its front. Behind this the plasma begins to decay.

Figure 7 shows an oscilloscope trace of the signal used to probe a discharge at 120 Torr with $t_{pul}=40\ \mu$s: trace *1* is the level in the absence of any probe signal and on trace *2* the characteristic induction denotes the beginning and end of the microwave pulse. Its initial horizontal section indicates the signal level in the absence of a discharge. The descending leading edge is caused by the discharge gradually cover-

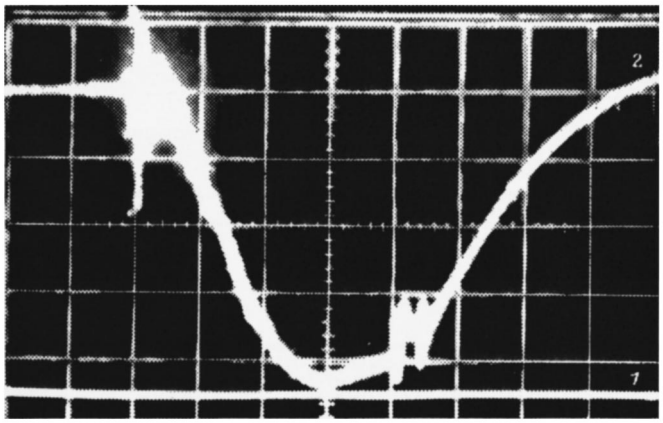


FIG. 7. Oscilloscope trace of $\lambda = 1$ cm radiation used to probe a microwave discharge at $p = 120$ Torr, $E_0 = 4$ kV / cm, and $t_{\text{pul}} = 40\ \mu$s.

ing the probe beam. This process lasts for approximately $30\ \mu$s which, from Eqs. (21) and (22) corresponds to $V_{\text{fr}} \cong 10^5$ cm / s. The maximum attenuation observed at the instant of complete overlap corresponds to $\overline{n_e} \cong 5 \times 10^{12}\ \text{cm}^{-3}$. It should be noted that this value is arbitrary because the discharge is initially extremely nonuniform (for the calculations of $\overline{n_e}$ the dimension of the discharge along the probing line was determined from a suitable photograph and the model of a planar nonreflecting plasma layer was used). During the last $10\ \mu$s of the microwave pulse the attenuation is reduced. We can postulate that this is because the plasma has started to decay beyond the discharge front. The scale of this process is around $100\ \mu$s. After the end of the microwave pulse, the decay is accelerated but still lasts for tens of microseconds. Such a long time may either be attributed to the thermal nature of the ionization in the streamers[28] or to the plasma-chemical processes in the plasma.[29]

CONCLUSIONS

We have seen that a method of using a microwave discharge to purify the Earth's atmosphere from Freons which are damaging to the ozone layer, may involve exposing a particular region of the troposphere to electromagnetic radiation and spatial scanning with a microwave beam. In order to reduce the energy cost of destroying the Freon molecules, the irradiation should be carried out at an electromagnetic wave intensity lower than the threshold for the breakdown of air over the entire beam and a discharge should be initiated in the region being purified, for example by using resonant vibrators. After being formed at the initiators, this discharge propagates in the opposite direction to the radiation in the form of growing and branching streamers. By scanning the beam, we can control its front and sustain continuous burning in a given volume.[30,31]

Different sections of this subcritical discharge must be considered from the point of view of electron generation. The first sections are the streamer heads where the processes take place at a level of induced field higher than critical and appreciably higher than the initial level. The second sections are the streamer channels adjacent to the heads. These have the structure of a thin plasma filament with a reduced molecular concentration surrounded by a diffuse sheath. The field in the channels is almost the same as the initial value. These are followed by the resonant streamer sections whose system positioned at the discharge front dissipates almost all the electromagnetic energy incident on the discharge. The sections between the streamers are filled with plasma, excited by ionizing ultraviolet radiation from the streamers. Finally, there is the region behind the discharge front which is filled with decaying plasma.

At the present time, we can estimate many characteristics of a subcritical microwave discharge, such as the average growth rate of its constituent streamers, the propagation velocity of the discharge front, the number of resonant streamers at this front, and so on. Nevertheless, its investigation can by no means be described as complete. For instance, the question of the highest possible subcriticality and the wavelength range of the field in which it is achieved has not been resolved. Direct experiments to study the destruction of Freons in this type of discharge as a function of the degree of subcriticality are also required and the energy cost of the process needs to be determined.

[1] V. M. Smirnov, *The Earth's Atmosphere and Energetics*, Novoe o Nauke, Zhizni i Tekhnike. Ser. Fizika. No. 3. [in Russian] Znanie, Moscow (1979), 64 pp.
[2] K. T. Tsand, D. D.-M. Ho, F. Y. Wong *et al.*, *ibid.*, pp. 142–157.
[3] G. A. Askaryan, G. M. Batanov, A. É. Barkhudarov *et al.*, Tr. Inst. Obshch. Fiz. Ross. Akad. Nauk **47**, 23 (1994).
[4] G. A. Askaryan, G. M. Batanov, D. F. Bykov *et al.*, Tr. Inst. Obshch. Fiz. Ross. Akad. Nauk **47**, 9 (1994).
[5] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **55**, 389 (1985) [Sov. Phys. Tech. Phys. **30**, 224 (1985)].
[6] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **64**(1), 74 (1994) [Tech. Phys. **39**, 40 (1994)].
[7] G. M. Batanov, S. I. Gritsinin, I. A. Kossyĭ *et al.*, Tr. Fiz. Inst. Akad. Nauk SSSR **60**, 170 (1985).
[8] V. G. Brovkin, Yu. F. Kolesnichenko, and D. V. Khmara, *Ball Lightning in the Laboratory* [in Russian], Khimiya, Moscow (1994), pp. 110–136.
[9] N. D. Borisov, A. V. Gurevich, G. M. Milikh, *An Artificial Ionized Region in the Atmosphere* [in Russian], Institute of Terrestrial Magnetism, the Ionosphere, and Radio Wave Propagation, Academy of Sciences of the USSR, Moscow, (1985), 184 pp.
[10] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **65**(7), 60 (1995) [Tech. Phys. **40**, 666 (1995)].
[11] L. P. Grachev, I. I. Esakov, K. V. Khodataev *et al.*, MRTI Preprint No. 9005 [in Russian], Moscow Radio Electronics Institute, Academy of Sciences of the USSR, Moscow (1990), 14 pp.
[12] L. P. Grachev, I. I. Esakov, K. V. Khodataev *et al.*, Fiz. Plazmy **18**, 411 (1992) [Sov. J. Plasma Phys. **18**, 216 (1992)].
[13] A. D. MacDonald, *Microwave Breakdown in Gases* (Wiley, New York, 1966; Mir, Moscow, 1969, 205 pp.).
[14] Yu. P. Raĭzer, *Physics of Gas Discharges* [in Russian], Nauka, Moscow (1987), 591 pp.
[15] Yu. A. Lupan, Zh. Tekh. Fiz. **46**, 2321 (1976) [Sov. Phys. Tech. Phys. **21**, 1370 (1976)].
[16] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1960; Nauka, Moscow, 1982, 531 pp.).
[17] V. B. Gil'denburg, I. S. Gushchin, S. A. Dvinin *et al.*, Zh. Eksp. Teor. Fiz. **97**, 1151 (1990) [Sov. Phys. JETP **70**, 645 (1990)].
[18] S. I. Gritsinin, A. A. Dorofeyuk, I. A. Kossyĭ *et al.*, Teplofiz. Vys. Temp. **25**, 1068 (1987).
[19] V. G. Brovkin, D. F. Bykov, S. K. Golubev *et al.*, Zh. Tekh. Fiz. **61**(2), 153 (1991) [Sov. Phys. Tech. Phys. **36**, 213 (1991)].
[20] G. T. Markov and D. M. Sazonov, *Antennae* [in Russian], Énergiya, Moscow (1975), 529 pp.

[21] V. B. Gil'denburg and A. V. Kim, Fiz. Plazmy **6**, 904 (1980) [Sov. J. Plasma Phys. **6**, 496 (1980)].
[22] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **66**(8), 73 (1996) [Tech. Phys. **41**, 785 (1996)].
[23] K. V. Khodataev, Khim. Fiz. **12**(3), 303 (1993).
[24] A. V. Kim and G. M. Fraĭman, Fiz. Plazmy **9**, 613 (1983) [Sov. J. Plasma Phys. **9**, 358 (1983)].
[25] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **65**(5), 21 (1995) [Tech. Phys. **40**, 416 (1995)].
[26] N. A. Bogatov, S. V. Golubev, and V. G. Zorin, Pis'ma Zh. Tekh. Fiz. **9**, 888 (1983) [Sov. Tech. Phys. Lett. **9**, 382 (1983)].
[27] A. L. Vikharev, A. M. Gorbachev, A. V. Kim *et al.*, Fiz. Plazmy **18**, 1064 (1992) [Sov. J. Plasma Phys. **18**, 554 (1992)].
[28] L. P. Grachev, I. I. Esakov, G. I. Mishin *et al.*, Zh. Tekh. Fiz. **59**, 149 (1989) [Sov. Phys. Tech. Phys. **34**, 1181 (1989)].
[29] P. V. Vedenin, L. P. Grachev, I. I. Esakov *et al.*, in *Abstracts of the Second International Symposium on Theoretical and Applied Plasma Chemistry (ISTAPC) (IGKhTA*, Ivanovo, 1995), p. 23.
[30] G. I. Batskikh and K. V. Khodataev, in *Abstracts of the Third International Workshop*, Institute of Applied Physics, Russian Academy of Sciences, Nizhniĭ Novgorod, Russia, 1996, Abstract D-5.
[31] G. I. Batskikh and K. V. Khodataev, in *Abstracts of the Second Practical Conference on Sustainable Development: System Analysis in Ecology*, Sevastopol, Ukraine (Ukrainian Ministry of Education, 1996), p. 60.

Translated by R. M. Durham

Azimuthal surface waves in a magnetized plasma

I. A. Girka and P. K. Kovtun

Kharkov State University, 310077 Kharkov, Ukraine
(Submitted April 7, 1997)
Zh. Tekh. Fiz. **68**, 25–28 (December 1998)

A theoretical analysis is made of the dispersion properties of surface waves propagating along the azimuth in magnetized cylindrical plasma waveguides. It is shown that surface oscillations of the ion component may propagate in these waveguides. © *1998 American Institute of Physics.* [S1063-7842(98)00412-7]

INTRODUCTION

Extraordinary electromagnetic waves propagating azimuthally along the plasma boundary in cylindrical waveguides have been termed azimuthal surface waves (ASWs).[1–4] These waves have recently been actively studied mainly because of their possible use for the development of plasma[5] and semiconductor electronics devices.[4] These studies considered the case of a dense plasma ($\omega_{pe}^2 \gg \omega_{ce}^2$, where $\omega_{p\alpha}$ and $\omega_{c\alpha}$ are the plasma and cyclotron frequencies of α-species particles where $\alpha=e$ for electrons and $\alpha=i$ for ions) for which the frequency of the oscillations studied lies above the electron cyclotron frequency, so that the motion of the plasma ion component can be neglected. This relationship between the plasma and cyclotron frequencies can easily be satisfied in a semiconductor plasma which typically has a high concentration n_e of free electrons. However, in a laboratory gas plasma this relationship is frequently not satisfied (a concentration of the order $n_e \sim 10^{15}\,\mathrm{cm}^{-3}$ is achieved in the best controlled thermonuclear fusion experiments, but the use of a strong magnetic field for plasma confinement yields the inequality $\omega_{pe}^2 \lesssim \omega_{ce}^2$). Here we investigate the dispersion properties of ASWs in magnetized plasma waveguides with $\omega_{pe}^2 \lesssim \omega_{ce}^2$ and we show that surface oscillations caused by motion of the ion component may exist in these waveguides.

FORMULATION OF THE PROBLEM

We shall investigate the propagation of electromagnetic waves near the boundary of a homogeneous plasma cylinder of radius a separated from a coaxial ideally-conducting metal chamber of radius b by a layer of dielectric with the permittivity ε. We shall assume that the waveguide is homogeneous along the axis $\partial/\partial z \equiv 0$ and an external static magnetic field is oriented along the waveguide axis $\mathbf{B}_0 \| \mathbf{e}_z$. We assume that the components of the electric induction and the electric field strength vectors of the **E** wave are related by the permittivity tensor of a cold weakly collisional plasma.

In this case, the system of Maxwell equations is broken down into two independent systems which describe an E-wave with the components E_r, E_φ, and H_z, and an H-wave with the components H_r, H_φ, and E_z. The properties of the H-wave do not depend on the external magnetic field. A surface H-wave does not propagate[2] in waveguides with a narrow dielectric gap ($b-a \ll a$, $b-a \ll c/\omega$). Thus, we shall confine our analysis to E-type waves. Since ASWs with different azimuthal mode numbers m propagate independently, on the basis of the symmetry of the problem, we shall seek a solution of the Maxwell equations for the axial component of the ASW magnetic field in the form $H_z = H_z^{(m)}(r)\exp(im\varphi - i\omega t)$.

The components of the ASW electric field are expressed in terms of H_z as follows:

$$E_r = \frac{k}{k_0^2}\left(\frac{\varepsilon_2}{\varepsilon_1}\frac{\partial H_z}{\partial r} + \frac{m}{r}H_z\right), \tag{1}$$

$$E_\varphi = i\frac{k}{k_0^2}\left(\frac{\partial H_z}{\partial r} + \frac{m}{r}\frac{\varepsilon_2}{\varepsilon_1}H_z\right), \tag{2}$$

where $k=\omega/c$, k_0^{-1} determines the depth of penetration of the field in the plasma, $k_0^2 = (\omega/c)^2(\varepsilon_2^2-\varepsilon_1^2)/\varepsilon_1$, and $\varepsilon_{1,2}$ are the components of the plasma permittivity tensor.

DERIVATION OF THE DISPERSION EQUATION

We impose the constraint that the following boundary conditions are satisfied for the components of the ASW field: the tangential component of the electric field of the wave is zero at the metal surface

$$E_\varphi(r=b)=0; \tag{3}$$

the tangential components of the electric and magnetic fields of the wave are continuous at the plasma–vacuum interface

$$E_\varphi(r=a+0)=E_\varphi(r=a-0), \tag{4}$$

$$H_z(r=a+0)=H_z(r=a-0); \tag{5}$$

and the wave fields are bounded in the entire waveguide volume and particularly on the axis

$$H_z(r=0)<\infty. \tag{6}$$

From the Maxwell equations we obtain a Bessel equation for the amplitude of the magnetic field $H_z^{(m)}$ in the plasma region, whose solution which satisfies condition (6) is expressed in terms of a modified mth order Bessel function $I_m(k_0 r)$

$$H_z^{(m)} = C_1 \cdot I_m(k_0 r), \tag{7}$$

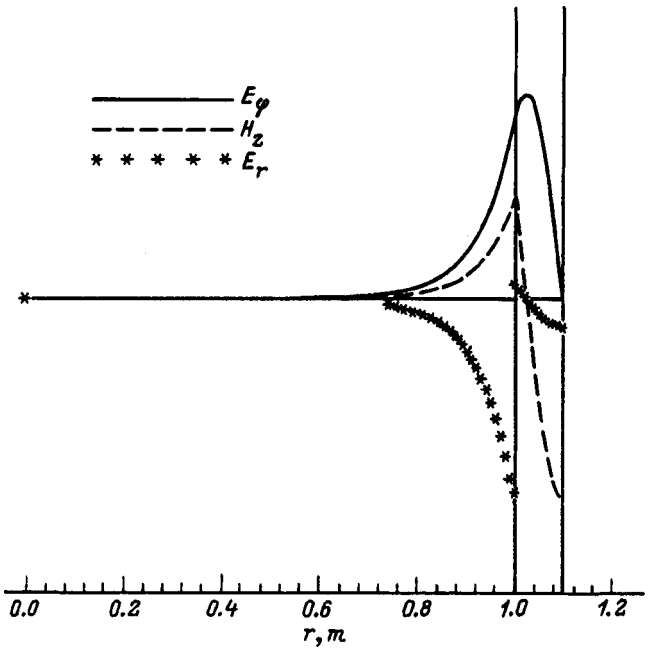


FIG. 1.

where C_1 is a normalization factor. The solution (7) has the form of a surface wave when $k_0^2>0$.

From this it follows that ASWs may exist in the following frequency ranges:

$$\omega_{lh}<\omega<\omega_1 \quad \text{and} \quad \omega_{hh}<\omega<\omega_2, \tag{8}$$

which we shall call low-frequency and high-frequency, respectively. Here ω_{lh} and ω_{hh} are the lower and upper hybrid frequencies: $\omega_{1,2}=\pm 0.5\omega_{ce}+\sqrt{0.25\omega_{ce}^2-\omega_{ci}\omega_{ce}+\omega_{pe}^2}$ are the cutoff frequencies of the bulk waves.[6] In a rarefied plasma with $\omega_{pe}<|\omega_{ce}|$ the high-frequency range becomes very narrow and is close to the electron cyclotron frequency and the ASW frequencies in this range depend weakly on the plasma parameters. In this case, the lower hybrid frequency is of the order of the ion cyclotron frequency and conditions are created for the existence of surface oscillations of the plasma ion component.

In the dielectric region, the magnetic field of the wave is expressed in terms of a linear combination of mth-order Bessel $J_m(\kappa r)$ and Neumann $N_m(\kappa r)$ functions, i.e., the ASWs are only surface waves in the plasma and are bulk waves in the dielectric

$$H_z^{(m)}=C_2\,(J_m(\kappa r)N_m'(\kappa b)+J_m'(\kappa b)N_m(\kappa r)), \tag{9}$$

where $\kappa(\omega/c)\sqrt{\varepsilon}$ and C_2 is a normalization factor.

Figure 1 gives the ASW fields as a function of the waveguide radius. The following values of the waveguide parameters were used for the calculations: $n_e=2\times 10^{12}\,\mathrm{cm}^{-3}$, $B_0=50\,\mathrm{kG}$, $a=1\,\mathrm{m}$, $b=1.1\,\mathrm{m}$, $m=-3$.

We stress that the azimuthal component of the electric field is nonzero in the bulk of the waveguide and becomes zero at the metal chamber. In this respect the field topology in a cylindrical plasma differs from the propagation of surface waves across a static magnetic field parallel to a planar plasma boundary.[3] For the selected values of the waveguide parameters, the ASWs are essentially nonpotential: the electromagnetic energy stored in the magnetic component is not small compared with the electric field energy of the wave.

Having imposed the constraint that the boundary conditions (4) and (5) are satisfied, we obtain the ASW dispersion equation in the form

$$D^{(\mathrm{pl})}=D^{(\mathrm{vac})}, \tag{10}$$

where

$$D^{(\mathrm{pl})}=\frac{1}{k_0}\frac{I_m(k_0a)}{I_m(k_0a)}+\frac{m\cdot\varepsilon_2}{k_0^2a\cdot\varepsilon_1}, \tag{11}$$

$$D^{(\mathrm{vac})}=\frac{1}{\kappa}\frac{J_m'(\kappa a)\,N_m'(\kappa b)-J_m'(\kappa b)\,N_m'(\kappa a)}{J_m'(\kappa b)\,N_m(\kappa a)-J_m(\kappa a)\,N_m(\kappa b)}. \tag{12}$$

ANALYSIS OF THE DISPERSION EQUATION

For a narrow dispersion gap ($m\Lambda\ll 1$, $\kappa(b-a)\ll|m|$, where $\Lambda\equiv(b-a)/a$), when the influence of the plasma parameters on the dispersion properties of the ASWs is strongest, the right-hand side of the dispersion equation is simplified

$$D^{(\mathrm{vac})}\approx(b-a)\left(\frac{m^2}{\kappa a^2}-1\right). \tag{13}$$

From this it can be seen that ASWs do not propagate in broad waveguides with a narrow gap since the left-hand side of Eq. (10) is positive because of the inequality ($k_0a\gg|m|$), and the right-hand side is negative when $\omega_{lh}>(|m|c)/(a\sqrt{\varepsilon})$ so that the dispersion equation has no solution (see also the numerical calculations plotted in Fig. 5).

A simple analytic solution of Eq. (10) for ASWs in the low-frequency range may be obtained for the case of a thin cylinder $\kappa b\ll 1$, $k_0a\ll 1$

$$\frac{\omega}{\omega_{ci}}\approx 1+\frac{|\gamma|}{|\gamma|+\varepsilon}\frac{\omega_{pi}^2}{\omega_{ci}^2}, \tag{14}$$

where

$$\gamma=\frac{(a/b)^{2m}-1}{(a/b)^{2m}+1}.$$

For a narrow gap we have $\gamma\approx m\cdot\Lambda\ll 1$. In this case, as can be seen from Eq. (14) the natural frequency of the waveguide is determined not by the absolute width of the dielectric gap $(b-a)$ but by the relative width Λ. In this case, the waves are unidirectional: the sign of the azimuthal wave number can only be negative ($m<0$). The direction of propagation of the ASWs is the same as the direction of ion cyclotron rotation and their frequency may be of the order of the ion cyclotron frequency. The ASWs propagate in a thin cylinder (the dispersion equation (10) has a solution in the form of a surface wave $\omega>\omega_{lh}$) if the plasma density is not too low or the external magnetic field is not too strong $m_i/m_e>\omega_{pi}^2/\omega_{ci}^2\gtrsim\{1+(\varepsilon/|\gamma|)^2\}$.

A necessary condition for the existence of ASWs near a dense plasma–metal interface ($b=a$) is the presence of an

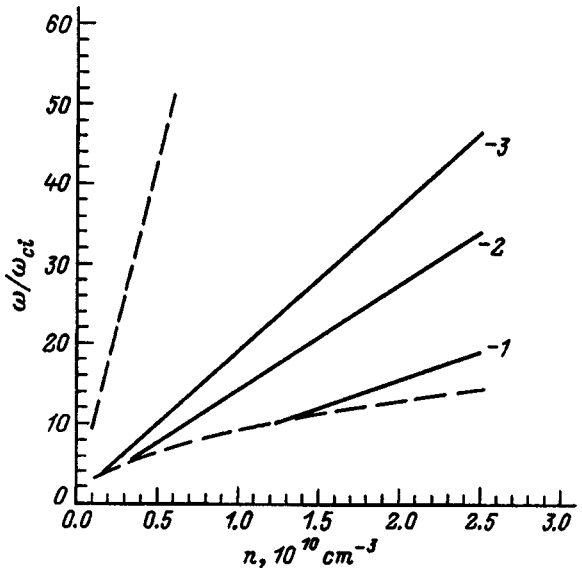


FIG. 2.

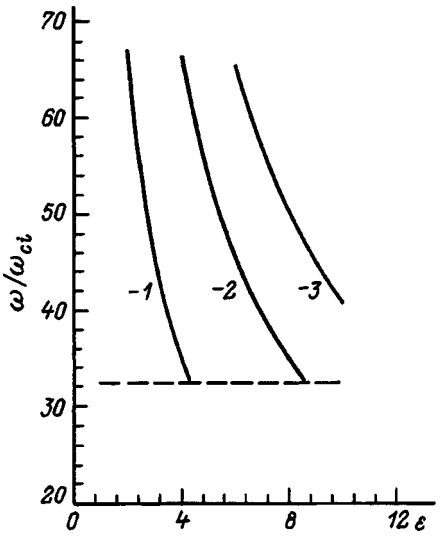


FIG. 4.

external magnetic field,[3] although for a thin magnetized plasma cylinder the dispersion equation has no solution.

For a thin plasma cylinder ($k_0 a \ll |m|$) the ASWs may be considered to be potential and the magnetic field of the wave is much smaller than the electric field. The amplitude of the electric fields of ASWs with the azimuthal number $m=-1$ remains almost constant even in the plasma region.

The results of a numerical solution of the dispersion equation (10) are plotted in Figs. 2–4. The ASW frequency is plotted as a function of the parameters of the waveguiding structure: Fig. 2 gives the density Fig. 3 the external magnetic field B_0, and Fig. 4 the permittivity of the dielectric gap ε. The following values of the waveguide parameters were taken: for all the figures $a=10$ cm, $\Lambda=0.1$, for Fig. 2 $B_0=3000$ G, and the density is normalized to $10^{10}\,\mathrm{cm}^{-3}$, for Figs. 3 and 4 $n_e=10^8\,\mathrm{cm}^{-3}$, and for Fig. 4 $B_0=50$ G. The numbers on the figures indicate the mode number m and the dashed lines denote the boundaries of the low-frequency range.

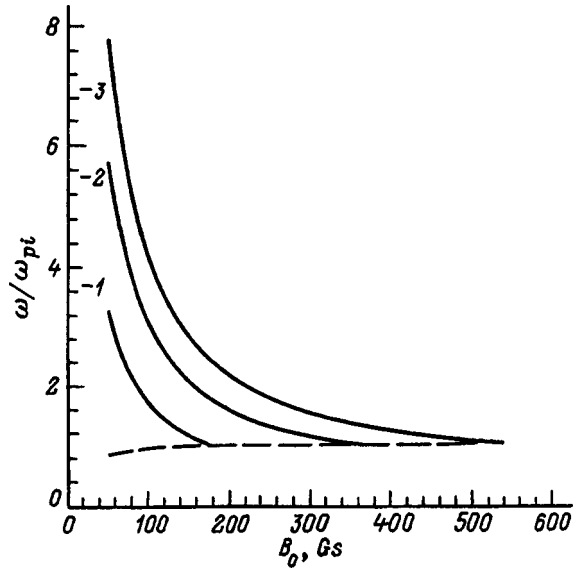


FIG. 3.

In waveguides with a wide dielectric gap the ASW properties depend weakly on the plasma parameters and are mainly determined by the properties of the gap. For a given azimuthal number it is possible for higher radial harmonics of the ASWs to exist and a solution of Eq. (10) exists for

$$(2n+1)\frac{\pi}{2}<\frac{\omega}{c}(b-a)<(2n+2)\frac{\pi}{2},$$

$$n=1,2,3\ldots, \quad \kappa a \gg |m|.$$

In this case, the ASWs are essentially nonpotential. In wide waveguides inserted in a strong magnetic field $k_0 a \gg |m|$, the depth of penetration of the field in the plasma is small and the analytic expression for the frequency of the nth radial mode has the form $\omega=\omega_{0n}+\delta\omega$, where

$$\omega_{0n}=\frac{\pi n}{2}\frac{c}{b-a},$$

$$\delta\omega=\frac{k_0}{\kappa}\frac{c}{b-a}\left(1-\frac{m\varepsilon_2}{k_0 a\varepsilon_1}\right)_{\omega=\omega_{0n}}. \tag{15}$$

For waves with a small depth of penetration the frequency dependence on m is weaker and decreases even more strongly as the plasma density increases. It can be seen from Eq. (15) that $\delta\omega$ for positive m is greater than that for negative m and the difference in the frequency of waves with opposite azimuthal numbers is greatest near the upper boundary of the low-frequency range where $\varepsilon_1+\varepsilon_2=0$.

The results of a numerical solution of the dispersion equation for a wide dielectric gap are plotted in Fig. 5. The following values of the parameters were used: $a=1$ m, $\Lambda=0.1$, $B_0=50$ kG, and $m=3$.

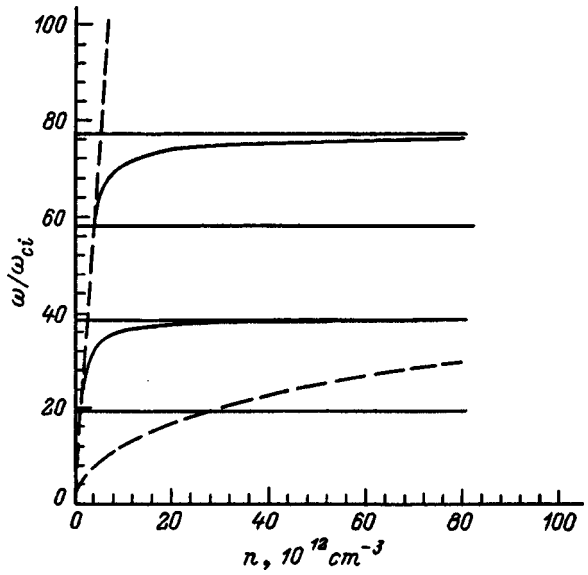


FIG. 5.

CONCLUSIONS

To sum up, we have investigated the dispersion properties of extraordinary electromagnetic oscillations propagating along the azimuth near the boundary of a magnetized plasma partially filling a cylindrical waveguide. We studied the field distribution (see Eqs. (7) and (9) and Fig. 1) and the dependence of the ASW natural frequency (see Eqs. (14) and (15) and Figs. 2–5) on the parameters of the waveguiding structure.

As an example we have reported calculations of the natural frequency of an ASW with $m=-3$ propagating in a completely ionized helium plasma with $a=10\,\mathrm{cm}$, $B_0=3000\,\mathrm{G}$, $T=1\,\mathrm{eV}$, $n_e=10^{10}\,\mathrm{cm}^{-3}$, and $\Lambda=0.1$. For this case the ASW frequency is $\omega=19\omega_{ci}=140\times10^6\,\mathrm{s}^{-1}$, which is much higher than the collision frequency $\nu\approx30n_eT^{-3/2}\,\mathrm{s}^{-1}$ in a cold plasma.

These results may find applications in the design of plasma electronics devices using a magnetized plasma as an element of the waveguiding structure (see Refs. 7 and 8, for example).

The authors are grateful to V. A. Girka for useful discussions of the results.

This work was supported by the Scientific Physicotechnological Center of the National Academy of Sciences and the Ministry of Education of the Ukraine.

[1] V. A. Girka, I. A. Girka, A. N. Kondratenko, and V. I. Tkachenko, Radiotekh. Elektron. **33**, 1031 (1988).
[2] V. A. Girka and I. A. Girka, Radiotekh. Elektron. **36**, 1997 (1991).
[3] N. A. Azarenkov, A. N. Kondratenko, and K. N. Ostrikov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **36**, 335 (1993).
[4] V. A. Girka, I. A. Girka, and V. I. Tkachenko, Zh. Tekh. Fiz. **66**(4), 114 (1996) [Tech. Phys. **41**, 357 (1996)].
[5] V. A. Girka, I. A. Girka, V. P. Olefir, and V. I. Tkachenko, Pis'ma Zh. Tekh. Fiz. **17**(1), 87 (1991) [Tech. Phys. Lett. **17**, 35 (1991)].
[6] N. A. Krall and A. W. Trivelpiece, *Principles of Plasma Physics* (McGraw-Hill, New York, 1973; Mir, Moscow, 1975).
[7] N. I. Karbushev, in *Book of Abstracts of the 11th International Conference on High Power Particle Beams*, Prague, 1996, pp. 1–55.
[8] I. N. Onishchenko, V. A. Balakirev, A. M. Korostelev *et al.*, *ibid.* pp. 1–61.

Translated by R. M. Durham

Spherical microwave electrode discharge. Phenomenology and results of probe measurements

L. Bardosh

Uppsala University, Uppsala, Sweden,

Yu. A. Lebedev

Institute of Petrochemical Synthesis, Russian Academy of Sciences, 117912 Moscow, Russia
(Submitted July 8, 1997)
Zh. Tekh. Fiz. **68**, 29–33 (December 1998)

A description is given of an initiated microwave ball discharge generated at the end of an antenna in a metal discharge chamber whose dimensions are substantially greater than those of the glowing plasma zone. The experiments were carried in hydrogen at pressures of 1–15 Torr. The double electric probe method was used to investigate the plasma structure. The Boltzmann equation was used to estimate the microwave field strengths corresponding to the measured values of the electron temperature near the glowing region of the discharge. It is shown that the spatial structure of the field corresponds to the field structure of a surface wave propagating along the discharge surface. The existence of such a wave can explain the increase in the plasma radiation intensity at its boundary. © *1998 American Institute of Physics.* [S1063-7842(98)00512-1]

INTRODUCTION

Initiated microwave discharges, i.e., discharges which only exist in the presence of an initiator at powers below those needed to sustain the discharge without such an initiator, have recently attracted considerable attention among researchers.[1,2] These initiators can be external sources of ionization, objects inserted in the discharge chamber, dust particles, and so on. These discharges possess various interesting qualities: they can exist at anomalously low powers, they can create a nonequilibrium plasma at pressures up to atmospheric, and they provide the possibility of effectively controlling the plasma parameters, and so on. At present this is the least studied type of microwave discharge. No studies have been made of the physical processes leading to its formation, its structure, and parameters. In some cases, it has been assumed that when a passive initiator is used, this is a self-sustained discharge surrounded by a region of nonself-sustained discharge.[1]

Despite the lack of sufficient information on the nature of these discharges, their extensive possibilities have already been demonstrated in plasma chemistry, for example.[3,4] The properties of an initiated discharge are clearly observed in the results of Ref. 5 which demonstrated high conversion of methane into acetylene in a microwave discharge in a discharge chamber containing nickel plates.

Here we report results of an investigation of the parameters of a microwave discharge generated at reduced pressures at the ends of a metal antenna.

EXPERIMENTAL SETUP

The apparatus is shown schematically in Fig. 1. The discharge chamber was a metal cylinder of ~ 10 cm radius, excited from the end of a pin antenna comprising a stainless steel tube. The 6 mm diameter tube was an element of a coaxial waveguide junction which could be tuned by means of short-circuiting plungers. The dimensions of the discharge chamber were not specially selected. The discharge system was similar to that described in Ref. 6 except that the distance between the lower end of the chamber and the antenna was more than 15 cm and was kept constant. Hence, the discharge was generated at the end of an isolated antenna.

We used a microwave generator with an output power up to 2.5 kW and frequency 2.45 GHz. The absorbed power was determined using a reflectometer by subtracting the reflected power from the incident power. Straightaway we note that the power thus calculated cannot be defined *a priori* as that absorbed by the plasma, since losses may occur in the energy transmission elements in the chamber and at the walls.[7]

We used a double electric probe to measure the parameters of the plasma electron component. The measurement method for a microwave plasma was described in detail in Ref. 8. Tungsten probes of 100 μm diameter in quartz capillaries with an uninsulated section 1 mm long and spaced 3 mm apart were inserted through a port in the side of the chamber and could be moved along the radius of the discharge. Each probe was inserted in a metal screen. The measurements were made in the plane 3 mm from the exciting electrode and perpendicular to its axis.

We shall now make some observations on the role of the electrode in the microwave plasma. A fairly detailed study of this problem has been made for rf discharges (see, for example, Ref. 9). As the frequency of the field increases, the thickness of the space charge layer at the electrode decreases, as does the energy of the ions bombarding the electrode. In the microwave range the conductivity of the layer is low and

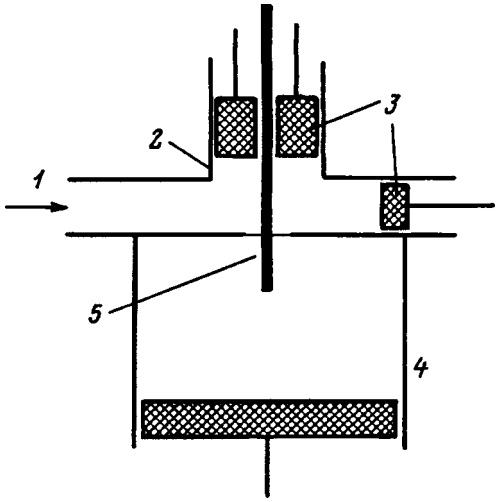


FIG. 1. Experimental setup: *1* — unit for coupling electromagnetic energy into waveguide, *2* — coaxial waveguide junction, *3* — short-circuiting plungers, *4* — discharge chamber, and *5* — antenna.

continuity of the current is provided by displacement currents through the layer. In this case, the electrode does not play any role in generating the charged particles (γ-processes caused by secondary electron emission from the electrodes are unimportant) and the discharge exists in the α form (volume ionization processes play a role). This is confirmed by experimental data showing the lack of electrode erosion in a microwave plasma in a hydrogen–methane mixture.[6] The electrode evidently only determines the structure of the electromagnetic field and in particular, generates a nonuniform microwave field with increasing strength near its surface.

Here we present results obtained for a hydrogen discharge.

PHENOMENOLOGY OF THE DISCHARGE

The main features of the discharge may be described as follows.

The spherical discharge forms at the end of the exciting pin, partially covering it.

A discharge may form at the point where the exciting pin enters the chamber and may be shifted toward its end by moving short-circuiting plungers in the waveguide channel at constant incident power. When the incident power was increased, the discharge moved back to the point where the pin enters the chamber and could only be displaced toward its end by altering the position of the tuning plunger. Thus, the microwave discharge and the chamber (including the external tuning elements) are a strongly coupled microwave system.

At low incident powers (insufficient to excite the discharge) a discharge could be initiated at the end of the pin using a Tesla coil.

The diameter of the discharge increases with decreasing pressure and increasing incident power.

The discharge glow is inhomogeneous: the exciting pin at the center of the discharge is covered with a bright glowing film. With increasing distance from the axis, the plasma radiation intensity decreases and then increases toward the edge of the spherical formation. The discharge boundary is clearly defined. The glow along the surface of the discharge sphere is inhomogeneous, being less bright near the antenna surface on the side of the microwave generator and brighter on the opposite side.

A dumbbell type of discharge structure could be obtained, comprising two spheres connected by a faintly glowing neck. One sphere was located at the end of the exciting pin and the other was situated on the pin at a distance of approximately $\lambda/4$ ($\lambda \approx 12.2$ cm is the wavelength of the electromagnetic radiation in free space). A distinguishing feature from the frequently observed standing-wave structure was that the plasma formation on the pin was a sphere of the same size and had exactly the same luminescence structure as the first (a brightly glowing film on the surface of the pin at the center of the sphere and a sharply defined boundary on the side where energy is supplied along the pin).

A probe positioned near the discharge may alter the structure of the electromagnetic field and in many cases, probe measurements could not be made because any displacement of the probe even outside the growing region led to changes in the discharge glow and even quenched it.

Under conditions when displacement of the probes outside the discharge caused no visible changes, the discharge behaved as a liquid droplet toward the probes. When the probes came in contact with the glowing zone, the discharge began to move away from them and collapsed inward. When the glowing zone was pierced by the probes, the plasma wrapped round them and as they moved outward, the plasma was dragged behind them, as if on a wettable surface. At a certain distance from the center, the plasma became detached from the probes and assumed a spherical shape, although the dimensions were not established immediately but after a few oscillations (a few seconds).

Two types of discharge exist, depending on the electrodynamic characteristics of the system used to transfer microwave energy into the gas. One type occurs with high reflection coefficients in the microwave channel, exists only for one position of the matching plungers, pressure, and power, and is characterized by low energy inputs to the plasma. The other type has lower levels of reflection from the discharge section and the energy input to the plasma can be varied whilst keeping the system geometry constant.

RESULTS OF PROBE MEASUREMENTS

The charged particle concentration can be estimated from the saturation ion current using the well-known relation

$$I_i = 0.61 N_i e S_p v_i, \tag{1}$$

where N_i is the charged particle concentration, e is the electron charge, $S_p = \pi dl = 4.7 \times 10^{-7}\ \mathrm{m}^2$, $v_i \approx 10^4 (T_e/A)^{1/2}$ is the average ion velocity in the plasma allowing for their acceleration in the prelayer, and A is the ion mass number.

The values of the saturation ion current plotted in Fig. 2 correspond to probe voltages of 50 V and are averages over the values calculated using both branches of the two-probe characteristic. In some cases, we used a single-probe system

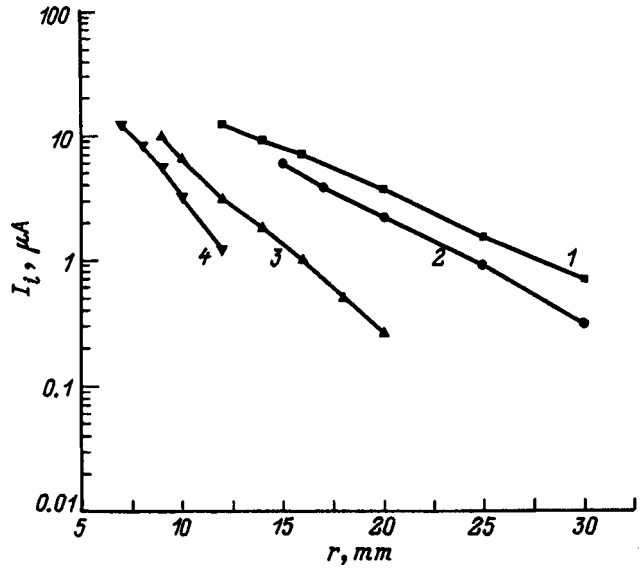


FIG. 2. Saturation ion current of probe as a function of distance from the center of the discharge at pressures of 1 (*1*), 3 (*2*), 8 (*3*), and 15 Torr (*4*).

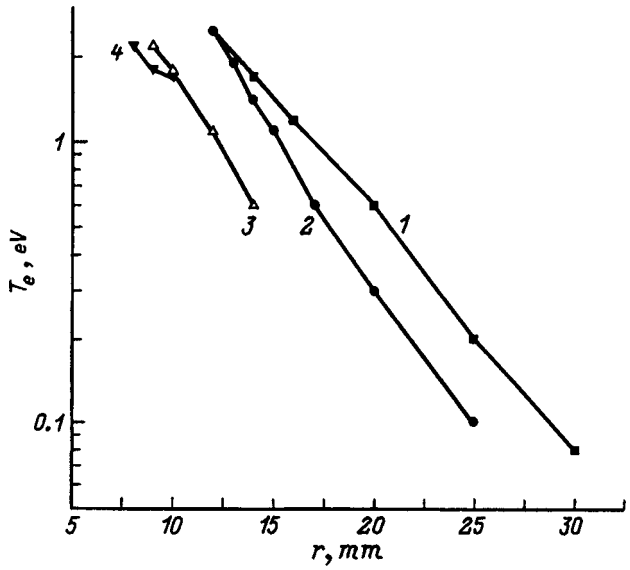


FIG. 3. Electron temperature as a function of distance from the center of the discharge (*1–4* — as in Fig. 2).

in which the counterprobe was the metal discharge chamber. The error in the determination of the ion current was ~10%. In calculations of the charged particle concentration it is important to know the plasma ion composition. In a hydrogen plasma with degrees of dissociation less than 50%, the H_3^+ ion formed as a result of ion conversion[10] predominates, and for order-of-magnitude estimates of the charged particle concentration ($n_e \approx n_i$) from the saturation ion current we can use $N_i[\text{cm}^{-3}] \approx 2\times 10^9 I_i[\mu\text{A}]$, assuming that $T_e \sim 3$ eV. The electron temperature was determined from the current–voltage characteristic of the double probe at the floating potential using a conventional technique (see, for example, Ref. 11).

Most of the measurements were made in the discharge zone outside the plasma formation. This particularly applies to measurements of the electron temperature (the absolute values were determined to within ~25%). This is because, as we know, the probe measurements are exposed to the action of the microwave field[11] and the ion component of the probe current-voltage characteristic is distorted considerably less. Thus, whereas the charged particle concentration can be determined fairly reliably, the average electron energy (temperature) may be highly exaggerated. This first statement was checked out experimentally: the measured double-probe and single-probe characteristics were the same for the ion components. Since the microwave field differed in the layer near the probe (in the first case, it was determined by the difference in the strengths of the microwave field at the closely spaced probes whereas in the second case, it was determined by the microwave field acting between the probe and the counterprobe or chamber), this indicates that the ion component of the current–voltage characteristic is not sensitive to the microwave action under the experimental conditions described here.

The space outside the glowing zone is isotropic with respect to the potential and the double-probe characteristic passes through zero current at zero potential. Within the discharge the double-probe characteristic was shifted along the voltage axis, which indicates that the plasma is nonuniform and that a plasma potential difference exists at the locations of the probes.

Figure 2 gives the saturation ion current as a function of the radial position of the probe. Most of the data refers to the region outside the glowing discharge zone (the first point on the descending section approximately corresponds to the boundary of the glowing zone and this dimension can be used to estimate the discharge radius).

These estimates showed that the electron concentration ($n_i \approx n_e$) on the discharge axis at pressures of 1 and 3 Torr is close to critical (at 2.45 GHz we have $n_{ec} \approx 7\times 10^{10}\,\text{cm}^{-3}$). Similar measurements could not be made at higher pressures because the probe distorted the plasma and the electrodynamics of the system so severely that the reflection from the discharge section was altered. In some cases, the ion component of the probe current–voltage characteristic was also distorted and had the characteristic form described in Ref. 12.

These experimental results can be used to estimate the power P absorbed by the plasma. For this we can use the well-known relation

$$P = P_{ab}V = n_c \bar{\varepsilon}\,\delta\nu_{\text{ef}}V,$$

where $\bar{\varepsilon}$ is the average electron energy, δ is the average fraction of the energy lost by an electron in a single collision, ν_{ef} is the effective electron–heavy particle collision frequency, P_{ab} is the specific power absorbed by the plasma, and V is the plasma volume. The values of δ and ν_{ef} were obtained by self-consistent modeling of a hydrogen plasma.[10]

Estimates made at pressures of 1 and 3 Torr ($\delta\nu_{\text{ef}} \sim 3\times 10^7$ and $6\times 10^7\ \text{s}^{-1}$, respectively) taking into account the plasma dimensions (determined from Fig. 2) showed that the measured absorbed power agrees with the calculated value to within 20%. This is within the error band of the estimates and we can say that the power dissipated in the system is the power absorbed by the plasma. Similar estimates could not

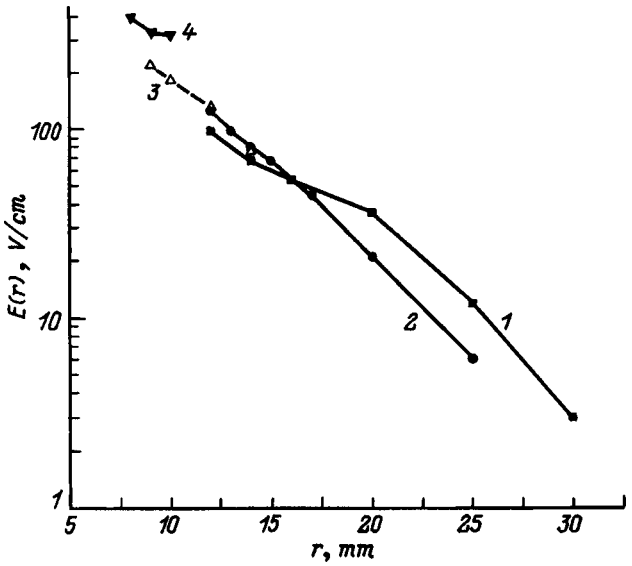


FIG. 4. Distribution of electric field strength near plasma formation (*1–4* — as in Fig. 2).

be made at pressures of 8 and 15 Torr because of the lack of data on the charged particle concentrations on the discharge axis (see above), but we can postulate that in these cases, the energy losses in the elements of the microwave channel are low.

These results allow us to determine the structure of the electromagnetic field in the space surrounding the glowing discharge zone. Assuming that there is no electric field outside the glowing zone, the characteristic scale of variation of the average electron energy is given by $v/\delta\nu_{\text{ef}}$ where v is the electron velocity in the direction being studied. The electron moves toward the wall as a result of ambipolar diffusion and at a pressure of 1 Torr we have $v_{\text{dif}} \sim 10^4$ cm/s. Taking the values of $\delta\nu_{\text{ef}}$ given above, the characteristic scale of variation of the average energy in the radial direction is $\sim 10^{-3}$ cm. It can be seen from Fig. 3 that the electron temperature varies considerably more slowly as a result of electron heating in the electromagnetic field.

The electric field strength $E(r)$ corresponding to the measured electron temperature ($T_e = 2\overline{\varepsilon}/3$) can be determined using the Boltzmann equation. The Boltzmann equation was solved by a numerical method described in detail in Ref. 10 allowing for elastic and inelastic collisions between electrons and heavy particles. The method of successive approximations was used to find $E(r)$ and the results are plotted in Fig. 4. It can be seen that the electric field strength decreases exponentially with increasing r. The only type of wave having this structure is a surface wave.[13] Thus, a surface wave concentrated in the near-surface layer propagates along the surface of the plasma formation. This clearly explains the increase in the radiation intensity in this region of the discharge.

The role of the surface wave in sustaining the discharge is unclear. Is its existence necessary or is its presence simply a consequence of different spatial modes of the electromagnetic waves, including surface modes, being formed in the plasma region when the homogeneity of the microwave transmission line is impaired? Further experimental and theoretical investigations are required to answer this question.

This work was supported financially by the Scientific Exchange Program of the Royal Swedish Academy of Sciences and the Russian Fund for Fundamental Research (Grant No. 96-02-184609).

The authors would like to thank Yu. F. Kolesnichenko, V. G. Brovkin, and G. S. Solntsev for useful discussions.

[1] S. V. Golubev, S. I. Gritsinin, V. G. Zorin *et al.*, *Rf Discharge in Wave Fields* [in Russian], Institute of Applied Physics, Academy of Sciences of the USSR, Gorky (1988), pp. 136–197.
[2] V. G. Brovkin, Yu. F. Kolesnichenko, and D. V. Khmara, *Ball Lightning in the Laboratory* [in Russian], Khimiya, Moscow (1994), pp. 119–135.
[3] G. M. Batanov, N. K. Berezhetskaya, E. F. Bol'shakov *et al.*, Plasma Sources Sci. Technol. **2**, 164 (1993).
[4] V. G. Brovkin, Yu. F. Kolesnichenko, and D. V. Khmara, Prikl. Fiz. No. 4, 5 (1994).
[5] V. I. Fedoseev, Yu. I. Aristov, Yu. Yu. Tanashev *et al.*, Kinet. Katal. **37**(6), 869 (1997).
[6] L. Bardos, H. Barankova, Yu. A. Lebedev *et al.*, in *Proceedings of the Seventh European Conference on Diamond, Diamond-like and Related Materials*, 1996, p. 4.1.
[7] Yu. A. Lebedev, Plasma Sources Sci. Technol. **4**, 4740 (1995).
[8] Yu. A. Lebedev, Teplofiz. Vys. Temp. **33**, 850 (1995).
[9] Yu. P. Raĭzer, M. N. Shneĭder, and N. A. Yatsenko, *Rf Capacitive Discharge: Physics, Engineering, and Applications* [in Russian], Moscow Physicotechnical University Press, Moscow (1995).
[10] Yu. A. Lebedev and I. L. Epstein, J. Mosc. Phys. Soc. **5**(1), 103 (1995).
[11] Yu. A. Ivanov, Yu. A. Lebedev, L. S. Polak, *Methods of Contact Diagnostics in Nonequilibrium Plasma Chemistry* [in Russian], Nauka, Moscow (1981), 142 pp.
[12] V. A. Dovzhenko, S. I. Kuznetsov, P. P. Mel'nichenko, and G. S. Solntsev, Zh. Tekh. Fiz. **47**, 2506 (1977) [Sov. Phys. Tech. Phys. **22**, 1457 (1977)].
[13] M. Moisan, A. Shivarova, and A. W. Trivelpiece, Plasma Phys. **24**, 1331 (1982).

Translated by R. M. Durham

Damage to optically transparent crystals with a macroscopic crack under pulsed laser irradiation

V. A. Fedorov, I. V. Ushakov, and V. P. Shelokhvostov

G. R. Derzhavin Tambov State University, 392622 Tambov, Russia
(Submitted February 19, 1997; resubmitted September 30, 1997)
Zh. Tekh. Fiz. **68**, 34–37 (December 1998)

An investigation was made of the influence of a macroscopic crack on the mechanical strength of NaCl, LiF, and $CaCO_3$ single crystals exposed to pulsed laser irradiation. The probability of growth of the initial macrocrack being activated with optical breakdown of the sample is estimated. A mechanism is proposed for the growth of an initial macroscopic crack in which cavities, microcracks, and macrocracks formed as a result of the heating and burnout of absorbing inclusions combine with the initial crack. © *1998 American Institute of Physics.* [S1063-7842(98)00612-6]

The presence of cracks in a transparent dielectric reduces its optical strength for various reasons. First, as a result of the multiple reflection of light from the walls of the crack, the electric field strength at certain points in the material will substantially exceed that of the incident wave, which increases the probability of multiphoton ionization, electron avalanche, and so on.[2] Second, after grinding and polishing, the surface of samples consists of relief-polycrystalline, cracked, and elastically strained layers. The surface layer inevitably contains suspended particles. The cracked layer contains microcracks lying in different planes.[1] As a result of the surfaces of the cracks undergoing chemical reactions with the surrounding medium, the absorption of light in the cracked layer increases,[2] the emission of electrons by the growing cracks reduces the optical breakdown threshold at the surface of the irradiated crystal,[3] and crack growth induced by thermal stresses created at absorbing inclusions may ultimately result in damage to the crystals.[4]

However, in most transparent materials the optical breakdown threshold is determined by the absorbing inclusions. Determinations of their sizes and the results of laser irradiation of these inclusions were reported in Refs. 1, 5, and 6.

A consequence of optical breakdown is mechanical damage to the material, but the contribution of initial cracks to the damage process at this stage remains unclear.

The aim of the present study is to investigate the damage to optically transparent crystals with an initial macroscopic crack under the action of pulsed laser radiation.

EXPERIMENTAL METHOD

The investigations were carried out using the following single crystals: NaCl, which were divided into three groups according to their impurity content: 1a — 10^{-2} wt % Cr^{3+}, 1b — 10^{-3} wt % Cr^{3+}, and 1c — Ca^{2+}, Mg^{2+}, and Ba^{2+} impurity contents each not exceeding 10^{-3} wt %; LiF having Ca^{2+}, Mg^{2+}, and Ba^{2+} impurity contents each not exceeding 10^{-3} wt %, and $CaCO_3$ with an impurity content of $\approx 10^{-2}$ wt %.

By cleaving a $15\times30\times16$ mm crystal, we obtained two identical samples each measuring $5\times30\times8$ mm. In one of these, a macroscopic crack 4–6 mm long was introduced in the {100} plane. The (100) plane was oriented perpendicular to the direction of propagation of the radiation. The samples were irradiated in pairs, one with a crack and one without.

We used a GOS-1001 pulsed laser with $\lambda = 1060$ nm. The pulse energy was varied between 40 and 400 J and the area of the irradiated surface was varied between 12 and 100 mm^2 (the pulse length varied between ≈ 0.6 ms at a pump energy of 40 J and ≈ 1.5 ms at a pump energy of 400 J). The crystals were irradiated by pulses starting at $E = 0.7E_{th}$ where E_{th} is the threshold pulse energy, and the energy of each successive pulse was increased by 10% until the sample fractured.

The size of the absorbing inclusions in the group 1a crystals did not exceed 10 μm. Figure 1 shows one large absorbing inclusion in a $CaCO_3$ crystal.

EXPERIMENTAL RESULTS

1) Growth of the initial macrocracks under the action of pulsed laser radiation was observed in the experiments. Table I gives experimental data on the probability of laser-

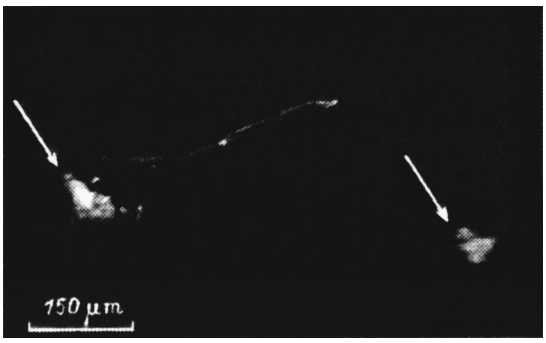


FIG. 1. Two absorbing inclusions consisting of filament-like veins in the bulk of $CaCO_3$. Their thickness at various points is 0.02–0.04 mm. In addition to the inclusions shown the crystal also contains inclusions of between 5 and 25 μm.

TABLE I. Statistics of initial macrocrack growth during laser-induced damage to samples.

Crystal group No.	During crystal damage initial crack grew to form main crack	During crystal damage initial crack grew by 2–3 mm	Crystal damaged without initial crack growth	*P*
NaCl (1a)	78.6%	7.14%	14.28%	0.88
NaCl (1b)	80%	Not identified	20%	0.8
NaCl (1c)	37.5%	37.5%	25%	0.7
$CaCO_3$	60%	Not identified	40%	0.6

Note: *P* is the probability of initial crack growth being activated during optical breakdown which caused the damage to the sample (including the case where damage is caused only by initial crack growth).

induced growth of an initial macrocrack when the samples are irradiated. Table II gives the statistics of damage to crystals with and without macrocracks exposed to a series of laser pulses. The coefficient $C=n_T/n_\sigma$ is used in the table to assess the degradation of the samples under exposure to a series of radiation pulses.

2) The damage to the samples is generally caused by the formation of laser-induced cracks at absorbing inclusions, which combine with the initial crack.

A laser-induced crack may be generated in the plane parallel to the plane of the initial crack and may combine with it as a result of transition cracks in the {110} and {100} planes (Fig. 2a). When a fairly advanced laser-induced crack lying in the (001) plane combines with an initial crack, this frequently results in part of the crystal being cleaved (Fig. 3a).

The onset of appreciable damage (the formation of large laser-induced cracks and, at the epicenter of optical breakdown, a friable mass of material fragments) almost always involves the initial crack. However, cases were observed when an initial crack propagated to form a main crack in the absence of any major damage centers.

If the crack changed its plane of propagation from {100} to {110} during the growth process, this always resulted in a dislocation pile-up near its tip. However, these sections of crack growth are of negligible length. Main cracks formed in the {100} planes while the cracks located in the {110} planes were only fairly well-developed in the immediate vicinity of the optical breakdown zone.

TABLE II. Statistics of laser-induced damage to samples with and without an initial crack.

Crystal group No.	1	2	3	*C*
NaCl (1a)	25	25	50	1.03
NaCl (1b)	15	65	20	0.75
NaCl (1c)	30	40	30	0.9
$CaCO_3$	50	17	33	0.96

Note: 1 — number of samples with and without cracks damaged under the same number of irradiation cycles (%); 2 — number of samples with cracks damaged as a result of fewer irradiation cycles compared with corresponding samples without cracks (%) (see experimental method); 3 — number of samples without cracks damaged as a result of fewer irradiation cycles compared with samples with cracks (%); $C=n_T/n_\sigma$ where n_T is the number of irradiation cycles before damage for crystals with an initial crack and n_σ is the number of irradiation cycles before damage for crystals without an initial crack.

DISCUSSION OF RESULTS

The experimental results suggest that there is a range of stresses at absorbing inclusions $\sigma_n \leqslant \sigma \leqslant \sigma_v$ (where σ_n is the elastic limit and σ_v is the stress sufficient to fracture the crystal) in which damage to the sample is caused by activation of initial crack growth. The nonlinear increase in the thermal stresses at absorbing inclusions above a certain critical optical intensity[7] indicates that the created stresses may be both lower than σ_n and higher than σ_v. In this case, 1) appreciable damage to the sample is usually accompanied by initial crack growth; 2) the probability of initial crack growth being activated during laser-induced damage to a sample reaches 80% (Table I), when $\sigma \gg \sigma_v$, damage is caused by the appearance of laser-induced cracks at absorbing inclusions, and in this case there is a high probability of these cracks interacting and combining with the initial crack; 3) there is no clear indication that the strength of an irradiated crystal tends to be reduced by an initial crack (Table II, coefficient *C*).

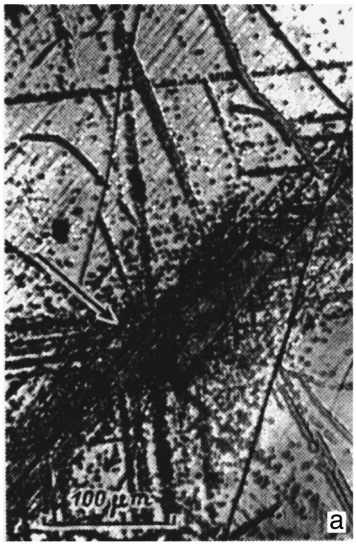
For NaCl crystals (Table II) we established a specific trend that as the impurity content in the sample decreases, the macrocrack plays a greater role in the degradation of the material under pulsed irradiation.

The probability of activation of initial crack growth depends on the thermal stresses created at absorbing inclusions and on the position of the absorbing inclusion relative to the crack. If heating of an absorbing inclusion in front of the tip of the initial crack gives rise to a laser-induced crack lying in the plane of the initial crack, these combine at the tip of the initial crack. In general (Figs. 2a and 2b), the plane of the laser-induced crack will be situated at a certain distance S_1 from the plane of the initial crack. When the laser-induced crack reaches the cross section *S* (Fig. 2b), there is a probability that the growing crack will change its plane of propagation to (110) or (010) and will combine with the initial crack. According to Gilman,[8] the probability of combining is especially high if the distance is $S_1 < 0.3S_2$ (Ref. 8).

When optical breakdown occurs at an absorbing inclusion near the surface of the initial crack, this crack grows as a result of the pressure of the forming gas (plasma) acting on the surface of the crack (Fig. 4).

Although there is a high probability of the cracks combining when $S_1 < 0.3S_2$, this condition is not sufficient. The probability of the cracks combining will increase with decreasing S_1 and increasing length of the laser-induced crack.

Crack growth in a solid is determined by $\sigma \geqslant \sigma_{cr}$



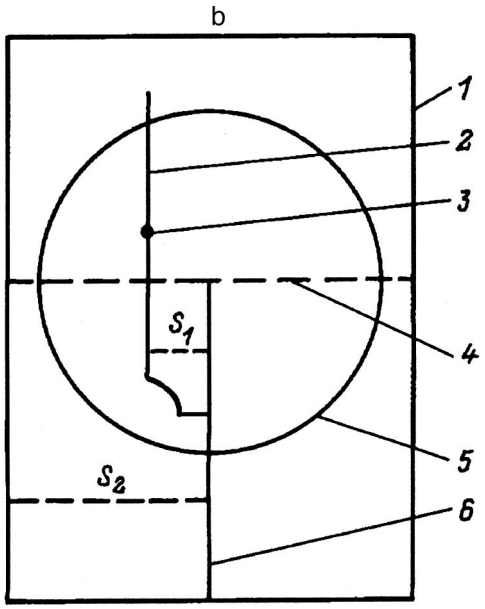
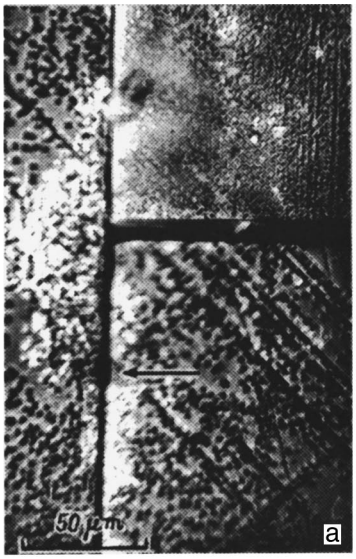


FIG. 2. a — Combining of a laser-induced crack with an initial one shown by the arrow; LiF(010) crystal surface; the laser-induced crack was formed as a result of heating of an absorbing inclusion positioned ahead of the tip of the initial crack; the initial and laser-induced cracks are located in the (100) plane; b — schematic showing growth of an initial crack by combining with a laser-induced one; *1* — crystal, *2* — laser-induced crack, *3* — absorbing inclusion, *4* – cross section *S*, *5* — irradiation zone, and *6* — initial macrocrack.



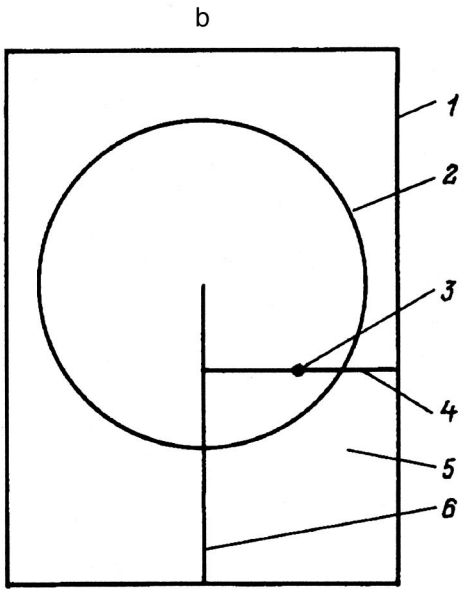


FIG. 3. a — Combining of a laser-induced crack with an initial one shown by the arrow; LiF(010) crystal surface; the absorbing inclusion at which the laser-induced crack forms is positioned ahead of the tip of the initial crack; cracks located in (100) plane; b — schematic of laser-induced crack combining with initial crack: *1* — crystal, *2* — irradiation zone, *3* — absorbing inclusion, *4* — laser-induced crack, *5* — cleaved section of crystal, and *6* — initial macrocrack.

$=(\alpha\gamma E/L_{\max})^{1/2}$ (Ref. 9), where α is a coefficient, γ is the effective surface energy, E is Young's modulus, and σ_{cr} is the minimum stress required for crack growth which is inversely proportional to the crack length $L_{\max}$. The growth of a laser-induced crack is accompanied by a reduction in the mechanical stresses mainly as a result of cooling of the gas or plasma formed by the heating of absorbing inclusions and by the plastic deformation which accompanies crack growth. As a result, the crack growth may cease (it was noted that fairly large cracks of up to 5–8 mm appeared and did not turn into main cracks). When a laser-induced crack combines with an initial one, this reduces the stress σ_{cr} required for its growth which helps to convert this crack into a main one.

The formation of a laser-induced crack (at an absorbing inclusion positioned ahead of the tip of the initial crack) in the (001) plane (Figs. 3a and 3b) causes cleaving of part of the crystal.

Under the action of radiation pulses of constant energy the cracks and cavities formed at an absorbing inclusion can be stabilized.[5] The presence of an initial crack ensures that there is a high probability of the resulting damage combining and interacting with the initial crack, which promotes damage, i.e., the initial macrocrack destabilizes the macrodamage formed in the bulk of the sample.

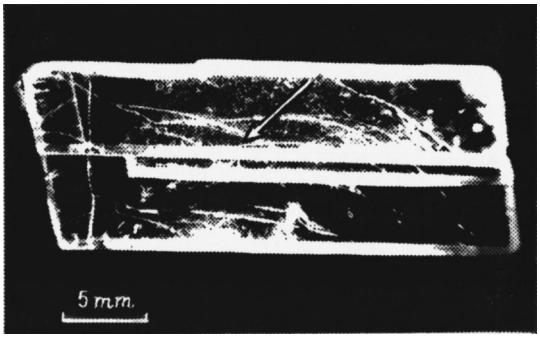


FIG. 4. Damage to $CaCO_3$ crystal as a result of optical breakdown at two large absorbing inclusions one of which was located near the surface of the initial crack. The position of the tip of the initial crack (before irradiation) is shown by the arrow.

CONCLUSIONS

1) For $CaCO_3$, NaCl, and LiF single crystals with the impurity content indicated above, the activation of initial macrocrack growth is predominantly caused by absorbing inclusions located in the bulk of the material.

2) The probability of initial crack growth being activated depends on the thermal stresses formed at the absorbing inclusion and on the position of the inclusion relative to its tip.

3) Laser-induced damage observed in the bulk of the sample becomes destabilized as a result of combining with the initial macrocrack.

[1] M. S. Bakharev, L. I. Mirkin, S. A. Shesterikov, and M. A. Yumasheva, *Structure and Strength of Materials under Laser Irradiation* [in Russian], Moscow University Press, Moscow (1988), 224 pp.
[2] S. V. Bilibin, V. N. Egorov, A. A. Katsnel'son *et al.*, Kvantovaya Élektron. (Moscow) **9**, 1912 (1982) [Sov. J. Quantum Electron. **12**, 1245 (1982)].
[3] V. N. Smirnov, Pis'ma Zh. Tekh. Fiz. **14**, 316 (1988) [Sov. Tech. Phys. Lett. **14**, 139 (1988)].
[4] V. A. Feodorov, I. V. Ushakov, and V. P. Shelohvostov, in *Proceedings of the Eighth International Symposium on the Influence of Structure Defects on the Damage Threshold of Transparent Dielectrics*, Braunschweig, Germany, 1997, paper WPB2-17. 3-9801433-8-4.
[5] V. P. Krutyakova and V. N. Smirnov, Zh. Tekh. Fiz. **48**, 844 (1978) [Sov. Phys. Tech. Phys. **23**, 495 (1978)].
[6] V. P. Krutyakova and V. N. Smirnov, Zh. Tekh. Fiz. **54**, 323 (1984) [Sov. Phys. Tech. Phys. **29**, 189 (1984)].
[7] S. I. Anisimov and B. I. Makshantsev, Fiz. Tverd. Tela (Leningrad) **15**, 1090 (1973) [Sov. Phys. Solid State **15**, 743 (1973)].
[8] J. J. Gilman, *Atomic Mechanism of Fracture* [Metallurgiya, Moscow, 1963, pp. 220–250].
[9] V. I. Vladimirov, *Physical Nature of Metal Fracture* [in Russian], Metallurgiya, Moscow (1984), 280 pp.

Translated by R. M. Durham

High-speed superplasticity of microcrystalline alloys under conditions of local grain boundary melting

V. N. Perevezentsev and Yu. V. Svirina

A. A. Blagondravov Institute of Mechanical Engineering, Nizhniĭ Novgorod Branch, 603024 Nizhniĭ Novgorod, Russia

(Submitted July 15, 1997)
Zh. Tekh. Fiz. **68**, 38–42 (December 1998)

A model is proposed for the high-speed superplasticity of materials under conditions of local grain boundary melting at temperatures close to solidus. It is shown that the local melting of grain boundaries containing segregations of impurity atoms, results in the formation of a structure consisting of liquid-phase regions and solid intergranular bridges which provide cohesion of the grains during the deformation process. The equilibrium concentration, dimensions, and activation energy for the formation of solid bridges are determined as a function of the temperature, initial impurity concentration in the boundary, and the boundary thickness. A mechanism is proposed for grain-boundary slip under conditions of local grain boundary at anomalously high strain rates. © *1998 American Institute of Physics.* [S1063-7842(98)00712-0]

INTRODUCTION

It has long been assumed that the deformation of materials under conditions of structural superplasticity takes place at homologous temperatures in the range 0.4–0.6 and strain rates $\dot{\varepsilon}=10^{-4}$–$10^{-2}\,\mathrm{s}^{-1}$. Recently, it was observed that some microcrystalline aluminum alloys and composites demonstrate all the characteristic features of superplasticity at temperatures close to the solidus temperature.[1–5] Superplastic deformation can then take place at anomalously high strain rates $\dot{\varepsilon}=10^{0}$–$10^{3}\,\mathrm{s}^{-1}$. This influence was called high-speed superplasticity (HSS). The behavior of materials under HSS conditions is similar to that under conditions of normal structural superplasticity although it has some characteristic features. In particular, the threshold HSS stress increases with decreasing grain size and is almost a linearly decreasing function of temperature. On passing through a certain temperature $T=T_i$, the activation energy of the deformation process which is determined from the slope of the curve $\ln\dot{\varepsilon}(1/T)$ shows an abrupt severalfold increase, and this is accompanied by a drop in the deforming stresses and an increase in the rate sensitivity coefficient $m=\partial\log\sigma/\partial\log\dot{\varepsilon}$. It was observed experimentally[1–5] that a material enters the HSS state as a result of local grain boundary melting at temperatures $T\geqslant T_i$, where T_i may differ from the solidus temperature by tens of degrees as a result of a difference in the chemical composition of the boundaries and the body of the grains. For aluminum HSS alloys, the local melting effect is caused by segregation at intercrystallite and interphase boundaries of magnesium, silicon, and various other elements. The optimum deformation regime corresponding to the maximum rate sensitivity coefficient and the limiting strain before rupture is achieved at temperatures within the range $T_i<T<T_S$, where the optimum temperature at a given strain rate correlates with the local melting temperature of the boundaries.

It follows from this reasoning that HSS theory should include a description of local grain boundary melting processes and an analysis of the material deformation mechanisms under these conditions. In Ref. 6 Perevezentsev proposed an HSS model based on the premise that the main mechanism responsible for HSS is slip along grain boundaries which comprise liquid layers containing a specific fraction of solid bridges which are needed to sustain the cohesion of neighboring grains during the plastic flow process. The aim of the present paper is to develop this model of the grain boundary structure and grain boundary slip under conditions of local melting of the boundaries.

MODEL OF LOCAL GRAIN BOUNDARY MELTING

It was suggested in Ref. 6 that at $T_i\leqslant T\leqslant T_S$ the grain boundary is a liquid layer containing islands of solid phase which connect neighboring grains (subsequently called *S*-bridges).

The existence of these bridges may be explained on the basis of the following reasoning. We shall postulate that at $T_i\leqslant T\leqslant T_S$ the grain boundary is a liquid layer of thickness δ containing impurity atoms whose volume concentration is C_i. It is easy to see that this structure is unstable with respect to fluctuations of the chemical composition. If, as a result of fluctuations of the chemical composition, the impurity concentration decreases to a certain level $C'<C_i$ in a certain volume element of the boundary $V>V^*$ (where V^* is the critical volume of a solid-phase nucleus), this volume of liquid is supercooled by a certain amount $\Delta T=T_S(C)-T$, where $T_S(C)$ is the solidus temperature of the alloy at $C=C'$, and there is a probability that it will crystallize. If δ is fairly small, the liquid crystallization process in local volumes of the boundary may result in the formation of solid bridges between grains. As an increasing number of new solid-phase bridges appear, the energy of the system reduces as a result of the release of crystallization heat but at the same time increases because of the increased chemical po-

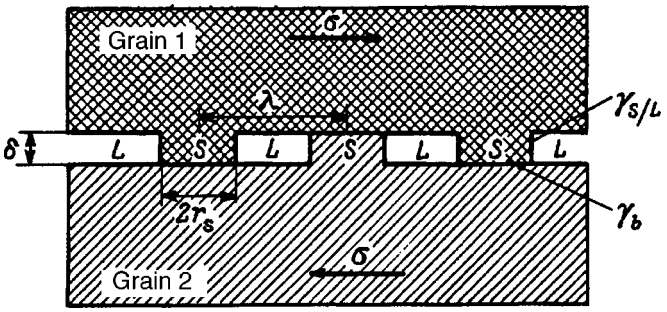


FIG. 1. Schematic of two-phase (solid–liquid) boundary: *L* — liquid phase and *S* — solid-phase bridges.

tential of the impurity atoms caused by their increased concentration in the liquid sections of the boundary (the total number of impurity atoms in the boundary is assumed to be constant).

We shall assume for simplicity that *S*-bridges are only formed when all the impurity atoms leave the volume of crystallizing liquid, and that each bridge is a tiny cylinder of height δ and radius r_S (see Fig. 1). Then, the change in the Helmholtz energy ΔF per unit area of the boundary accompanying the formation of *S*-bridges can be estimated from

$$\Delta F = n_S[-\pi\delta r_S^2\lambda_m\rho(1-T/T_S)+2\pi\delta r_S\gamma_{\mathrm{LIS}} + \pi r_S^2(\gamma_b-2\gamma_{\mathrm{LIS}})]+\Delta G_i. \quad (1)$$

Here n_S is the number of bridges per unit area of the boundary, λ_m is the specific heat of solidification, ρ is the density, and γ_{LIS} and γ_b are the specific surface energies of the liquid/crystal interface and the grain boundary, respectively. The first term in expression (1) describes the change in energy caused by the release of crystallization heat and the creation (disappearance) of new interfaces accompanying the formation of *S*-bridges. The second term describes the change in the chemical potential of the impurity atoms caused by an increase in their concentration in the liquid phase surrounding the bridge. Using the approximation of a strongly diluted homogeneous solution ($C_i a^3 \ll 1$), the quantity ΔG_i can be estimated using the expression

$$\Delta G_i = (C_i\delta)kT\ln(1+\Delta C_i/C_i), \quad (2)$$

where $C_i\delta$ is the number of impurity atoms per unit area of the boundary, ΔC_i is the increase in the concentration of impurity atoms in the liquid sections of the boundary caused by the formation of crystalline bridges,

$$\Delta C_i = C_i\pi r_S^2 n_S/(1-\pi r_S^2 n_S). \quad (3)$$

Assuming that the fraction of the solid phase in the total volume of the boundary is fairly small ($\pi r_S^2 n_S \ll 1$), we can use the approximate expressions $\Delta C_i/C_i = \pi r_S^2 n_S(1+\pi r_S^2 n_S)$, $\ln(1+\Delta C_i/C_i)\approx\Delta C_i/C_i$. Substituting these expressions into Eq. (2) gives

$$\Delta G_i \approx (C_i\delta)kT\pi r_S^2 n_S(1+\pi r_S^2 n_S). \quad (4)$$

By analyzing expressions (1) and (4), we can easily confirm that the condition of thermodynamic equilibrium between the solid bridges and the surrounding melt

$$\Delta F = 0, \quad \frac{\partial\Delta F}{\partial r_S}\delta r_S + \frac{\partial\Delta F}{\partial n_S}\delta n_S = 0$$

is satisfied for values of n_S^* and r_S^* given by

$$n_S^* = (\lambda_m\rho\Delta T/T_S+\Delta\gamma/\delta-C_i kT)^3/27\pi\gamma_{\mathrm{LIS}}^2 kTC_i \quad (5)$$

$$r_S^* = 3\gamma_{\mathrm{LIS}}/(\lambda_m\rho\Delta T/T_S+\Delta\gamma/\delta-C_i kT), \quad (6)$$

respectively, where

$$\Delta\gamma = 2\gamma_{\mathrm{LIS}}-\gamma_b, \quad \Delta T = T_S - T.$$

The fraction occupied by solid bridges in the total volume of the boundary $\alpha = n_S\pi r_S^2$ has the form

$$\alpha = (\lambda_m\rho\Delta T/T_S+\Delta\gamma/\delta-C_i kT)/3kT\,C_i. \quad (7)$$

The maximum of ΔF_S on the curve $\Delta F(r_S, n_S=n_S^*)$ may be taken to be the energy of formation (disappearance) of crystalline bridges for a given boundary thickness and impurity concentration. The analytic expression for ΔF_S has the form

$$\Delta F_S = \eta\pi\gamma_{\mathrm{LIS}}^2/(\lambda_m\rho\Delta T/T_S+\Delta\gamma/\delta-C_i kT), \quad (8)$$

where η is a numerical coefficient

$$\eta = \frac{3}{16}((\sqrt{3}-1)^4+(\sqrt{3}-1)-12(\sqrt{3}-1)^2).$$

Physical concepts indicate that a close correlation should exist between the impurity concentration C_i, the liquid layer thickness δ, and the temperature of local grain-boundary melting T_i. We shall estimate the thickness of the liquid layer δ for a given initial concentration of impurity atoms C_{i0} (C_{i0} is the concentration for the initial boundary thickness $\delta=\delta_0$) and temperature T assuming that the total number of impurity atoms in the boundary remains unchanged as the boundary layers melt and the boundary expands. This last assumption is justifiable if the profile of the impurity atom concentration as a function of the distance along the normal to the boundary decreases so rapidly that the contribution made by impurity atoms in the bulk of the grains to the value of C_i can be neglected. The relation between C_i and δ can then be written as

$$C_i = C_{i0}\frac{\delta_0}{\delta}. \quad (9)$$

To a first approximation, the driving force for expanding the liquid layer, acting per unit area of the boundary, is estimated from the expression

$$P_i = -\frac{\partial}{\partial\delta}\left(kTC_{i0}\delta_0\ln\left(\frac{C_i}{C_{i0}}\right)+\lambda_m\rho(\delta-\delta_0)\frac{\Delta T}{T_S}\right), \quad (10)$$

where the first term in parentheses describes the reduction in the chemical potential of the system caused by a decrease in the concentration of impurity atoms and the second term describes the increase in energy caused by melting of the boundary layers of grains as the liquid grain-boundary layer expands.

At a given temperature and initial impurity concentration, the boundary stops expanding when the condition

$P_i \leqslant 0$ is satisfied. Assuming that $P_i = 0$, we find the dependence of the liquid layer thickness on the temperature T and the initial impurity concentration C_{i0}

$$\frac{\delta}{\delta_0} = \frac{C_{i0} k T_S}{\lambda_m \rho} \frac{T}{\Delta T}. \tag{11}$$

An estimate of δ for typical values of the parameters $\lambda_m \rho a^3 \sim 1.5 k T_S$, $\delta_0 \sim 2a$, $C_{i0} a^3 \sim 0.1$ and typical temperatures for HSS $T \sim (0.96\text{–}0.99) T_S$ (Ref. 1) gives $\delta \sim (3 - 12)a$, which correlates with the experimentally determined thicknesses of the liquid layer in various alloys.[4]

Using the condition $\delta \geqslant \delta_0$ and expression (11), we find the temperature T_{Sb} at which a liquid layer of boundary containing an initial impurity concentration C_{i0} begins to expand

$$T_{Sb} = \frac{T_S}{1 + \dfrac{C_{i0} k T_S}{\lambda_m \rho}}. \tag{12}$$

Using expressions (5), (6), and (11), we find the dependence of the equilibrium number and size of *S*-bridges on the temperature and initial impurity concentration

$$n_S^* = \frac{1}{27\pi} \left(\frac{\Delta \gamma}{C_{i0} \delta_0 k T} \right)^3 \left(\frac{\lambda_m \rho}{\gamma_{\mathrm{LIS}}} \right)^2 \left(\frac{\Delta T}{T_S} \right)^2, \tag{13}$$

$$r_S^* = \frac{3 \gamma_{\mathrm{LIS}} C_{i0} \delta_0}{\lambda_m \rho} \frac{k T_S}{\Delta \gamma} \frac{T}{\Delta T}. \tag{14}$$

It can be seen from these expressions that as the temperature increases, the radius of the *S*-bridges increases but their number decreases. In this case, the fraction of the solid phase α in the total volume of the boundary decreases

$$\alpha = \frac{2 \gamma_{\mathrm{LIS}} - \gamma_b}{3 C_{i0} \delta_0 k T}. \tag{15}$$

It can be seen from formula (15) that the inequality $2\gamma_{\mathrm{LIS}} > \gamma_b$ must be satisfied for the existence of solid islands. Otherwise, the boundary becomes completely wetted with liquid, so that the solid bridges disappear. An estimate of the values of n_S^*, r_S^*, and α for typical values of the parameters $\lambda_m \rho a^3 \sim 1.5 k T_S$, $\delta_0 \sim 2a$, $C_{i0} a^3 \sim 0.1$, $\Delta \gamma a^2 \sim 0.1 k T_S$, and $\gamma_{\mathrm{LIS}} a^2 \sim 0.2 k T_S$ in the temperature range $T \sim (0.96\text{–}0.98) T_S$ gives $n_S^* a^2 \sim (3 \times 10^{-4})\text{–}(4 \times 10^{-5})$, $r_S^*/a \sim 12\text{–}39$, and $\alpha \sim 0.16$.

GRAIN-BOUNDARY SLIP UNDER LOCAL MELTING CONDITIONS

In this model the rate of grain-boundary slip is limited by the rate of propagation of the grain-boundary shear across the solid sections of the boundary (*S*-bridges). One possible mechanism for this process is the melting of a monatomic layer of material in any cross section of the *S*-bridge. This process will probably take place by thermally activated melting of the grain boundary in the region of contact between a solid bridge and one of the grains. In this last case, the energy dissipated as a result of the melting and formation of new interfaces between the liquid and solid phases will be partially compensated by a decrease in the energy surfaces $\pi r_S^2 \gamma_b$ as the relevant section of the boundary disappears. The activation energy of this process F_a may be written as

$$F_a = \pi (r_S^*)^2 \left(\lambda_m \rho \frac{\Delta T}{T_S} \delta_0 + 2\gamma_{\mathrm{LIS}} - \gamma_b \right) \tag{16}$$

or

$$F_a = \pi (r_S^*)^2 \lambda_m \rho \delta_0 \frac{T_{Sb}^0 - T}{T_S}, \tag{17}$$

where the following notation is introduced

$$T_{Sb}^0 = T_S \left(1 - \frac{\gamma_b - 2\gamma_{\mathrm{LIS}}}{\lambda_m \rho \delta_0} \right).$$

It is easy to see that T_{Sb}^0 has the meaning of the melting point of a boundary of thickness δ_0 containing no segregated impurity atoms. Using formula (14), we write the expression for F_a in the form

$$\frac{F_a}{k T_S} = 9\pi \left(\frac{\gamma_{\mathrm{LIS}}}{\Delta \gamma} \right)^2 \left(\frac{k T_S}{\lambda_m \rho a^3} \right) \left(\frac{\delta_0}{a} \right)^2 (C_{i0} a^3)^2 \left(\frac{T}{\Delta T} \right)^2 \left(\frac{T_{St}^0 - T}{T_S} \right). \tag{18}$$

An estimate of F_a for typical values of the parameters $\lambda_m \rho a^3 \sim 1.5 k T_S$, $\delta_0 \sim 2a$, $C_{i0} a^3 \sim 0.05$, $\Delta \gamma a^2 \sim 0.1 k T_S$, $\gamma_{\mathrm{LIS}} a^2 \sim 0.2 k T_S$, $T_{Sb}^0 \approx 0.996 T_S$ in the temperature range $T \sim (0.94\text{–}0.98) T_S$ gives $F_a \sim (4\text{-}16) k T_S$. It can be seen from Eq. (18) that the activation energy of grain-boundary slip initially increases with temperature and then decreases to zero at T_{Sb}^0, when the liquid layer extends over the entire length of the boundary.

The process of grain-boundary slip can be broadly represented as follows.[6] In a field of external stresses σ, the thermally activated melting of an *S*-bridge in its region of contact with one of the grains leads to a relative displacement of the grains along a zone of characteristic dimensions 2λ (where λ is the distance between the *S*-bridges) accompanied by a localized power shift $B = (\lambda \sigma)/G$ along the perimeter of this zone which, for simplicity, will subsequently be considered to be a ''dislocation loop'' with the Burgers vector B. Further expansion of this loop in the plane of the boundary may take place in two different regimes. The first is achieved when the external stress acting on the ''dislocation'' is sufficiently high for it to propagate across the *S*-bridge by force (nonactivation regime). The second regime is observed at lower external stresses when the shear propagates through the *S*-bridge by a thermally activated mechanism. The first regime is the most interesting to explain the experimentally observed high strain rates. It was shown in Ref. 6 that in order to achieve this regime, the external stress σ should exceed a certain threshold σ_0

$$\sigma_0 = (GQ)^{1/2} (n_S^*)^{3/4}, \tag{19}$$

where G is the shear modulus and Q is the energy needed to cross a single *S* bridge.

Assuming that the *S*-bridge is crossed by the above mechanism and neglecting the work of the external stresses, to a first approximation we can write $Q \approx F_a$. Then, using expressions (18) and (19), we obtain

$$\frac{\sigma_0}{G}=K\left(\frac{T}{T_S}\right)^{3/4}\frac{(T_S-T)^{1/2}(T_{Sb}^0-T)^{1/2}}{T_S}, \tag{20}$$

where

$$K\approx 6\times 10^{-2}\frac{\lambda_m\rho a^3}{\sqrt{Ga^3\gamma_{\mathrm{LIS}}a^2}}\left(\frac{\Delta\gamma a^2}{C_{i0}a^3kT_S}\right)^{5/4}. \tag{21}$$

For typical values of the parameters used above and $Ga^3\sim 50kT_S$, $C_{i0}a^3\sim 0.1$, and the temperature $T=0.96T_S$, we obtain $\sigma_0\sim 1.2\times 10^{-3}G$. From expression (20), we can see that the threshold stress decreases with increasing temperature until it reaches zero at $T=T_{Sb}^0$. This may well explain the experimentally observed substantial drop in the flow stress with increasing temperature in superplastic alloys which exhibit high-speed superplasticity.[1] It is also interesting to note that this expression can explain the experimentally observed unexpected increase in the threshold stress with decreasing grain size. In fact, assuming that the number of impurity atoms is constant and that they all segregate at boundaries, a reduction in the grain size leads to a drop in the initial impurity concentration C_{i0} in the boundary ($C_{i0}=C_{i\nu}(d/\delta)$, where $C_{i\nu}$ is the volume-averaged impurity concentration). It can be seen from expressions (20) and (21) that a decrease in C_{i0} increases the threshold stress σ_0.

Assuming that the rate of expansion of dislocation loops propagating in the force regime is limited by the viscous resistance of the liquid, the expression for the strain rate limited by grain boundary slip may be given as[6]

$$\frac{\partial\varepsilon}{\partial t}=\left(\frac{\sigma-\sigma_0}{G}\right)^2\left(\frac{2\pi D_{\mathbf{l}}}{ad}\right)\left(\frac{Ga^3}{kT}\right)\exp\left(-\frac{F_a}{2kT}\right), \tag{22}$$

where $D_{\mathbf{l}}$ is the coefficient of self-diffusion in the liquid phase and d is the average grain size.

An estimate of $\partial\varepsilon/\partial t$ for typical values of the parameters $D_{\mathbf{l}}\sim 10^{-5}\,\mathrm{cm}^2/\mathrm{s}$, $a\sim 3\times 10^{-8}\,\mathrm{cm}$, $Ga^3/kT\sim 50$, the typical values for HSS $(\sigma-\sigma_0)/G\sim 10^{-3}$, $d\sim 500a$, and the theoretically determined values $F_a\sim(4-16)kT_S$ gives $\partial\varepsilon/\partial t\sim 10^1-10^3\,\mathrm{s}^{-1}$. This estimate shows that this mechanism of grain-boundary slip can give the very high strain rates observed experimentally.

An additional analysis of the accommodation processes of grain-boundary slip at grain joints[7,8] and allowance for the influence of disperse second-phase particles on the threshold stress σ_0 is required to obtain a more adequate description of the strain rate as a function of the stress, temperature, and grain size.

CONCLUSIONS

1. A model has been proposed for the local melting of grain boundaries containing an excess concentration of impurity atoms. It has been shown that under conditions of local melting, the grain boundary is a layer of liquid containing solid-phase islands which sustain the cohesion between the grains during deformation.

2. An analysis has been made of a mechanism of grain-boundary slip along a two-phase (liquid–solid) boundary which essentially involves thermally activated melting of solid bridges followed by forcible propagation of the incipient dislocation loops in the boundary.

3. Expressions have been obtained to estimate the activation energy and threshold stress for grain-boundary slip and to analyze their dependence on the temperature and initial concentration of impurity atoms in the boundary.

[1] K. Higashi, Mater. Sci. Forum **170–172**, 131 (1994).
[2] J. Koike, M. Mabuchi, and K. Higashi, Acta Metall. Mater. **43**, 199 (1995).
[3] M. Mabuchi and K. Higashi, Philos. Mag. Lett. **70**, 1 (1994).
[4] K. Higashi, T. G. Nieh, M. Mabuchi, and J. Wadsworth, Scr. Metall. **32**, 1079 (1995).
[5] K. Higashi, T. G. Nieh, J. Wadsworth, Acta Metall. Mater. **43**, 3275 (1995).
[6] V. N. Perevezentsev, Zh. Tekh. Fiz. **66**(12), 140 (1996) [Tech. Phys. **41**, 1275 (1966)].
[7] V. N. Perevezentsev, V. V. Rybin, and V. N. Chuvil'deev, Acta Metall. Mater. **40**, 887 (1992).
[8] V. N. Perevezentsev, V. N. Chuvil'deev, and S. A. Larin, Mater. Sci. Forum **170–172**, 613 (1994).

Translated by R. M. Durham

Magnetic field diffusion during a metal–insulator phase transition

Yu. B. Kudasov

Russian Federal Nuclear Center, 607190 Sarov, Nizhniĭ Novogorod Region, Russia
(Submitted August 19, 1997)
Zh. Tekh. Fiz. **68**, 43–48 (December 1998)

An analysis is made of the diffusion of a static (slowly varying) magnetic field in a conductor in which a first-order phase transition to the insulating state takes place under the action of the Joule heating. An investigation is made of the case of subsonic propagation of the phase boundary. A $(V_{1-X}Cr_X)_2O_3$ solid solution is analyzed as a model substance. The application of this effect in pulsed high-current circuit breakers is discussed. © *1998 American Institute of Physics.* [S1063-7842(98)00812-5]

1. INTRODUCTION

An analysis is made of the nonlinear steady-state diffusion of a magnetic field in a substance in which a first-order metal–insulator phase transition takes place under the action of Joule heating. This effect is studied with the prospect of constructing pulsed high-current circuit breakers in mind.

Many different materials are now known in which the conductivity decreases substantially under the action of temperature and strong magnetic fields, i.e., factors normally encountered during the action of pulsed high-current devices. These include oxides and sulfides of transition metals,[1] manganites of lanthanides (see, for example, Ref. 2), and so on. Thus, two possible types of solid-state pulsed circuit breakers, thermal and magnetic, can be arbitrarily distinguished. Combined regimes are also possible. In the present study a $(V_{1-X}Cr_X)_2O_3$ solid solution will be analyzed for the numerical estimates presented below. This material has been fairly widely used in low-power thermistors.[3] In this material, a first-order metal–semiconductor phase transition with conduction losses of between two and three orders of magnitude takes place at temperatures around 400 K (Ref. 4) so that it can be used as a thermal circuit breaker, i.e., using loss of conduction induced by the action of Ohmic heating. Changes in the chromium concentration in the solution continuously shift the phase transition temperature. Below room temperature, the conduction jump may reach 4–5 orders of magnitude. In addition, it was recently shown that a metal–semiconductor phase transition can also be stimulated in this material by strong magnetic fields, especially near the critical point (magnetic circuit breaker[5]). The following analysis will only consider the regime of a thermal circuit breaker, i.e., the dependence of the phase transition on the magnetic field will be neglected. It will be assumed that the durations of the processes being studied are between a few microseconds and tens of microseconds, which is typical of explosive magnetic current generators such as the MK-2 (Ref. 6). In addition to specific results for $(V_{1-X}Cr_X)_2O_3$ solid solutions, general conclusions and constraints on the circuit breaker material will be put forward.

2. MODEL OF STEADY-STATE PROPAGATION OF A PHASE TRANSITION WAVE

An analysis will be made of the steady-state one-dimensional planar diffusion of a static magnetic field in a material in which a first-order metal–insulator phase transition takes place under the action of Joule heating (Fig. 1). The propagation velocity of the phase boundary is assumed to be lower than the velocity of sound in the metal and insulator (homogeneous) phases. The magnetic field **B** depends on the x coordinate along which diffusion takes place and is directed along the z axis. In the insulator phase the magnetic field is constant. As will be seen subsequently, rapid propagation of a phase boundary in a material is accompanied by the creation of appreciable pressures because of the change in the specific volume as a result of the phase transition. Thus, effects associated with strength will be neglected and the hydrodynamic approximation will be used.

We shall write the equation of continuity and the Euler equation in the form

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) = 0, \quad \frac{\partial u}{\partial t} + u\nabla u = -\frac{1}{\rho}\nabla P, \tag{1}$$

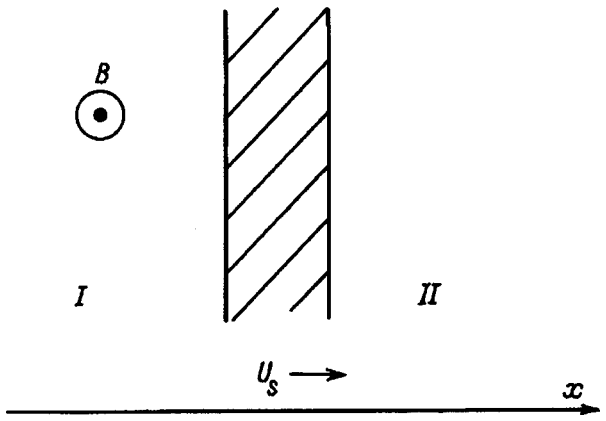


FIG. 1. One-dimensional steady-state magnetic field diffusion: I — metal phase, II — insulator phase, shaded region — heterogeneous phase mixture (with slow wave propagation).

where ρ is the density, u is the velocity of the medium, $P=P_h+P_m, P_h$, and $P_m=B^2/8\pi$ are the hydrodynamic and magnetic pressures, respectively.

We are interested in the steady-state solutions of these equations. We convert to a reference frame moving at constant velocity u_s in step with the phase boundary. For steady-state propagation of the phase boundary we obtain from system (1)

$$u=\frac{A}{\rho}=AV, \quad P=-Au+P_0=-A^2V+P_0, \tag{2}$$

where V is the specific volume, P_0 is the constant component of the pressure, and A is a constant determined by the boundary conditions: $u_S\rho|_{x\to\infty}$.

We shall assume that the phase transition temperature T_C is not too high and the heating of the material during operation of the circuit breaker is not too great (~ 100 K). In this case, the specific heat can be considered to be constant. For a $(V_{1-X}Cr_X)_2O_3$ solid solution T_C is around 400 K. We shall assume that for $(V_{1-X}Cr_X)_2O_3 C_P=110$ J/K·mol for both the metal and the insulator phases.[7] Since the volume jump ΔV induced by the phase transition ($\approx$1% from Ref. 8) appreciably exceeds the change in volume caused by thermal expansion, we shall neglect the latter. We then have $C_P=C_V$. In addition, the propagation velocity is assumed to be sufficiently fast so that heat conduction can be neglected. This is valid if the characteristic time of propagation of the phase boundary τ considerably exceeds a^2C_p/λ, where λ is the thermal conductivity and a is the characteristic dimension of the problem. In our case, a should be taken to be the size of the transition region, which we know to be of the order of 1 mm. No reference data are available on the thermal conductivity of V_2O_3. As an estimate we take the thermal conductivity of V_2O_5 which is of the order of 4 W/m·K in the range 300–400 K (Ref. 9). Note that the electron component of the heat conduction in V_2O_3 is small because of its comparatively poor conductivity and has no significant influence on the thermal conductivity. We then obtain characteristic heat propagation times of the order of 1 s which is much greater than the typical operating time of a pulsed circuit breaker.

We shall assume that the compressibility of both phases of $(V_{1-X}Cr_X)_2O_3$ is $\varkappa=9.735\times10^{-8}$ cm^5/g·bar (Ref. 10) and in our model there is no difference between the adiabatic and isothermal compressibility because we neglect the thermal expansion. Under all these conditions, the model $P-V$ diagram will have the form shown in Fig. 2.

Note that in general three regions may exist: a metal phase I, a insulator phase II, and a mixed (heterogeneous) region. We shall assume that the conductivity σ_0 in the metal phase is constant (i.e., independent of temperature and pressure). Efficient operation of a circuit breaker requires that the skin-layer depth in the insulator phase should be much greater than the wall thickness of the breaking element, so that we can take the conductivity in the insulator phase to be zero. If necessary, the residual conductivity can easily be taken into account. In the mixed phase, neglecting the percolation threshold effect, we assume that the conductivity is

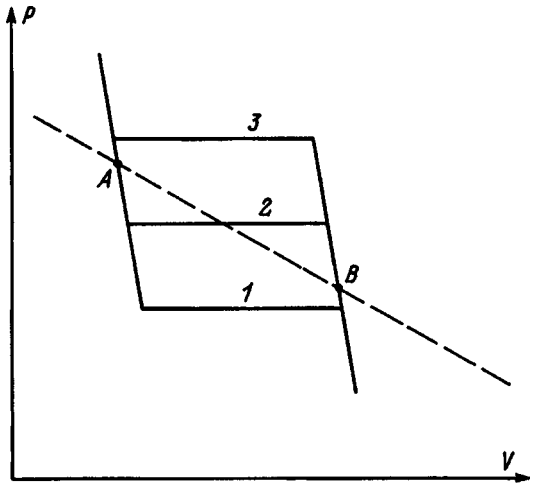


FIG. 2. Model of the equation of state: *1–3* — isotherms in the mixed phase at various temperatures (in order of increase).

simply proportional to the bulk fraction of the metal phase. The conductivity as a function of the specific volume along the line AB will then have the form

$$\sigma=\sigma_0\left(1-\frac{V-V_A}{\delta V'}\right), \tag{3}$$

where $\delta V'=V_B-V_A=\Delta V/(1-A^2\varkappa)$ and $V_{A(B)}$ is the volume at point $A(B)$.

Solving the magnetic field diffusion equation in the metal phase is elementary; difficulties arise when studying diffusion in the mixed phase since in this case, the conductivity and velocity of the medium depend on the coordinate. The magnetic field diffusion equation is written as

$$\frac{\partial\mathbf{B}}{\partial t}=\mathrm{curl}(\mathbf{u}\times\mathbf{B})-\mathrm{curl}(\nu_m\,\mathrm{curl}\,\mathbf{B}), \tag{4}$$

where $\nu_m=c^2/4\pi\sigma$ is the magnetic viscosity.

In this case, the conductivity in the mixed phase depends on the coordinate. We shall seek a steady-state solution in a fixed reference frame, i.e., of the type $\mathbf{B}(x-u_{st})$. The characteristic velocities of the medium u will then not exceed $u_S\Delta V/V$. In any case, for $(V_{1-X}Cr_X)_2O_3$ the magnetic Reynolds number is $\mathrm{Re}_m=ul_0/\nu_m\ll1$, where l_0 is the length scale over which the magnetic field changes and the first term on the right-hand side of Eq. (4) can be neglected. Equation (4) can then be written in the form

$$\nabla^2B+\left(E\sigma-\frac{\nabla\sigma}{\sigma}\right)\nabla B=0, \tag{5}$$

where $E=4\pi u_S/c^2$. Here and subsequently, the equation is written for a single (z) component of the vector **B**.

Finally, the system of equations (1) and (5) must be supplemented by the heat balance equation. Here we obtain a very complex system of equations since Eq. (5) is coupled to the system (1) via both the heat balance equation and the magnetic pressure appearing in Eq. (1). In this case, the hydrodynamic pressure in the mixed phase will be defined by a certain curve in the $P-V$ plane. However, we shall see that

in the cases of interest, the magnetic pressure gradient in the mixed phase can be neglected, i.e., we shall assume that $P \propto P_h$ in the mixed phase.

Along the line AB (Fig. 2) the process in the mixed region takes place with the absorption of Joule heat. From this it follows that a necessary condition for the existence of a mixed region is that the slope of the line AB should be less than the slope of the adiabatic curve in the mixed phase. We calculate the absorbed specific heat Q along the line AB

$$dQ|_{AB}=(TdS)_{AB}=T\left\{\left(\frac{\partial S}{\partial T}\right)_V dT+\left(\frac{\partial S}{\partial V}\right)_T dV\right\}_{AB}$$

$$=\left(C_V' dT+T\frac{\Delta S}{\Delta V}dV\right)_{AB}, \tag{6}$$

where ΔS is the entropy jump accompanying the phase transition (for $(V_{1-X}Cr_X)_2O_3$ $\Delta S=0.63$ J/K·mol (Ref. 3) and C' is the specific heat of the phase mixture at constant volume.

Since the specific heats of the metal and insulator phases are assumed to be the same, the entropy of the phase mixture may be given as $S=S_I(T)+\alpha\Delta S$, where $S_I(T)$ is the entropy of the metal phase and α is the specific fraction of the insulator phase in the mixture. Using the Clausius–Clapeyron formula $\partial P/\partial T=\Delta S/\Delta V$, we then obtain

$$C_V'=T\left\{\left(\frac{\partial S_I}{\partial T}\right)_V+\left(\frac{\partial \alpha}{\partial T}\right)_V \Delta S\right\}=C_V+T\left(\frac{\Delta S}{\Delta V}\right)^2 \varkappa. \tag{7}$$

In expression (6) V and T are independent variables so that we again use the Clausius–Clapeyron formula and Eq. (2), and then the specific heat can be represented as a function only of the specific volume

$$dQ|_{AB}=\left(-A^2 C_V'\frac{\Delta V}{\Delta S}+T\frac{\Delta S}{\Delta V}\right)dV. \tag{8}$$

When the velocity of the phase transition wave reaches the adiabatic velocity of sound in the mixed phase, the expression in parentheses in Eq. (8) becomes zero. The mixed region then disappears since the Joule heat absorbed in this region should be zero. Thus, two propagation regimes of the phase transition wave can be identified: a slow regime when a mixed region exists between the metal and insulator phases, and a fast regime when there is no mixed region and a discontinuity forms. The critical velocity is the adiabatic velocity of sound in the mixed phase. In our model, this has the form

$$u_c=\frac{\Delta S}{\rho \Delta V}\sqrt{\frac{T}{C_V'}}. \tag{9}$$

We express the Joule heat in terms of ∇B

$$\rho\frac{\partial Q}{\partial t}=\frac{c^2(\nabla B)^2}{16\pi^2\sigma}. \tag{10}$$

Finally, we equate the heat released by diffusion of the magnetic field (10) and that absorbed by the process taking place along line AB from Eq. (8), and we obtain the heat balance equation

$$\frac{(\nabla B)^2}{\sigma}=D\nabla\sigma, \tag{11}$$

where

$$D=-\frac{16\pi^2\delta V'}{c^2 u_S \sigma_0 \rho}\left(-A^2 C_V'\frac{\Delta V}{\Delta S}+T\frac{\Delta S}{\Delta V}\right),$$

and V was replaced by σ using Eq. (3).

Equations (5) and (11) form a system for the functions $B(x)$ and $\sigma(x)$ which can be used to find a solution for steady-state diffusion of the magnetic field in the mixed phase.

3. STRUCTURE OF THE PHASE TRANSITION WAVE

1. Fast propagation of the phase transition wave $u_S > u_C$. Diffusion of the magnetic field takes place in the metal phase with the boundary condition $x_{MI}(t)=-u_S t$, where x_{MI} is the boundary between the metal and the insulator phases. The solution of Eq. (5) then has the form

$$B=B_0\exp\left(-\frac{u_S}{\nu_m}[x-u_S t]\right), \tag{12}$$

where B_0 is the magnetic field in the insulator phase.

Note that this expression contains two independent variables B_0 and u_S. The temperature as a function of the coordinate can easily be calculated under the conditions formulated above

$$T=\frac{B_0^2}{8\pi\rho C_V}\exp\left(-2\frac{u_S}{\nu_m}[x-u_S t]\right)+T_0. \tag{13}$$

At the point x_{MI} the temperature should be equal to the phase transition temperature T_{MI}. Thus, the additional constraint is imposed

$$T_{MI}=\frac{B_0^2}{8\pi\rho C_V}+T_0, \tag{14}$$

and only one independent variable remains. In principle T_{MI} may be a function of pressure.

The structure of a fast wave is shown schematically in Fig. 3 and Fig. 4 shows the correlation between the external magnetic field and the wave propagation velocity, and also gives the pressure jump at the phase boundary at a function of the wave velocity.

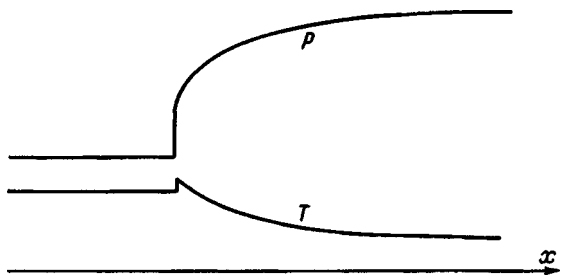


FIG. 3. Structure of fast phase transition wave. A pressure jump occurs at the phase boundary; the subsequent smooth increase in pressure is caused by the magnetic field gradient.

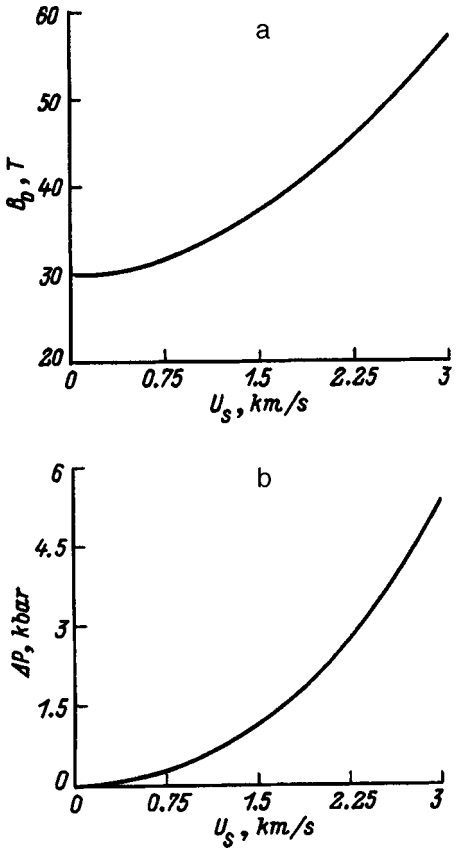


FIG. 4. Fast wave in a $(V_{1-X}Cr_X)_2O_3$ solid solution: a — external magnetic field as a function of wave propagation velocity; b — pressure jump at phase boundary as a function of wave propagation velocity.

2. Slow propagation of phase transition wave $u_S < u_C$. The magnetic field and temperature distribution in the metal phase can be determined as for fast propagation. The system of equations (5) and (11) must be solved to find the solutions in the mixed region. We first compare the two terms in parentheses in Eq. (5). We introduce a characteristic number having the same form as the Reynolds number, $\mathrm{Re}'_m = u_S l_0 / \nu_m$, where $l_0 = \sigma/\nabla\sigma$ is the characteristic length at which the conductivity changes. Depending on this characteristic number, we can neglect one of the terms in parentheses in Eq. (5). If $\mathrm{Re}'_m \gg 1$, the second term can be dropped. Then, making the substitution

$$\nabla B = f, \tag{15}$$

we can easily obtain a solution of the system (5) and (11) in the form

$$f = C_2 \exp\left(\frac{x - u_S t}{DC_I^2}\right), \tag{16}$$

where C_1 and C_2 are constants.

It is easily established that for $(V_{1-X}Cr_X)_2O_3$ solid solutions the condition $\mathrm{Re}'_m \gg 1$ is satisfied and the solution (16) is found in the mixed region.

In principle, it is also possible to have the case $\mathrm{Re}'_m \ll 1$. Thus, we also find its solution. We can now drop the first term in parentheses in Eq. (5). We again change the variables (15) and also change $|f| = \sqrt{|g|}$, $\mathrm{sign}(g) = \mathrm{sign}(f)$. The system of equations then has the form

$$g' + 2E\sigma g = 0, \quad g = D\sigma\sigma'. \tag{17}$$

We substitute g from the second equation (17) into the first and obtain an equation with separable variables

$$g' = -2ED\sigma^2\sigma', \tag{18}$$

which gives

$$g = -\frac{2}{3}ED\sigma^3 + C_1, \tag{19}$$

where C_1 is the integration constant.

Using Eqs. (18) and (19) and changing the variables $x = 3\xi/2E$, we obtain an equation with separable variables

$$\frac{\partial\sigma}{\partial\xi} = \frac{C'^3_1 - \sigma^3}{\sigma}, \tag{20}$$

where $C'_1 = (3C_1/2ED)^{1/3}$.

By integrating Eq. (20), we can determine the implicit dependence of the conductivity on the coordinate

$$\xi + C_2 = \int \frac{\sigma d\sigma}{C'^3_1 - \sigma^3} = \frac{1}{6C'_1}\ln\frac{C'^2_1 + C'_1\sigma + \sigma^2}{(C'_1 - \sigma)^2} - \frac{1}{\sqrt{3C'_1}}\arctan\frac{2\sigma + C'_1}{\sqrt{3C'_1}}. \tag{21}$$

Here C_2 is the integration constant. From this we can also find the distribution of the magnetic pressure field, and so on. Note that when $\sigma \to 0$, i.e., near the boundary of the mixed and insulator phases, the dependence of the conductivity on the coordinates may be expressed in the explicit form

$$\sigma = \sqrt{2C'^3_1\xi}. \tag{22}$$

It is interesting to note that, as can be seen from Eq. (22), at the boundary of the mixed and insulator phases where $\mathrm{Re}'_m \ll 1$, the derivatives of the conductivity, density, and pressure increase without bound whereas the functions themselves are bounded and continuous.

4. DISCUSSION AND CONCLUSIONS

We shall first consider the assumption that $P \propto P_h$ in the mixed phase, which was used to investigate the slow wave. This assumption is valid when the wave propagation velocity is not too low compared with the critical value. Then, the mixed region is narrow and the magnetic field gradient is small compared with the hydrodynamic pressure gradient (in the limit $u_S \to u_C$ we have for the pressure change in the mixed region: $\Delta P_h \to \mathrm{const}$, $\Delta P_m \to 0$). Thus for a $(V_{1-X}Cr_X)_2O_3$ solid solution with $u_S = 0.8u_C$ we find that the typical distance over which the conductivity varies, DC_1^2 from formula (16), is of the order of 0.002 cm and the variation of the magnetic field over this length is less than 100 G. In this case, the magnetic pressure gradient is negligible compared with the hydrodynamic pressure gradient in the

mixed region (around 2.5 kbar/cm). It can be seen from this example that the range of validity of the solutions (16) and (21) is fairly wide, at least for $(V_{1-X}Cr_X)_2O_3$.

For $(V_{1-X}Cr_X)_2O_3$ the hydrodynamic velocity of sound is around 6.5 km/s and the critical velocity of the phase transition wave calculated from Eq. (9) is only 100 m/s. Thus, there is a very wide range of solutions with fast wave propagation. The fast wave regime is preferable for a circuit breaker because the mixed region causes expansion of the skin layer and increases the breaking time. It can be seen from formula (9) that the critical velocity is proportional to the ratio $\Delta S/\Delta V$. In $(V_{1-X}Cr_X)_2O_3$ this ratio is anomalously small[3] which ensures a large range of solutions with fast wave propagation.

For $(V_{1-X}Cr_X)_2O_3$ there is another specific constraint on the range of steady-state solutions with fast wave propagation. Let us assume that when $x\to\infty$ the material is fixed (anchored boundary) and when $x\to-\infty$, the pressure in the material is equal to atmospheric pressure P_0 (free boundary). Then the pressure in the material on the metal side of the phase boundary ($x\to+x_{MI}$) is $P_0+\Delta P$, where ΔP is the pressure jump at the phase boundary. However, as the pressure increases, the metal–insulator phase transition line for $(V_{1-X}Cr_X)_2O_3$ ends at a critical point.[1] The maximum value of ΔP depends on the initial temperature but in any case, cannot exceed 8 kbar (Ref. 1). In this case, the wave velocity is around 4 km/s. At higher velocities the propagation of the phase transition wave may have singularities.

In order to calculate the characteristics of a circuit breaker, we need to investigate the essentially transient processes of wave formation and its propagation to the boundary. Nevertheless, the main parameters of a circuit breaker can be estimated from these results. Let us assume that a barrier (circuit breaker) made of a $(V_{1-X}Cr_X)_2O_3$ solid solution separates the space in which the initial current pulse (and magnetic field) is generated from the load space. When the barrier is transferred rapidly to the insulator state, a peaked current pulse (magnetic field) forms in the load. This operating regime is typical of explosive magnetic generators.[6] We shall first estimate the highest possible peaking coefficient of the magnetic field (current) pulse, i.e., the ratio of the time of the initial pulse in the flux compression space τ_1 to the pulse time in the load space τ_2. A circuit breaker in the metal phase should effectively shield the initial magnetic field pulse, for example $d>2\lambda_1$, where d is the barrier thickness and λ_1 is the skin layer depth for the initial magnetic field pulses. In the insulator phase the circuit breaker should transfer the magnetic field into the load space without appreciable losses within the time $\sim\tau_2$, for example $D<\lambda_2/2$, where λ_2 is the skin layer depth for the magnetic field pulse in the load space. Since $\lambda\propto(\tau/\sigma)^{1/2}$, we find that the maximum ratio τ_1/τ_2 is $1/16(\sigma_m/\sigma_i)\sim10$ at temperatures of 300–400 K. Here $\sigma_{m(i)}$ is the conductivity in the metal (insulator) phases. Let us assume that the initial magnetic field is 60 T which corresponds to a phase boundary velocity of 3 km/s. The depth of penetration of the magnetic field in the metal phase, as given by formula (12) is $\lambda=\nu_m/u_S\approx0.3$ cm. The magnetic field rise time in the load space can then be estimated as the time taken for the wave to cover a distance of the order of 2λ. From this it can be seen that the magnetic field (current) rise time in the load space may be a few microseconds.

These results can also be used to investigate a slowly varying magnetic field $\mathbf{B}_0$. The condition for which $\mathbf{B}_0$ can be considered to be slow is obtained from Eqs. (4) and (12)

$$\left|\frac{\partial B_0}{\partial t}\ll u_S\nabla B|_{x=+x_{MI}}=\frac{B_0u_S^2}{\nu_m}\right|. \tag{23}$$

In particular, for the example examined above we find $|\partial B_0/\partial t|\ll60$ T/μs, i.e., the initial magnetic field pulse rises slowly (usually no faster than 10 T/μs) for almost any explosive magnetic current generator.[6]

To conclude, in the analysis described above we assumed that the phase transition takes place fairly rapidly and we neglected its kinetics. However, we know that the time required for a first-order phase transition in solids may be fairly long if the transition is caused by rearrangement of the crystal structure. The metal–paramagnetic insulator phase transition in $(V_{1-X}Cr_X)_2O_3$ mainly involves the electronic subsystem and takes place without any change in the crystal structure so that this phase transition is obviously fast.

The author would like to than A. K. Zvezdin and colleagues at the Institute of General Physics of the Russian Academy of Sciences for taking part in discussions of the phase diagram of $(V_{1-X}Cr_X)_2O_3$ and related compounds, and also delegates at the Russian-American Seminar on the ''Behavior of Materials under High-Intensity Pulsed Loads'' for discussions of the propagation dynamics of a phase transition wave. The author also thanks A. E. Dubinov and I. V. Makarov for valuable discussions and also J. Brooks and V. Levis for invaluable support.

This work was partially financed by the International Scientific-Technical Center (Project No. 829).

[1] N. F. Mott, *Metal–Insulator Transitions* (Barnes and Noble, New York, 1974; Nauka, Moscow, 1979, pp. 838–858).
[2] É. L. Nagaev, Usp. Fiz. Nauk **166**, 838 (1996).
[3] A. A. Bugaev, B. P. Zakharchenya, and F. A. Chudnovskiĭ, *Metal–Semiconductor Phase Transition and Its Applications* [in Russian], Nauka, Leningrad (1979).
[4] H. Kuwamoto, J. M. Honig, and J. Appel, Phys. Rev. B **22**, 2626 (1980).
[5] Yu. B. Kudasov, Fiz. Tverd. Tela (St. Petersburg) **38**, 1335 (1996) [Phys. Solid State **38**, 739 (1996)].
[6] A. I. Pavlovskiĭ, and R. Z. Lyudaev, in *Problems of Modern Experimental and Theoretical Physics*, edited by A. P. Aleksandrov [in Russian], Nauka, Leningrad (1984), pp. 206–270.
[7] *Handbook of Thermodynamic Properties of Individual Substances*, edited by V. P. Glushko Vol. 4, Book 2 [in Russian], Nauka, Moscow (1982), p. 67.
[8] D. B. McWhan, A. Menth, J. P. Remeika *et al.*, Phys. Rev. B **7**, 1920 (1973).
[9] *Handbook of the Thermal Conductivity of Solids*, edited by A. S. Okhotin *et al.* [in Russian], Energoatomizdat, Moscow (1984), p. 256.
[10] D. B. McWhan, J. P. Remeika *et al.*, Phys. Rev. B **2**, 3734 (1970).

Translated by R. M. Durham

Influence of microwave treatment on the electrophysical characteristics of technically important semiconductors and surface-barrier structures

A. A. Belyaev, A. E. Belyaev, I. B. Ermolovich, S. M. Komirenko, R. V. Konakova, V. G. Lyapin, V. V. Milenin, and E. A. Solov'ev

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

M. V. Shevelev

E. O. Paton Institute of Electrical Welding, National Academy of Sciences of Ukraine, 252005 Kiev, Ukraine

(Submitted September 2, 1997)
Zh. Tekh. Fiz. **68**, 49–53 (December 1998)

An investigation is made of changes in the electrophysical parameters of narrow-gap ($Cd_xHg_{1-x}Te$ $x=0.22$–0.24) and wide-gap (gallium arsenide, indium and gallium phosphides) semiconductor materials and Schottky-barrier diode structures based on these materials, stimulated by microwave electromagnetic radiation. It is shown that the parameters of materials and device structures may be improved by defect gettering. An analysis is made of possible mechanisms for the interaction between microwave radiation and the objects being studied.
 [S1063-7842(98)00912-X]

INTRODUCTION

The operating conditions of basic semiconductor elements envisage a certain level of stability to various types of irradiation, including ionizing radiation and electromagnetic radiation.[1] Since the fifties, an enormous number of theoretical and experimental studies have been devoted to the influence of ionizing radiation on semiconductor materials and devices. In contrast, the action of electromagnetic radiation, especially microwave, has not been considered so extensively. However, it is known that in some cases, microwave radiation acting on basic structures and complete devices (diodes, transistors, integrated circuits) causes catastrophic failure.[2] On the other hand, recent reports have indicated defect gettering effects and structural relaxation stimulated by microwave radiation in semiconductor materials.[3,4]

The aim of the present study is to investigate the influence of microwave radiation on the electrophysical characteristics of narrow-gap ($Cd_xHg_{1-x}Te$) and wide-gap (gallium arsenide, indium and gallium phosphides) solid semiconductor materials and surface-barrier structures using these.

SAMPLES AND METHODS

The samples were: 1) single-crystal wafers of $Cd_xHg_{1-x}Te$ ($x=0.21$–0.24), gallium arsenide, and indium and gallium phosphides with a free electron concentration of $\sim 10^{16}$–$2\times 10^{17}\,cm^{-3}$ and 2) Schottky-barrier device structures formed in a vacuum of 10^{-4} Pa by condensation of chromium on (100) surfaces of gallium arsenide and indium and gallium phosphides, and also by condensation of molybdenum, tungsten, platinum, aluminum, Au–Ti, and titanium nitride on the (100) surface of gallium arsenide. The thickness of the metal layer was 80–100 nm.

The samples were irradiated in a magnetron in the centimeter wavelength range (free-space irradiation regime).[5] The time of exposure was varied between 1 and 60 s for different samples. The generator output power was 5 kW. The intensity of the microwave irradiation was varied by varying the distance between the irradiated object and the waveguide exit or by varying the exposure time with the distance from the object kept constant. Before and after the irradiation of solid samples at temperature $T=77$ K, measurements were made of the photoluminescence in the spectral range 0.6–2.0 eV excited by a powerful PZh-100 incandescent lamp with $h\nu > 2.0$ eV. The Hall effect and the minority carrier lifetime $\tau_{p'}$ were also measured. For the surface-barrier structures, measurements were made of the current–voltage and capacitance–voltage characteristics and the diffusion length of the minority carriers $L_{p'}$. Auger electron spectra and distribution profiles of the components in the metal–semiconductor contact were also measured.

RESULTS OF MEASUREMENTS AND DISCUSSION

These experiments show that the results of the microwave irradiation of $Cd_xHg_{1-x}Te$ crystals depend strongly on the initial state of the samples. For instance, in the initial sample No. 1 (Table I) the presence of inhomogeneities capable of producing anomalies on the temperature dependence of the Hall constant $R_x(T)$ (Fig. 1), shows up after treatment as a change in the type of conductivity and an abrupt change in the minority carrier lifetime τ_p (Fig. 2). However, for fairly homogeneous samples the nature of the changes differs qualitatively. Microwave treatment of such samples (Nos. 2 and 3 in Table I) causes an appreciable increase in τ_p. Moreover, R_x shows a slight increase in the range of impurity conductivity while the mobility μ_p decreases slightly. The data obtained from the temperature dependences of τ_p in this

TABLE I. Influence of microwave irradiation on the electrophysical parameters of $Cd_xHg_{1-x}Te$ (ideality factor n, mobility μ_p and lifetime τ_p of minority carriers, energy E_t and concentration N_t of impurity centers) at $T=77$ K.

Sample No.	Composition x	Type of treatment	n, cm^{-3}	μ_p, cm^2/Vs	τ_p, μs	E_t, eV	N_t, cm^{-3}
1	0.24	Initial	6.9×10^{15}	2×10^4	0.61		
		Microwave	2.54×10^{16}	6×10^3	0.246		
2	0.21	Initial	5.12×10^{14}	2×10^5	1.4	0.1	1.1×10^{14}
		Microwave	4.9×10^{14}	1.5×10^5	2.4	0.1	3.6×10^{13}
3	0.22	Initial	4.82×10^{14}	1.4×10^5	2.1	0.07	6.25×10^{13}
		Microwave	4.6×10^{14}	9.6×10^4	3.0	0.07	2.5×10^{13}

case indicate some reduction in the concentration of recombination-active centers having energy levels in the upper half of the band gap.

In order to envisage a possible mechanism for the influence of microwave treatment on the defect structure of these crystals, we shall consider nonuniform heating of the sample by absorption of microwave energy. In other words, we shall assume that the main mechanism responsible for the dissipation of absorbed microwave radiation energy is heating which stimulates the displacement of defects to sinks. These sinks may be dislocations, small-angle boundaries, or large clusters of point defects. This process is responsible for a slight increase in τ_p and R_x (see samples Nos. 2 and 3 in Table I). However, in initially inhomogeneous samples (sample No. 1 in Table I) containing additional conduction channels, microwave treatment intensifies the impurity atmospheres of the defects which are responsible for these channels. They become sufficiently powerful to abruptly change the type of conductivity and the lifetime of the minority carriers.

Thus, these results indicate that an important role must be ascribed to transformations of the crystal point-defect structure, which may be stimulated by microwave irradiation and diffusion of recombination-active point defects and impurities to sinks.

Data obtained by studying the photoluminescence spectra of gallium arsenide and indium and gallium phosphides also indicate that microwave irradiation has an appreciable influence on the point defect spectrum in crystals. For instance, the initial photoluminescence spectra of tin-doped gallium arsenide (Fig. 3) show two overlapping bands with maxima at $h\nu_{max}=1.15$–1.20 and 0.993-1.01 eV. As a result of irradiation for $t=6$ s, the position of the peaks becomes the same in all the samples (1.185 and 1.01 eV) with the 1.185 eV band being preferentially amplified. The bands become narrower which evidences some structural-impurity ordering of the surface layer of gallium arsenide. For $t\geqslant60$ s the intensities of both bands are reduced and they become broader, which indicates that degradation processes are intensified in the gallium arsenide. Similar changes in the 1.410, 1.150, and 0.820 eV bands after exposure to microwave radiation for 1–40 s are also observed for indium phosphide (100) single crystals (Fig. 4).

Gallium arsenide crystals initially doped with tellurium, which contain more defects, have a single band with $h\nu_{max}=1.20$ eV which remains almost unchanged after irradiation for 1–60 s. Slight changes in the structure of the local centers are observed in gallium phosphide single crystals as a result of microwave irradiation at the same exposures.

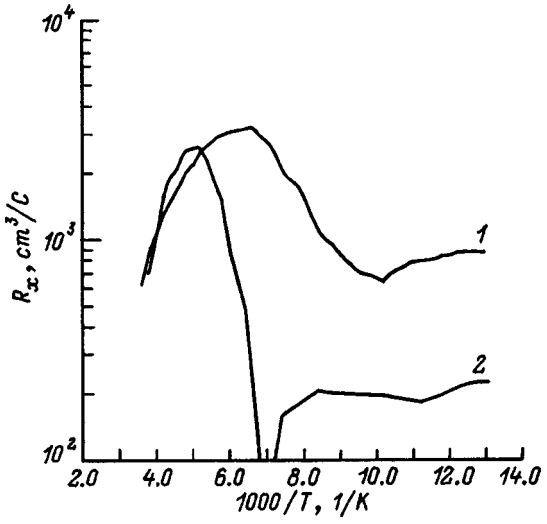


FIG. 1. Temperature dependence of the Hall constant of $Cd_{0.24}Hg_{0.76}Te$ samples: *1* — initial sample, *2* — after microwave irradiation for 5 s.

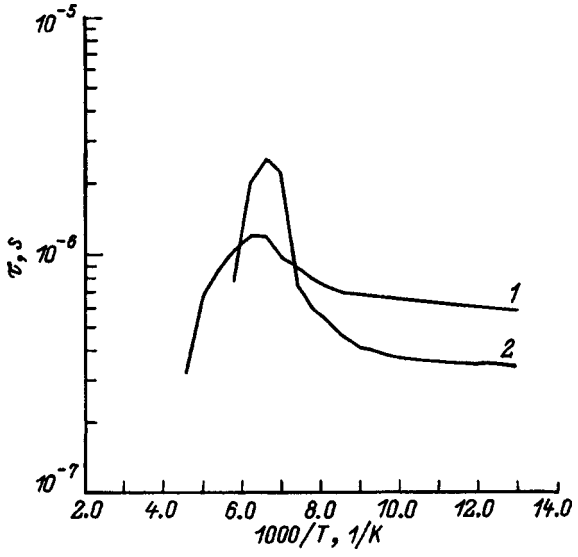


FIG. 2. Temperature dependence of the minority carrier lifetime in $Cd_{0.21}Hg_{0.79}Te$ samples (*1*, *2* — as in Fig. 1).

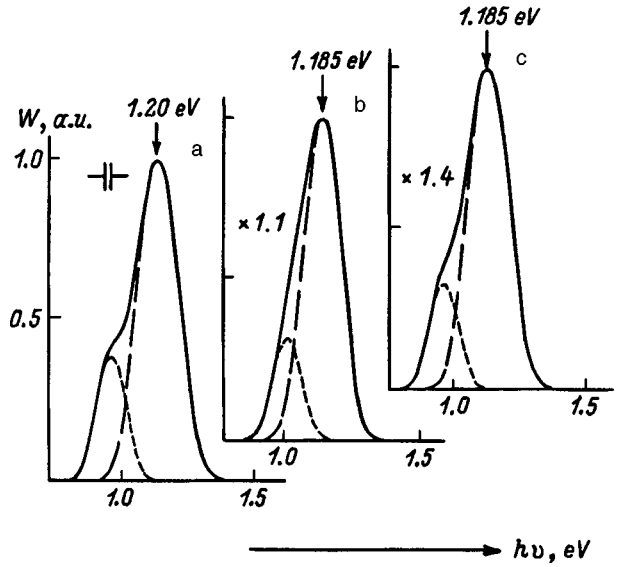


FIG. 3. Photoluminescence spectra of *n*-GaAs:Sn samples: a — initial sample, b, c — after microwave irradiation for 6 (62) s.

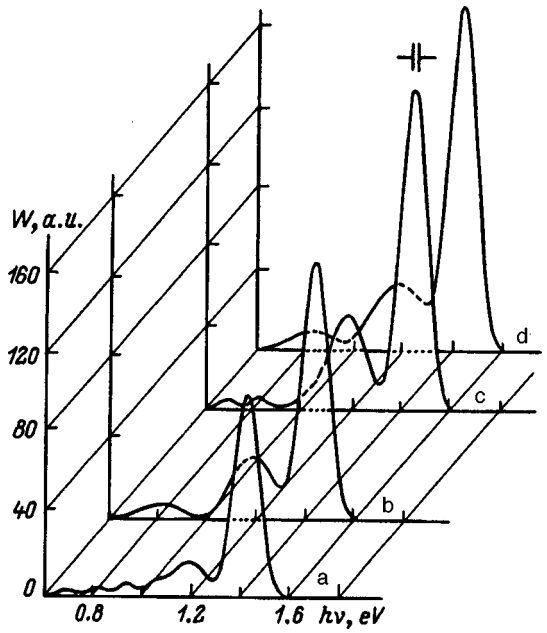


FIG. 4. Photoluminescence spectra of samples in *n*-InP(111): a — initial sample, b–d — after microwave irradiation for 10, 30, and 40 s.

Table II gives the results of microwave irradiation for 1–2 s on the electrophysical parameters of Schottky surface-barrier diodes. Table II shows that as a result of microwave irradiation, the properties of the interphase boundaries and the surface layers of the semiconductor changed appreciably, and this was accompanied by changes in the main parameters of the barrier. Specifically, the barrier height φ_B increased, the ideality factor n decreased, and the minority carrier diffusion length L_p increased. The latter is indicative of gettering processes stimulated by the microwave radiation in the surface layers of the gallium arsenide and the indium and gallium phosphides.

Figure 5a shows typical current–voltage characteristics of a surface-barrier diode. It can be seen that after microwave irradiation for 2 s, the length of the exponential section of the forward branch of the characteristic increased by an order of magnitude, the barrier height φ_B remained almost the same, and the ideality factor n decreased. The return branch of the characteristic (Fig. 5b) revealed an appreciable decrease in the return current I_R. The slope of the capacitance–voltage characteristic (Figs. 5c and 5d) remained almost unchanged. This indicates that as a result of the action of microwave radiation the dopant concentration is the same as that in the initial material.

The observed changes in the parameters of the diode structures are a consequence of structural-chemical rear-

TABLE II. Influence of microwave electromagnetic irradiation on the electrophysical parameters of Schottky diodes.

Type of structure	Type of treatment	Barrier height, φ_B eV	Ideality factor n	Diffusion length of minority carriers L_p, μm
Cr−GaAs	Initial	0.73−0.75	1.17−1.24	1.2−1.4
	Microwave	0.76−0.77	1.08−1.09	1.5−1.7
Mo−GaAs	Initial	0.68−0.69	1.16−1.23	2.3−2.8
	Microwave	0.68−0.69	1.09−1.14	2.5−2.7
W−GaAs	Initial	0.65−0.66	1.20−1.40	1.7−2.0
	Microwave	0.69−0.70	1.09−1.12	2.1−2.2
Pt−GaAs	Initial	0.88−0.95	1.12−1.37	2.1−2.2
	Microwave	0.88−0.89	1.18−1.24	2.1−2.2
Al−GaAs	Initial	0.55−0.58	1.68−2.20	1.6−1.9
	Microwave	0.57−0.58	1.30−1.40	2.0
Au−Ti−GaAs	Initial	0.70−0.76	1.3−1.4	1.7−1.8
	Microwave	0.70−0.76	1.32−1.33	1.85−1.92
TiN−GaAs	Initial	0.70−0.75	1.24−1.35	1.60−1.75
	Microwave	0.70−0.75	1.08−1.10	1.80−1.82
Cr−InP	Initial	0.67−0.69	1.50−1.80	1.52−1.60
	Microwave	0.63−0.65	1.20−1.40	1.70−1.75
Ga−GaP	Initial	1.63−1.67	1.50−1.72	0.53
	Microwave	1.85−1.88	1.32−1.40	0.87

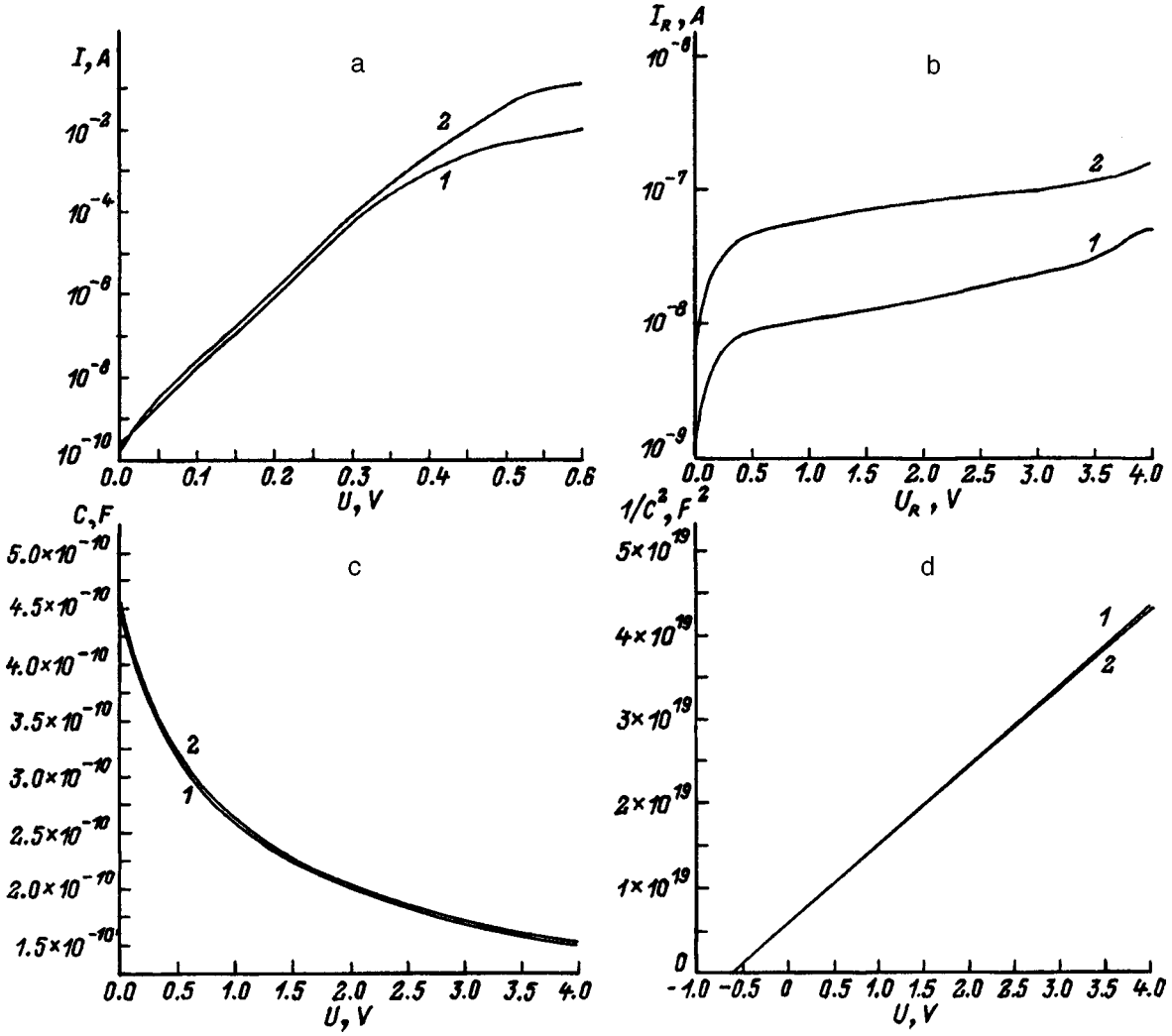


FIG. 5. Current–voltage and capacitance–voltage characteristics of a TiN–GaAs Schottky diode: a, b — forward (return) branch of current–voltage characteristic, c — capacitance–voltage characteristic, d — dependence of $1/C^2$ on V; *1* — initial sample and *2* — after microwave irradiation for 2 s.

rangement of the interphase boundaries at the metal–semiconductor contact, stimulated by the microwave radiation. In order to identify the characteristics of the interphase interactions, we analyzed the Auger profiles of the metal–semiconductor components for contacts formed by platinum, titanium, and aluminum with gallium arsenide. It was found that for each of these contacts we could identify characteristics not found in the other contacts. For example, for a Pt–GaAs contact the interphase interactions are described by the reaction

$$\mathrm{Pt+GaAs\rightarrow PtAs_2+(PtGaAs)},$$

whereas the following substitution reaction is typical of an Al–GaAs contact

$$y\mathrm{Al+GaAs\rightarrow}(y-x)\mathrm{Al+Al}_x\mathrm{Ga}_{1-x}\mathrm{As}+x\mathrm{Ga},$$

although these reactions require comparatively high temperatures which are not attainable in our experiments and also times longer than 1 s. In Au–Ti–GaAs structures the predominant process was microwave-induced mixing of the contact components with an appreciable fraction of the titanium oxide phase present after the microwave treatment. These interactions are clearly responsible for the transition from highly unstructured heterogeneous metal–semiconductor interfaces to more homogeneous ones and this promotes the changes observed in the parameters of basic structures.

We consider other possible mechanisms responsible for the changes observed in the properties of metal–semiconductor contacts.

1. A thermal mechanism caused by heating as a result of the absorption of microwave energy.[6] An analysis of the distribution profiles of the structure components in the contacts before and after microwave irradiation and a comparison with the results of a layer-by-layer analysis of thermally annealed contacts indicates that this factor has little influence.

2. An electrostatic mechanism caused by the real voltage drop at the barrier.[6] Even in the absence of critical electric fields which determine the avalanche and tunnel breakdown mechanisms, this mechanism may have a substantial influence on the diffusive redistribution of the contact components, as was shown in Ref. 6. Intensive interdiffusion between the metal and gallium arsenide is initiated at an

absorbed power level which is approximately 2/3 of the critical breakdown power. However, estimates made in Ref. 6 indicate that this factor does not play an important role under our irradiation conditions.

3. An electrodynamic mechanism caused by departure of the electronic subsystem of the semiconductor from thermodynamic equilibrium as a result of the generation of hot carriers in the surface layer of the semiconductor and an ensuing change in the impurity-defect composition of this layer.[7] The similarity between the electrophysical parameters of surface-barrier structures exposed to microwave irradiation and ^{60}Co gamma quanta (where this mechanism predominates) indicates that this mechanism plays a dominant role in structural-impurity defect rearrangement process under microwave irradiation.[8]

4. The formation of transient elastic-stress gradients caused by the almost instantaneous heating of disordered regions of the semiconductor created during the formation of the contact structure. In this case, the distribution profiles of the components of a contact pair before and after irradiation may remain almost constant. However, taking into account the possibility of collective interactions in elastic stress fields, the barrier for defect annihilation or rearrangement of defect complexes can be reduced substantially.[9,10]

CONCLUSIONS

A study has been made of the influence of microwave irradiation in the centimeter wavelength range on the electrophysical properties of the solid semiconductors $Cd_xHg_{1-x}Te$ ($x=0.22-0.24$), gallium arsenide, and indium and gallium phosphides, and also Schottky-barrier structures based on gallium arsenide and indium and gallium phosphides. It has been shown that the parameters of the materials and diode structures may be improved by gettering of defects stimulated by microwave irradiation. Specifically, the minority carrier lifetime increases, the return current decreases, and the spread of the Schottky barrier parameters is reduced (barrier height and ideality factor).

Further investigations are required before the processes taking place in microwave-irradiated semiconductors and devices based on these can be determined for certain. However the changes observed in the materials and at phase boundaries suggest that this irradiation may be used for technological purposes.

[1] L. O. Myrova and A. Z. Chepizhenko, *Ensuring the Stability of Communications Equipment to Ionizing and Electromagnetic Radiation* [in Russian], Radio i Svyaz', Moscow (1988) 296 pp.
[2] S. B. Bludov, N. P. Gadetskiĭ, K. A. Kravtsov *et al.*, Fiz. Plazmy **20**, 712 (1994) [Plasma Phys. Rep. **20**, 643 (1994)].
[3] D. E. Abdurakhimov, V. L. Vereshchagin, V. P. Kalinushkin *et al.*, Kratk. Soobshch. Fiz. No. 6, 27 (1991).
[4] T. G. Kryshtab, P. M. Lytvin, M. A. Masin, and I. V. Prokopenko, Metal Phys. Adv. Technol. **19**(3), 71 (1997).
[5] V. V. Milenin, R. V. Konakova, V. A. Statov *et al.*, Pis'ma Zh. Tekh. Fiz. **20**(4), 32 (1994) [Tech. Phys. Lett. **20**, 274 (1994)].
[6] V. V. Antipin, V. A. Golovitsyn, D. V. Gromov *et al.*, Zarub. Radioélektron. No. 1, 37 (1995).
[7] I. K. Sinishchuk, G. E. Chaĭka, and F. S. Shishiyanu, Fiz. Tekh. Poluprovodn. **19**, 674 (1985) [Sov. Phys. Semicond. **19**, 415 (1985)].
[8] O. Ya. Borkovskaya, N. L. Dmitruk, R. V. Konakova *et al.*, in *Submicron Devices, Proceedings of the NATO Advanced Study Institute*, edited by A.-P. Jauho and E. V. Buzaneva (Kluwer, Dordrecht, 1996), pp. 386–389.
[9] V. D. Skupov and D. I. Tetel'baum, Fiz. Tekh. Poluprovodn. **21**, 1495 (1987) [Sov. Phys. Semicond. **21**, 910 (1987)].
[10] A. E. Belyaev, A. A. Belyaev, E. F. Venger *et al.*, in *Proceedings of the Sixth International Crimean Conference on Microwave Engineering and Communications Technologies* [in Russian], Veber, Sevastopol (1967), pp. 71–89.

Translated by R. M. Durham

Optical vortices in the scattering field of magnetic domain holograms

N. A. Groshenko, O. S. Makalish, and A. V. Volyar

Simferopol State University, 333036 Simferopol, Ukraine
(Submitted February 7, 1997)
Zh. Tekh. Fiz. **68**, 54–58 (December 1998)

Experimental investigations and theoretical-model analyses have been made of the magnetooptic diffraction of light at ferrite garnet magnetic films with a banded domain structure which includes isolated magnetic grating defects in the form of ''forks'' and ''breaks.'' An analysis of the magnetic grating structure and the light diffraction field shows that in terms of its action on laser radiation, a banded domain grating is similar to a computer-synthesized phase hologram of an isolated pure screw-wavefront dislocation. It is shown that as a result of magnetooptic diffraction at a magnetic hologram, optical vortices may be reconstructed with a helicoidal wavefront carrying the topological charge $l=\pm 1,\pm 2$. [S1063-7842(98)01012-5]

INTRODUCTION

A Gaussian beam incident on a magnetic film with a banded domain structure breaks down into various laser beams characterized by their direction of propagation and intensity. This effect is usually called magnetooptic diffraction.[1] At low beam intensities the magnetooptic effect is based on the Faraday magnetic effect (it is assumed that in the optical wavelength range the magnetic permeability[2] is $\mu=1$.)

The magnetooptic diffraction process at a regular magnetic grating was studied in detail in Ref. 2, but in structures with banded magnetic domains, irregularities are usually encountered both as a result of crystal lattice defects in the film or substrate and also as a result of irregularities in the distribution of the magnetization vector **M** in the structure of the magnetic film. Groshenko *et al.*[3] observed that these defects in the distribution of uncertainties of the magnetization vector gradient may be of two types. The first type of singularity is associated with isolated defects in the banded domain structure and corresponds to indeterminacy of the *y*-component of the rate of change of the magnetization vector at the domain walls. This type of isolated defect is observed as ''forks'' in the magnetic grating structure or as ''breaks'' in the domain bands. The second type of defect is associated with pile-ups of forks and breaks in the banded magnetic structure. This type of defect severely distorts the phase structure of the magnetic grating and is transitional between a banded structure and a labyrinth structure of domain films.

The aim of the present study was to investigate characteristic features of the field distribution of a laser beam scattered by a magnetic grating with isolated fork structural defects.

1. The samples were $Bi_xLu_{3x}(Fe,Ga)_5O_{12}$ films grown by liquid-phase epitaxy on Ga–Gd substrates. The film thickness was $5-7\ \mu$m and the substrate thickness $150\ \mu$m. The period of the magnetic grating was varied under the action of an external magnetic field which could also change the profile of the magnetic grating. In the absence of the magnetic field, regions having average linear dimensions of 1.5–2 mm and a regular distribution of banded magnetic domains formed in the film. The orientation of the bands in one region could be turned through an arbitrary angle relative to another. The regions were separated by a pile-up of magnetic defects. The profile of a regular magnetic grating in the absence of a magnetic field was close to a stepped profile.

Figure 1a shows a region of magnetic film with a regular distribution of banded domains having a single fork dislocation. For comparison Fig. 1b shows a pile-up of magnetic defects. The average period of the magnetic grating in the absence of an external magnetic field was $3-3.5\ \mu$m.

The magnetic film was placed in one arm of a Mach–Zehnder interferometer which was excited by linearly polarized $\alpha=0.63\ \mu$m He–Ne laser light (Fig. 2). A sideways-looking microscope was used to select a region with a single fork defect and a smooth Gaussian beam (TEM_{00} mode) was directed onto this region. As a result of diffraction in the scattering far field we observed a system of diffraction orders. As usual, we assumed that the order to the right of the zeroth order was positive and that to the left was negative. The diffraction pattern is shown in Fig. 3a which reveals clearly defined intensity minima in $m\neq 0$ diffraction orders. The position of this minimum is determined by the relative position of the fork defect and the axis of the Gaussian beam. The intensity minimum can be positioned centrally by rearrangement of the domain film. Interference summation of one of the scattered diffraction orders with the Gaussian beam in the interferometer forms a pattern containing an interference fork for the $m=+1$ diffraction order, which has two branches (Fig. 3b). The interference fork in the $m=-1$ diffraction order is turned through 180°. The interference pattern in Fig. 3c is formed by the summation of the $m=\pm 2$ diffraction order with the Gaussian beam. In this case, the interference fork has three branches. The presence of a fork in the interference pattern indicates that the topological charge l is present in the scattered light. According to Berry,[4] the topological charge will be characterized by the difference between the number of lines entering and leaving

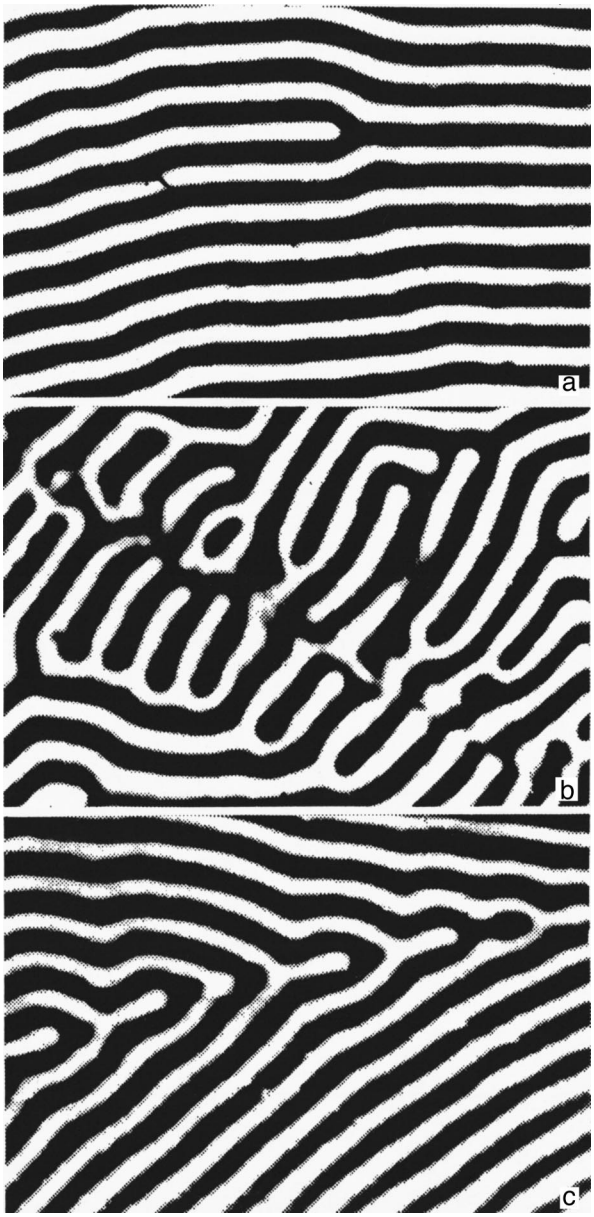


FIG. 1. Region of magnetic film with an isolated fork dislocation (a), pile-up of magnetic defects (b), and magnetic film with a complex fork dislocation (c).

the node of the fork. For example for the $m=\pm 1$ order the topological charge is $l=+1$. In general, for a magnetic grating we have $m=1$. If the reference and object beams have different wavefront curvatures in the Mach–Zehnder interferometer, the interference pattern is a spiral. The number of branches of the spiral is equal to the modulus of the topological charge, as shown in Figs. 3d and 3c.

2. We shall estimate the processes accompanying magnetooptic scattering at a magnetic grating singularity. Far from the singularity, the permittivity distribution in the plane of the film is characterized by the orientation of the magnetization vector relative to the plane of the film by means of a periodic function of the angle. For approximate calculations we can confine ourselves to a harmonic distribution of the permittivity over the area of the film. According to Born and Wolf,[5] the scattered light should only form two diffraction

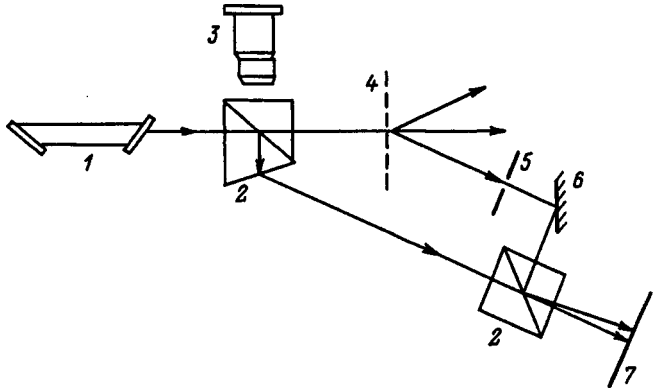


FIG. 2. Schematic of experimental apparatus: *1* — He–Ne laser, *2* — prisms, *3* — sideways-looking microscope, *4* — magnetic domain film, *5* — diaphragm, *6* — mirror, and *7* — screen.

orders at this type of inhomogeneity. We compare our experimental results with results[6,7] obtained in light diffraction experiments using a computer-synthesized phase hologram of a pure screw wavefront dislocation. Figure 4 shows a computer hologram calculated using the results reported in Ref. 6. Comparing the computer hologram (Fig. 4) with the magnetic grating (Fig. 1a), we find that light scattered by a magnetic grating with a single fork defect is similar in appearance to a wave carrying a pure screw wavefront dislocation. This field is known as a Laguerre–Gauss beam. The electric field strength of the mth order of a beam diffracted at the magnetic grating can then be expressed as a Laguerre–Gauss mode

$$E_l(\rho,\varphi,z)=E_l(\rho/\rho_0)^{|l|}\exp(il\varphi)\left(\frac{k\rho_0^2/2}{z+ik\rho_0^2/2}\right)^{(|l|+1)} \times\exp\left[i(|l|+1)\frac{\pi}{2}\right]\exp\left[\frac{-\rho^2}{\rho_0^2+2z/ik}\right] \quad (l=m), \tag{1}$$

where $\tan(\varphi)=y/x$ is the azimuthal coordinate of the beam, ρ is the beam radius, ρ_0 is the radius of the beam constriction, and k is the wave number.

We shall subsequently assume that a pure screw wavefront dislocation is present on the z axis if the following conditions are satisfied

$$\mathrm{Re}\,E_l(r,\varphi,z)=0, \quad \mathrm{Im}\,E_l(r,\varphi,z)=0. \tag{2}$$

The mathematical structure of the electric and magnetic fields near a pure screw dislocation of a wavefront with topological charge l can be given as

$$e\sim e_0\rho^{|l|}\exp(il\varphi), \quad h\sim h_0\rho^{|l|}\exp(il\varphi). \tag{3}$$

The equations (2) describe two surfaces whose intersection defines the line of the singularities or in our case, the line of propagation of a pure screw dislocation. Thus, this magnetic grating may be represented as a phase hologram whose transmission function in the xy plane is written as

$$U(x,y)=U_0-U_1\sin(kx-\varphi), \quad K=\frac{2\pi}{\Lambda}, \tag{4}$$

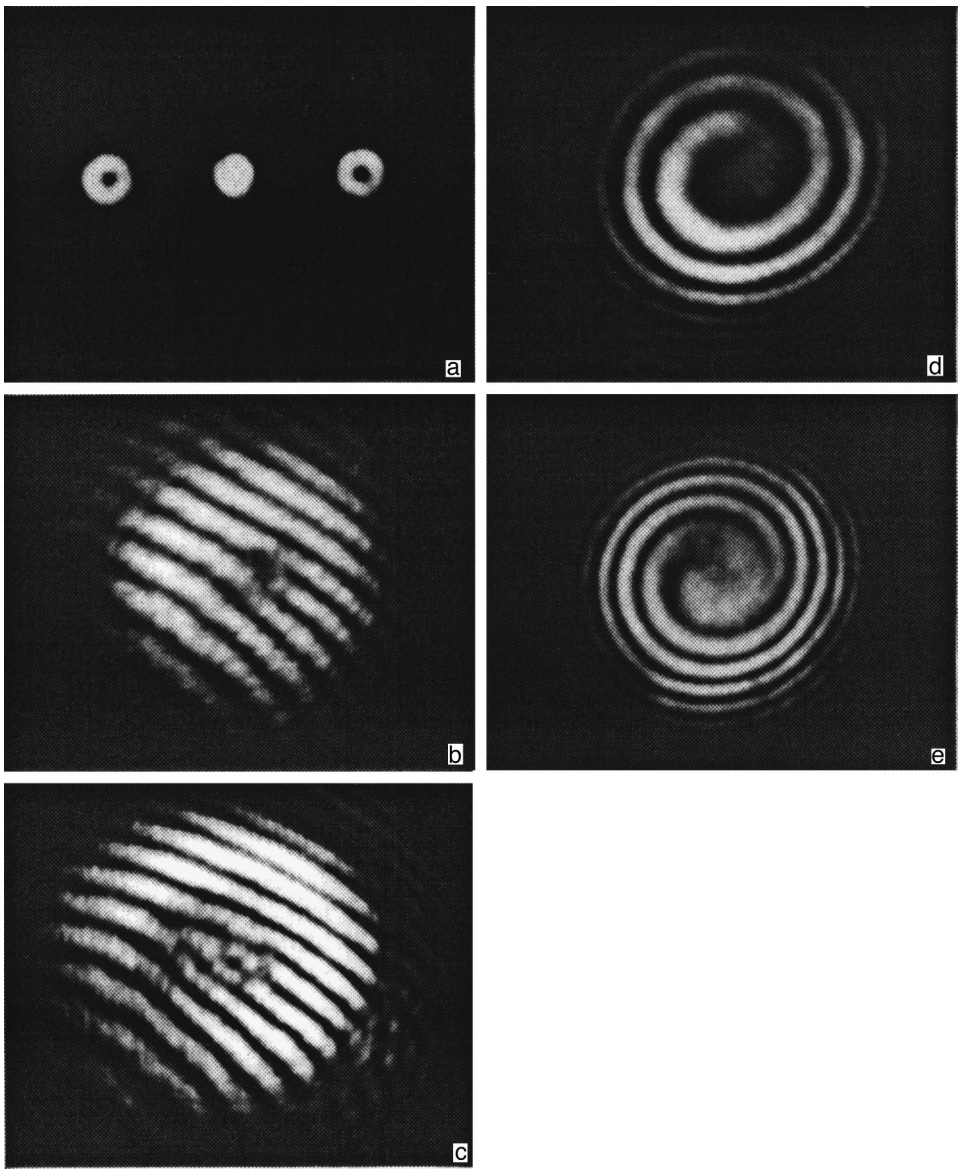


FIG. 3. Diffraction pattern (a), interference fork for $m=+1$ diffraction order (b), interference fork for $m=+2$ diffraction order (c), spiral interference pattern for $m=+1$ diffraction order (d), and spiral interference pattern for $m=+2$ diffraction order (e).

where Λ is the period of the magnetic grating, U_0 is the average transmission of the hologram, and $U_1(x,y)$ is the hologram profile.

The reconstructed beams at this hologram have a phase front in the form of a helicoidal surface shown in Fig. 5a. The spacing between the planes of the helicoid is equal to the wavelength of light. On the axis of the helicoid there is some indeterminacy of the wave phase. On passing through the point $x=y=0$ the phase jumps by π. When the helicoidal surface is crossed by a plane whose normal forms a certain angle to the axis of the helicoid, the lines of intersection describe lines of equal phase. In this case, these lines are quasiparallel bands near the singular point and form the immediate vicinity of a fork (Fig. 5b). If the helicoid intersects a system of concentric hemispheres, the projection of the line of intersection on the plane perpendicular to the axis of the helicoid is a spiral isophase line. In an interference experiment the isophase lines will correspond to the conditions of maximum intensity of the interference band.

From these investigations we can draw the following

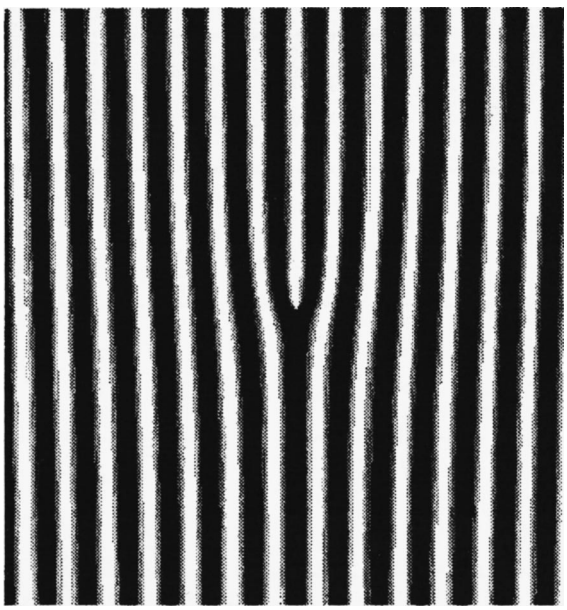


FIG. 4. Computer-synthesized hologram of pure screw wavefront dislocation.

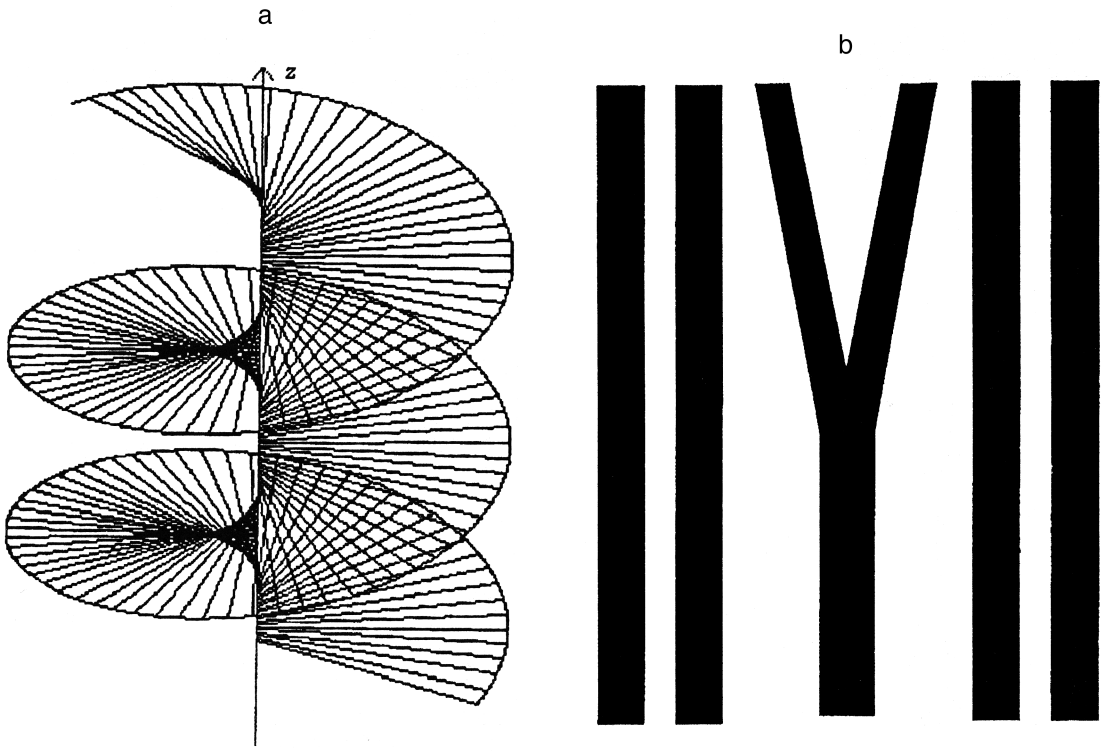


FIG. 5. Helicoidal surface (a) and quasiparallel bands near fork singular point (b).

conclusions. A magnetic domain grating with isolated singularities of the magnetization vector distribution is the phase hologram of an optical vortex. Although the diffraction efficiency of this hologram is low (in our experiment this was 2–3%), specific wave states, i.e., optical vortices with different topological charges, can be formed by means of magnetic holograms. This indicates that magnetic holograms have widespread applications in optoelectronics and optical communications. Unlike other types of light waves having no point singularities, the wavefront of an optical vortex generated by a magnetic hologram is a spatially continuous surface extending from the light source to the observer.

[1] Yu. F. Vilesov, V. G. Vishnevskiĭ, and N. A. Groshenko, Pis'ma Zh. Tekh. Fiz. No. 6, 1126 (1989) [*sic*].
[2] D. I. Simentsov and A. P. Gubarev, Opt. Spektrosk. **53**, 501 (1982) [Opt. Spectrosc. **53**, 295 (1982)].
[3] N. Groshenko, A. Volyar, and T. Fadeyeva, Proc. SPIE **2795**, 212 (1995).
[4] M. V. Berry, in *Physics of Defects*, Les Houches Summer School, Session XXXV (North Holland, Amsterdam, 1980), pp. 453–543.
[5] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, 1968; Nauka, Moscow, 1973).
[6] I. V. Basistiy, M. S. Soskin, and M. V. Vasnetsov, Opt. Commun. **119**, 604 (1955).
[7] V. Yu. Bazhenov, M. S. Soskin, and M. V. Vasnetsov, J. Mod. Opt. **39**, 985 (1992).

Translated by R. M. Durham

Three-dimensional diffuser and three-dimensional speckles

A. P. Vladimirov

Institute of Mechanical Engineering, Urals Branch of the Russian Academy of Sciences, 620219 Ekaterinburg, Russia
(Submitted September 24, 1997)
Zh. Tekh. Fiz. **68**, 59–63 (December 1998)

The well-known model of a diffusely scattering surface as a set of randomly distributed point scatterers is generalized to the case of a volume diffuser. Assuming that the coordinates of the coherent radiation source, the observer, and also the diffuser shape are arbitrary, a formula is obtained for the spatial correlation function of the scattered radiation intensity in a free field. Expressions to determine the transverse and longitudinal dimensions of the speckles are obtained for diffusers in the form of a rectangular parallelepiped and a cylinder. © *1998 American Institute of Physics.* [S1063-7842(98)01112-X]

INTRODUCTION

Various effects in speckle optics, speckle photography, speckle and holographic interferometry have been described using a model of a diffusely scattering object as a set of point scattering centers distributed over a certain surface and emitting coherent waves with random amplitudes and phases.[1–3] This model was used to study displacements and decorrelation of speckles caused by displacements and deformations,[4,5] and also by changes in the surface microrelief.[6] These data were used as the basis to develop real-time noncontact methods of determining displacements,[7] deformations,[8] and surface damage.[9]

It is a logical step to extend these methods[7–9] to determine the displacements, rotations, deformations, and damage in the bulk of a three-dimensional object. As a first step toward the development of such methods this model of a planar diffuser is generalized to a three-dimensional diffuser.

MODEL OF A THREE-DIMENSIONAL DIFFUSER

Let us consider a point source of coherent radiation at wavelength λ, having the coordinate $\mathbf{s}$ and illuminating point scatterers distributed randomly in a certain volume V near the origin, as shown in Fig. 1. We shall assume that the point scatterers are distributed fairly sparsely since no shading, screening, or multiple scattering are observed and waves from all scatterers with random amplitudes and phases are incident at an arbitrary point in space. We shall assume that the phases of the waves are uniformly distributed in the range between $-\pi$ and $+\pi$, the scatterers are uniformly distributed in the region V, and the probability density for the number of scatterers N obeys a Poisson distribution. We also assume that the refractive indices of the medium within the volume V and outside this volume are the same and are unity, and the random coordinates of the scatterers, the value of N, and the amplitudes and phases of an isolated wave and of different waves are statistically independent.

INTENSITY CORRELATION FUNCTIONS

The task for the first stage of our analysis is to calculate the spatial intensity correlation function of the scattered radiation at two arbitrary points in space $\mathbf{q}$ and $\mathbf{q}+\Delta\mathbf{q}$

$$B_{1,2}(\mathbf{q},\mathbf{q}+\Delta\mathbf{q})=\langle[I_1(\mathbf{q})-\langle I_1(\mathbf{q})\rangle]\times[I_2(\mathbf{q}+\Delta\mathbf{q})-\langle I_2(\mathbf{q}+\Delta\mathbf{q})\rangle]\rangle, \quad (1)$$

where $I_1(\mathbf{q})$ and $I_2(\mathbf{q}+\Delta\mathbf{q})$ are the intensities of the scattered radiation at the points $\mathbf{q}$ and $\mathbf{q}+\Delta\mathbf{q}$, respectively, and the angular brackets denote averaging over an ensemble of scatterers.

The intensity I at an arbitrary point $\mathbf{q}$ is given by

$$I(\mathbf{q})=A(\mathbf{q})A^*(\mathbf{q}),$$

where

$$A(\mathbf{q})=\sum_{j=1}^{N} a_j(\mathbf{q})$$

is the complex amplitude of the radiation at the point $\mathbf{q}$, $A^*(\mathbf{q})$ is the complex conjugate of $A(\mathbf{q})$, and $a_j(\mathbf{q})$ is the complex amplitude of a wave incident at point $\mathbf{q}$ from the jth scatterer.

It was shown in Ref. 2 that when $N\to\infty$ and assuming the properties of the amplitudes and phase described above, the combined probability density $\rho(A^r,A^i)$ of the real A^r and imaginary A^i parts of the complex amplitude is a circular Gaussian function with zero mean values $\langle A^r\rangle=\langle A^i\rangle=0$. We also note that $\langle A^rA^i\rangle=0$. For complex amplitudes which obey Gaussian statistics the intensity correlation function is equal to the square of the relative intensity modulus.[10] Thus, instead of the right-hand side of expression (1), we have

$$|\langle A(\mathbf{q})\,A^*(\mathbf{q}+\Delta\mathbf{q})\rangle|^2. \quad (2)$$

We then express $a_j(\mathbf{q})$ in the following form:

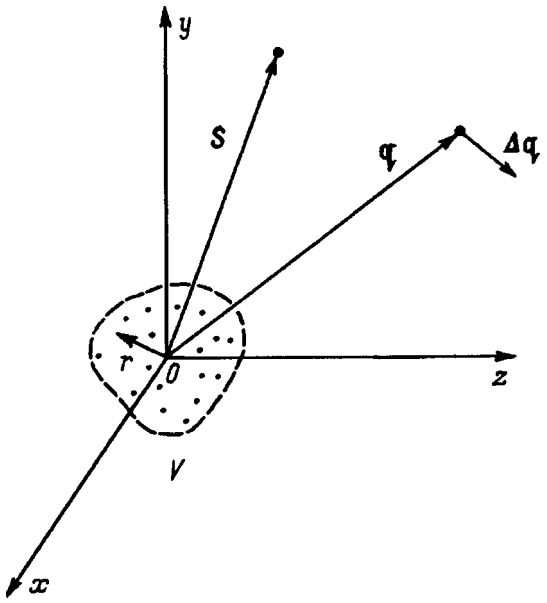


FIG. 1.

$$a_j(\mathbf{q})=[I_0(\mathbf{r}_j)]^{1/2}\xi_q(\mathbf{r}_j)\exp[i\varphi(\mathbf{r}_j)]\exp\{ik[|\mathbf{L}_s(\mathbf{r}_j)|+|\mathbf{L}_q(\mathbf{r}_j)|]\}=u_j\exp\{ik[|\mathbf{L}_s(\mathbf{r}_j)|+|\mathbf{L}_q(\mathbf{r}_j)|]\},$$

where $\mathbf{r}_j$ is the coordinate of the jth scatterer, $k=2\pi/\lambda$, $I_0(\mathbf{r}_j)$ is the intensity of the radiation illuminating the object at the point $\mathbf{r}_j$, $\xi_q(\mathbf{r}_j)$ is a coefficient which determines the fraction of the radiation originating from the jth scatterer to the point $\mathbf{q}$, $|\mathbf{L}_s(\mathbf{r}_j)|$ and $|\mathbf{L}_q(\mathbf{r}_j)|$ are the distances between point $\mathbf{s}$ and between point $\mathbf{r}_j$ and between point $\mathbf{r}_j$ and point $\mathbf{q}$, respectively, u_j and $\varphi(\mathbf{r}_j)$ are the complex amplitude and phase of the scattered wave in the immediate vicinity of the jth scatterer, respectively.

For the complex amplitude $A(\mathbf{q})$ we have

$$A(\mathbf{q})=\sum_{j=1}^{N} u_j\exp\{ik[|\mathbf{L}_s(\mathbf{r}_j)|+|\mathbf{L}_q(\mathbf{r}_j)|]\}. \tag{3}$$

Substituting Eq. (3) into Eq. (2), we have

$$B_{1,2}(\mathbf{q},\mathbf{q}+\Delta\mathbf{q})=\left|\int_{-\infty}^{+\infty}\cdots\int\sum_{j=1}^{N}\sum_{m=1}^{N}u_ju_m^*\exp\{ik[[\mathbf{L}_s(\mathbf{r}_j)|-|\mathbf{L}_s(\mathbf{r}_m)|]\}\exp\{ik[|\mathbf{L}_q(\mathbf{r}_j)|-|\mathbf{L}_{q+\Delta q}(\mathbf{r}_m)|]\}\times\rho(x_j,y_j,z_j,u_j,x_m,y_m,z_m,u_m^*,N)\,dx_1\ldots dN\right|^2, \tag{4}$$

where $\rho(x_j,\ldots,N)$ is the combined probability density of the random quantities $x_j,y_j,z_j,u_j,x_m,y_m,z_m,u_m^*$, N; $j=1,2,\ldots,N$; $m=1,2,\ldots,N$.

For the following calculations we shall assume that the complex amplitudes $u(\mathbf{r})$ of waves propagating from different scatterers are uncorrelated and the value of $\xi_q(\mathbf{r})$ is constant for all scatterers. Then, as in Ref. 2, we have

$$\langle u(\mathbf{r}_j)u^*(\mathbf{r}_m)\rangle=k_1[I_0(\mathbf{r}_j)I_0(\mathbf{r}_m)]^{1/2}\cdot\delta(\mathbf{r}_j-\mathbf{r}_m), \tag{5}$$

where k_1 is a coefficient and $\delta(\mathbf{r}_j-\mathbf{r}_m)$ is a three-dimensional delta function.

Bearing in mind the statistical independence of the random quantities appearing in formula (4) and expression (5), we have instead of the right-hand side of Eq. (4)

$$\left|k_1\int_{-\infty}^{+\infty}\cdots\int\sum_{j=1}^{N}I_0(\mathbf{r}_j)\exp\{ik[|\mathbf{L}_q(\mathbf{r}_j)|-|\mathbf{L}_{q+\Delta q}(\mathbf{r}_j)|]\}\times\rho(x_j,y_j,z_j)\rho(N)\,dx_1\ldots dN\right|^2, \tag{6}$$

Assuming that the size of the region V and the value of $|\Delta\mathbf{q}|$ are small compared with the distance between the scatterers and the point $\mathbf{q}$, we expand $|\mathbf{L}_q(\mathbf{r}_j)|$ and $|\mathbf{L}_{q+\Delta q}(\mathbf{r}_j)|$ as a Taylor series near the points $\mathbf{r}_j=0$ and $\mathbf{q}$, taking into account derivatives up to the third order inclusively. Omitting the subscript j, we have

$$\begin{aligned}
|\mathbf{L}_q(\mathbf{r})|&=[(x-q_x)^2+(y-q_y)^2+(z-q_z)^2]^{1/2}\\
&=[(x_1-q_x)^2+(x_2-q_y)^2+(x_3-q_z)^2]^{1/2}\\
&=f(x_1,x_2,x_3,q_x,q_y,q_z)=f(0,0,0,q_x,q_y,q_z)\\
&+\frac{1}{1!}\sum_{i=1}^{3}f'_{x_i}(x_1,x_2,x_3,q_x,q_y,q_z)\Big|_{(0,0,0,qx,q_y,q_z)}^{x_i}\\
&+\frac{1}{2!}\sum_{i=1}^{3}\sum_{m=1}^{3}f''_{x_ix_m}(x_1,x_2,x_3,q_x,q_y,q_z)\Big|_{(0,0,0,q_x,q_y,q_z)}^{x_ix_m}\\
&+\frac{1}{3!}\sum_{i=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3}f'''_{x_ix_mx_n}(x_1,x_2,x_3,q_x,q_y,q_z)\Big|_{(0,0,0,q_x,q_y,q_z)}^{x_ix_mx_n}\\
&=C_0+\sum_{i=1}^{3}C_i+\frac{1}{2}\sum_{i=1}^{3}\sum_{m=1}^{3}C_{im}+\frac{1}{6}\sum_{i=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3}C_{imn},
\end{aligned} \tag{7}$$

$$\begin{aligned}
|\mathbf{L}_{q+\Delta q}(\mathbf{r})|&=[(x-q_x-\Delta q_x)^2+(y-q_y-\Delta q_y)^2+(z-q_z-\Delta q_z)^2]^{1/2}\\
&=[(x_1-x_4)^2+(x_2-x_5)^2+(x_3-x_6)^2]^{1/2}\\
&=Q(x_1,x_2,x_3,x_4,x_5,x_6)\\
&=Q(a_1,a_2,a_3,a_4,a_5,a_6)\\
&+\frac{1}{1!}\sum_{i=1}^{6}Q'_{x_i}(x_1\ldots,x_6)\Big|_{(a_1\ldots,a_6)}^{(x_{i-a_i})}\\
&+\frac{1}{2!}\sum_{i=1}^{6}\sum_{m=1}^{6}Q''_{x_ix_m}(x_1,\ldots,x_6)\Big|_{(a_1\ldots,a_6)}^{(x_i-a_i)(x_m-a_m)}
\end{aligned}$$

$$+\frac{1}{3!}\sum_{i=1}^{6}\sum_{m=1}^{6}\sum_{n=1}^{6}Q'''_{x_i x_m x_n}(x_1\ldots,x_6)\Big|_{(a_1\ldots,a_6)}$$
$$\times(x_i-a_i)(x_m-a_m)(x_n-a_n)$$
$$=C_0+\sum_{i=1}^{6}C_i+\frac{1}{2}\sum_{i=1}^{6}\sum_{m=1}^{6}C_{im}$$
$$+\frac{1}{6}\sum_{i=1}^{6}\sum_{m=1}^{6}\sum_{n=1}^{6}C_{imn}, \tag{8}$$

where $(a_1,\ldots,a_6)=(0,0,0,q_x,q_y,q_z)$; q_x,q_y, and q_z are components of the vector $\mathbf{q}$; $\Delta q_x,\Delta q_y$ and Δq_z are components of the vector $\Delta\mathbf{q}$, and we also introduce the notation $x=x_1$, $y=x_2$, $z=x_3$, $q_x+\Delta q_x=x_4$, $q_y+\Delta q_y=x_5$, and $q_z+\Delta q_z=x_6$.

We now introduce the notation

$$\rho_q=(q_x^2+q_y^2+q_z^2)^{1/2},$$
$$l_x=\frac{q_x}{\rho_q},\quad l_y=\frac{q_y}{\rho_q},\quad l_z=\frac{q_z}{\rho_q}, \tag{9}$$

where ρ_q is the distance between the origin and the point $\mathbf{q}$, l_x,l_y, and l_z are components of the unit vector $\mathbf{l}$, directed from the point $\mathbf{r}=0$ to the point $\mathbf{q}$.

The explicit form of the functions $f(x_1,\ldots,q_z)$ and $Q(x_1,\ldots,x_6)$ was used in Ref. 11 to calculate the values of C_0, C_i, C_{im}, and C_{imn}. These values contain the variables x, y, and z to different powers, their products, the value of ρ_q, and also the components of the vectors $\mathbf{l}$ and $\Delta\mathbf{q}$. Substituting expressions (7) and (8) into (6) and combining terms with the same powers of the variables x_j,y_j, and z_j, we have:

$$\left|k_1\int_{-\infty}^{+\infty}\cdots\int\sum_{j=1}^{N}I_0(\mathbf{r}_j)\exp[-ik(a_{11}x_j^2+a_{22}y_j^2+a_{33}z_j^2\right.$$
$$+2a_{12}x_jy_j+2a_{13}x_jz_j+2a_{23}y_jz_j$$
$$\left.+2a_{14}x_j+2a_{24}y_j+2a_{34}z_j+a_{44})]\right.$$
$$\left.\times\rho(x_j,y_j,z_j)\rho(N)dx_1\ldots dN\right|^2$$
$$=\left|k_1\int_{-\infty}^{+\infty}\cdots\int\sum_{j=1}^{N}I_0(x_j,y_j,z_j)\right.$$
$$\times\exp[-ikF_0(x_j,y_j,z_j)]$$
$$\left.\times\rho(x_j,y_j,z_j)\rho(N)dx_1\ldots dN\right|^2, \tag{10}$$
$$j=1,\ldots,N$$

where $F_0(x_j,y_j,z_j)$ is the quadratic form of the equation $F_0(x_j,y_j,z_j)=0$, the values of the coefficients a_{ij} are also given in Ref. 11, and these coefficients contain ρ_q, and also the components of the vectors $\mathbf{l}$ and $\Delta\mathbf{q}$.

In expression (10) we introduce integration after the summation sign. Using the relation for the probability density

$$\int\int\int_{-\infty}^{+\infty}\rho(x,y,z)dxdydz=1$$

and assuming a uniform distribution of the random quantities x_j,y_j, and z_j in the region V, we have instead of expression (10)

$$\left|k_1\int_{-\infty}^{+\infty}\sum_{j=1}^{N}\left[\frac{1}{V}\int\int\int_V I_0(x_j,y_j,z_j)\right.\right.$$
$$\left.\left.\times\exp\{-ikF_0(x_j,y_j,z_j)\}dx_jdy_jdz_j\right]\rho(N)dN\right|^2. \tag{11}$$

Since the region of integration for all the integrals is the region V, all the integrals after the summation sign will be equal. Instead of expression (11), we finally have

$$\left|k_1\langle N\rangle\frac{1}{V}\int\int\int_V I_0(x,y,z)\exp[-ikF_0(x,y,z)]dxdydz\right|^2, \tag{12}$$

where $\langle N\rangle$ is the average number of scatterers in the volume V.

In general when the function $I_0(x,y,z)$ is given and the point $\mathbf{q}$ is selected arbitrarily, expression (12) can be calculated either by numerical integration of the integral therein or by reducing the quadratic form $F_0(x,y,z)$ to the canonical form, i.e., to a form containing no variables x, y, and z to the first power and the product x, y, and z. Such a transformation is implemented by successively transposing the origin and turning the axes relative to the origin. Such a transformation is clearly justified if, after this, the region V is converted into a new region convenient for integration.

PARTICULAR CASES

We shall analyze several particular cases for which simple calculations can be made. For the following transformations we shall assume that the region V is illuminated uniformly, i.e.,

$$I_0(x,y,z)=\begin{cases}I_0, & \text{if } x,y,z \text{ is in } V,\\ 0, & \text{otherwise.}\end{cases} \tag{13}$$

Let us assume that region V is illuminated in an arbitrary direction and the scattered radiation is observed along the z axis. Then we have

$$l_x=0,\quad l_y=0,\quad l_z=1,\quad a_{11}=\frac{\Delta q_z}{2\rho_q^2},$$
$$a_{22}=\frac{\Delta q_z}{2\rho_q^2},\quad a_{33}=a_{12}=0,$$
$$a_{13}=\frac{\Delta q_x}{\rho_q^2},\quad a_{23}=\frac{\Delta q_y}{\rho_q^2},$$
$$a_{14}=-\frac{\Delta q_x}{2\rho_q}+\frac{\Delta q_xq_z}{2\rho_q^2},\quad a_{24}=-\frac{\Delta q_y}{2\rho_q}+\frac{\Delta q_y\Delta q_z}{2\rho_q^2},$$

$$a_{34}=\frac{\Delta q_x^2+\Delta q_y^2}{4\rho_q^2},\quad a_{44}=\Delta q_z. \tag{14}$$

Let us assume that the points $\mathbf{q}$ and $\mathbf{q}+\Delta\mathbf{q}$ are located in a plane parallel to the plane $x0y$. Then we find $\Delta q_z=0$ and

$$F_0(x,y,z)=\frac{2\Delta q_x}{\rho_q^2}xz+\frac{2\Delta q_y}{\rho_q^2}yz-\frac{\Delta q_x}{\rho_q}x-\frac{\Delta q_y}{\rho_q}y+\frac{\Delta q_x^2+\Delta q_y^2}{2\rho_q^2}z=\left(\frac{2z}{\rho_q}-1\right)\times\left[\frac{\Delta q_x x+\Delta q_y y+(\Delta q_x^2+\Delta q_y^2)/4}{\rho_q}\right]+\frac{\Delta q_x^2+\Delta q_y^2}{4\rho_q}. \tag{15}$$

Taking the region V to be a parallelepiped of dimensions $2X_0\times 2Y_0\times 2Z_0$ with its center at the origin, and using relations (13) and (15), we have instead of expression (12)

$$\left|k_1\langle N\rangle I_0\exp\left[-ik\frac{\Delta q_x^2+\Delta q_y^2}{4\rho_q}\right]\frac{1}{2X_0 2Y_0 2Z_0}\times\int_{-X_0}^{+X_0}\int_{-Y_0}^{+Y_0}\int_{-Z_0}^{+Z_0}\exp\left\{-ik\left(\frac{2z}{\rho_q}-1\right)\times\left[\frac{\Delta q_x x+\Delta q_y y+(\Delta q_x^2+\Delta q_y^2)/4}{\rho_q}\right]\right\}dxdydz\right|^2. \tag{16}$$

In general, the three-dimensional integral in expression (16) can only be taken numerically. However, if the size of the diffusers $2Z_0$ and the value of ρ_q are such that $2z/\rho_q \ll 1$, this integral is easily taken. Then for the normalized intensity correlation function $\eta(\Delta q_x,\Delta q_y)$ we have

$$\eta(\Delta q_x,\Delta q_y)=\frac{\sin^2(\Delta q_x kX_0/\rho_q)}{(\Delta q_x kX_0/\rho_q)^2}\frac{\sin^2(\Delta q_y kY_0/\rho_q)}{(\Delta q_y kY_0/\rho_q)^2}. \tag{17}$$

If the region V is a cylinder of radius R and length $2Z_0$ centered at the origin, using expressions (12) and (13), the condition $2z/\rho_q \ll 1$, and converting to polar coordinates, we obtain an expression for the normalized intensity correlation function

$$\eta(\Delta r)=\frac{J_1^2(\pi\Delta rD/(\lambda\rho_q))}{(\pi\Delta rD/(\lambda\rho_q))^2}, \tag{18}$$

where $D=2R$, $\Delta r=(\Delta q_x^2+\Delta q_y^2)^{1/2}$, and $J_1(a)$ is a first-order Bessel function of the argument a.

Determining the intensity correlation intervals $\Delta\tilde{q}_x,\Delta\tilde{q}_y$, and $\Delta\tilde{r}$ from the conditions that the functions (17) and (18) decrease to zero the first time, we have

$$\Delta\tilde{q}_x=\frac{\lambda\rho_q}{X_0},\quad \Delta\tilde{q}_y=\frac{\lambda\rho_q}{Y_0},\quad \Delta\tilde{r}=1.07\frac{\lambda\rho_q}{D}.$$

We now return to the relations (14) and take the points $\mathbf{q}$ and $\mathbf{q}+\Delta\mathbf{q}$ along the z axis. Then we have $\Delta q_x,\Delta q_y=0$ and for $F_0(x,y,z)$ we have

$$F_0(x,y,z)=\frac{\Delta q_z(x^2+y^2)}{2\rho_q^2}+\Delta q_z. \tag{19}$$

Let us again assume that the region V is a rectangular parallelepiped of dimensions $2X_0\times 2Y_0\times 2Z_0$ centered at the origin. Substituting expression (19) into expression (12), using expression (13), and converting to the variables $s=x/X_0$ and $t=y/Y_0$, we have instead of expression (12):

$$\left|k_1 I_0\langle N\rangle\exp(-ik\Delta q_z)\frac{1}{\sqrt{\alpha}}\int_0^{\sqrt{\alpha}}\exp(-is^2)ds\times\frac{1}{\sqrt{\beta}}\int_0^{\sqrt{\beta}}\exp(-it^2)dt\right|^2=k_1^2\langle N\rangle^2 I_0^2|f(\alpha)f(\beta)|^2,$$

where

$$\alpha=\frac{\pi\Delta q_z X_0^2}{\lambda\rho_q^2},\quad \beta=\frac{\pi\Delta q_z Y_0^2}{\lambda\rho_q^2}.$$

Calculations made by the author showed that the function $g(\alpha,\beta)=|f(\alpha)f(\beta)|^2$ has a maximum at the origin and the contour lines for the expression $g(\alpha,\beta)$ are circles almost to the 0.1 level. The function reaches a first minimum when $g(\alpha,\beta)$ is approximately 0.08. We determine the intensity correlation interval from the condition that $g(\alpha,\beta)$ decreases to the 0.1 level. This occurs when $(\alpha^2+\beta^2)^{1/2}$ is five, to a high degree of accuracy. Thus, we have

$$\Delta\tilde{q}_z=\frac{5\lambda\rho_q^2}{\pi(X_0^4+Y_0^4)^{1/2}}. \tag{20}$$

Expression (20) determines the characteristic longitudinal dimension of the nonuniformity of the scattered radiation for a diffuser having a rectangular cross section of $2X_0\times 2Y_0$. In particular, for a diffuser with a square cross section with $X_0=Y_0=L_0$ we have

$$\Delta\tilde{q}_z=1.13\frac{\lambda\rho_q^2}{L_0^2}.$$

We shall estimate the longitudinal dimension of the speckles for a diffuser in the form of a cylinder of radius R and length $2Z_0$ with its center at the origin. Let us assume that the axis of the cylinder coincides with the Z axis. Converting to polar coordinates, using expressions (19), (13), and (12), we obtain for the normalized intensity correlation function

$$\eta(\Delta q_z)=\frac{\sin^2(k\Delta q_z R^2/(2\rho_q^2))}{[k\Delta q_z R^2/(2\rho_q^2)]^2}.$$

Determining the intensity correlation interval $\Delta\tilde{q}_z$ from the condition that the function $\sin x/x$ decreases to zero the first time, we have

$$\Delta\tilde{q}_z=\frac{\lambda\rho_q^2}{R^2}. \tag{21}$$

Formula (21) determines the characteristic longitudinal dimension of the speckles in the free field for a diffuser whose cross section is a circle of radius R.

It should be noted that for these diffusers the formulas for the characteristic transverse and longitudinal dimensions of the speckles remain the same for $2Z_0 \to 0$, i.e., on going to the limiting case of a planar diffuser. These formulas are the same as those familiar from the literature[2,4,12] which were derived for corresponding planar diffusers.

CONCLUSION

The well-known model of a diffusely scattering surface has been generalized to the case of a volume diffuser. It has been assumed that the size of the diffuser is small compared with the distances from the coherent radiation source and the observation points. For arbitrary directions of illumination and observation, intensity distribution of the illuminating radiation, and diffuser shape, a formula has been derived for the spatial intensity correlation function of the scattered radiation in the free field. For three-dimensional diffusers in the form of a rectangular parallelepiped and a cylinder, formulas were obtained for the transverse and longitudinal dimensions of the speckles. It has been shown that on going to the limit of two-dimensional scatterers, these formulas are the same as those familiar from the literature.

[1] L. J. Goldfisher, J. Opt. Soc. Am. **55**, 247 (1965).
[2] J. W. Goodman, in *Laser Speckle and Related Phenomena*, edited by J. C. Dainty (Springer-Verlag, New York, 1975), pp. 10–75.
[3] R. Jones and C. M. Wykes, *Holographic and Speckle Interferometry* (Cambridge University Press, Cambridge, 1983; Mir, Moscow, 1986, 328 pp.).
[4] V. V. Anisimov, S. M. Kozel, and G. R. Lokshin, Opt. Spektrosk. **27**, 483 (1969) [Opt. Spectrosc. **27**, 258 (1969)].
[5] I. Yamaguchi, Opt. Acta **28**, 1358 (1981).
[6] A. P. Vladimirov, in *Holography: Theoretical and Applied Problems* Collection of Scientific Papers [in Russian], Physicotechnical Institute, Academy of Sciences of the USSR, Leningrad (1988), pp. 100–106.
[7] I. Yamaguchi and T. Fujita, Proc. SPIE **1162**, 213 (1989).
[8] I. Yamaguchi, Opt. Eng. (Bellingham) **21**, 436 (1982).
[9] A. P. Vladimirov, Dissertation for Candidate's Degree [in Russian], Sverdlovsk (1989), 106 pp.
[10] E. Wolf, Philos. Mag. **2**, 351 (1957)
[11] A. P. Vladimirov, Deposited paper No. 1856-B96 [in Russian], VINITI Moscow (1996), 19 pp.
[12] M. Françon, *Laser Speckle and Applications in Optics*, trans. from the French (Academic Press, New York, 1979; Mir, Moscow, 1978, 171 pp.].

Translated by R. M. Durham

Multiwave excimer lamps using XeF/XeCl/KrF/KrCl molecules

A. K. Shuaibov

Uzhgorod State University, 294660 Uzhgorod, Ukraine, Russia
(Submitted October 6, 1997)
Zh. Tekh. Fiz. **68**, 64–67 (December 1998)

Results are presented of an investigation of the spectral and temporal characteristics of an electric-discharge excimer lamp emitting simultaneously on the 351 nm XeF, 308 nm XeCl, 249 nm KrF, and 222 nm KrCl transitions. The He/Kr/Xe/SF_6/HCl working mixtures were excited in a transverse discharge with ultraviolet spark preionization at a total pressure of 25–100 kPa. In order to obtain the same brightness the concentration ratio [Kr]/[Xe] for the excimer molecular transitions was 8/0.8 kPa, and that for the halogen-containing molecules was [SF_6]/[HCl][=[0.06–0.12]/[0.08–0.16] kPa. The duration of the radiation pulses for the excimer molecular transitions at atmospheric pressure was 100–200 ns and when the total mixture pressure was reduced to 250 kPa, this was doubled or trebled. The service life of the spontaneous radiation for B–X transitions in excimer molecules was 10^5 pulses. © *1998 American Institute of Physics.* [S1063-7842(98)01212-4]

INTRODUCTION

Sources of high-intensity spontaneous radiation using rare-gas monohalides (RX) together with corresponding laser emitters at $\lambda = 190–350$ nm are finding widespread applications in microelectronics, high-energy chemistry, biology, and medicine. Basov[1] suggested using pulsed excimer lamps to pump gas lasers. The design of these sources is simpler than laser ones and where there is no need to use the high spatial or spectral characteristics of RX ultraviolet radiation, and also if there is no need to transmit the radiation over large distances, these sources can successfully compete with excimer lasers.

So far, attention has been focused on studying suitable laser media and little importance has been attached to studying pulsed spontaneous RX radiation. Different types of pulsed volume discharges, including transverse ultraviolet-preionized discharges, have been effectively used to produce the active medium in excimer lamps.[2–7] These active media are He(Ne)/R/M mixtures, where R=Ar, Kr, Xe and M=F_2, NF_3, HCl, BCl_3, and so on, having compositions similar to those used in electric-discharge RX lasers but the range of working pressures is substantially broader 1–300 kPa. These electric-discharge excimer lamps emit at a single wavelength and multiwave operation when the active medium contains two working rare gases (such as Kr and Xe) and two different halogen carriers, has not been studied. Operating regimes with a single halogen carrier and different working rare gases have usually been studied for multiwave electric-discharge RX lasers.[8,9] Atezhev *et al.*[10] noted that the most efficient fluorine carrier (F_2) and chlorine carrier (HCl) are incompatible, probably because of the fast conversion of fluorine molecules into HF and the ensuing reduction in the efficiency and service life of the RF(B–X) lasing.

In Refs. 11 and 12 lasing in RF(B–X) was reported in a transverse discharge using mixtures with the weakly corrosive halogen carrier SF_6. The service life of an He/R/SF_6 laser was significantly better than that for equivalent media using F_2 molecules. The substitution and HF formation reactions in SF_6 and HCl mixtures are also less efficient that those in equivalent media using F_2 molecules. Thus, the halogen carriers SF_6 and HCl were selected to form both RF and HCl molecules.

Here we report results of an investigation of the spectral and temporal characteristics of the active medium of an excimer lamp operating simultaneously at four wavelengths: 351 nm XeF, 308 nm XeCl, 249 nm KrF, and 222 nm KrCl.

EXPERIMENTAL CONDITIONS AND METHOD

The excimer lamp had an active medium of dimensions $18 \times 2.2 \times (0.5–1.0)$ cm which was obtained using a transverse discharge with ultraviolet spark preionization. The interelectrode gap in the discharge was 2.2 cm. When radiation was coupled out from the end of the active medium, the excimer lamp uniformly illuminated an area 4 cm in diameter at a distance of 10 cm. The ultraviolet preionization system also allowed radiation to be coupled out from a 20×2 cm aperture which is of interest for pumping gas lasers. The lateral ultraviolet preionization from spark lines near the cathode had a density of one spark gap per 1.8 cm length of discharge zone.[13]

A high-voltage pulse generator was fabricated using a two-circuit LC capacitance charging circuit.[14] The 30 nF storage capacitance included three K15–10 capacitors. The 9.4 nF peaking capacitance consisted of KVI-3 capacitors. A TGI1-1000/25 hydrogen thyratron was used as the switch in the power supply system.

Pulse-periodic operation at pump pulse repetition frequencies of 5–50 Hz was achieved by using a module for transverse electrical pumping of the active mixture. The pumping module used a negative corona discharge in an electrode system comprising one row of needles and a grid

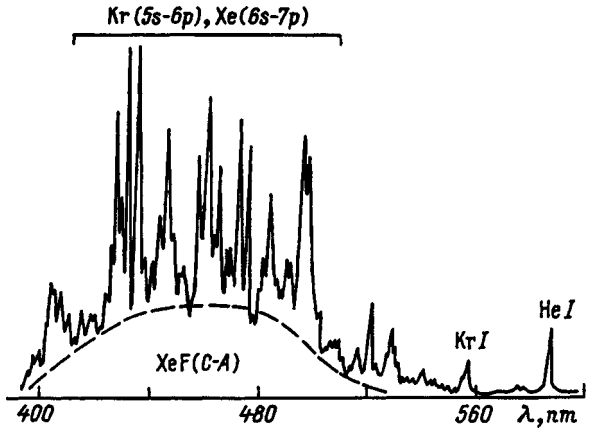


FIG. 1. General spectrum of plasma radiation for a transverse discharge in a $He/Kr/Xe/SF_6/HCl=92/8/0.8/0.12/0.08$ kPa mixture.

with an interelectrode gap of 2 cm. The pumping module was described in more detail in Ref. 15.

The current and voltage pulses were measured using a Rogowski loop and a low-inductance capacitive voltage divider. The plasma radiation spectrum of the transverse discharge was investigated using an MDR-2 monochromator, a Foton photomultiplier, and electrical system for recording low-repetition-frequency radiation pulses, and a KSP-4 automatic plotter. The temporal characteristics of the plasma radiation were measured using an ÉLU 14-FS electron linear multiplier. Pulsed measurements were made simultaneously using different channels of a 6LOR-04 oscilloscope. Estimates of the total ultraviolet radiation power of the excimer lamp were made using an IMO-2N meter 8 cm from the front end of the active medium. With allowance for the dependence of the excimer-lamp radiation intensity on distance,[5] the radiation power near the active medium (at a distance of 2 cm) at $U_C \geqslant 15$ kV may reach 20–50 kW. These output characteristics of a multiwave excimer lamp are similar to those obtained in Refs. 4–6 where the active media of single-wave excimer lamps were also pumped by a pulsed nanosecond discharge.

CHARACTERISTICS OF PLASMA RADIATION

An investigation was made of the spectral and temporal characteristics of the transverse discharge radiation in a $He/Kr/Xe/SF_6/HCl$ mixture at a total pressure of 25–100 kPa. Data from Refs. 16 and 17 was used to assign the radiation to the excimer molecular transitions. Figure 1 shows a typical general spectrum of the discharge radiation in the visible and Fig. 2 shows that in the ultraviolet. The difference between the radiation intensity in the visible and the ultraviolet reached two orders of magnitude. In the 400–500 nm range a line spectrum was observed against the background of a broad XeF(C–A) band. This was characterized by the highest-intensity lines of krypton and xenon atoms mainly on the Kr($5s$–$6p$) and Xe($6s$–$7p$) transitions. The highest-intensity KrI and XeI lines were similar to those observed in a transverse discharge using a He/Kr/Xe/HCl mixture,[15] where a detailed identification of the plasma radiation was given in the visible. Of the buffer gas lines only the

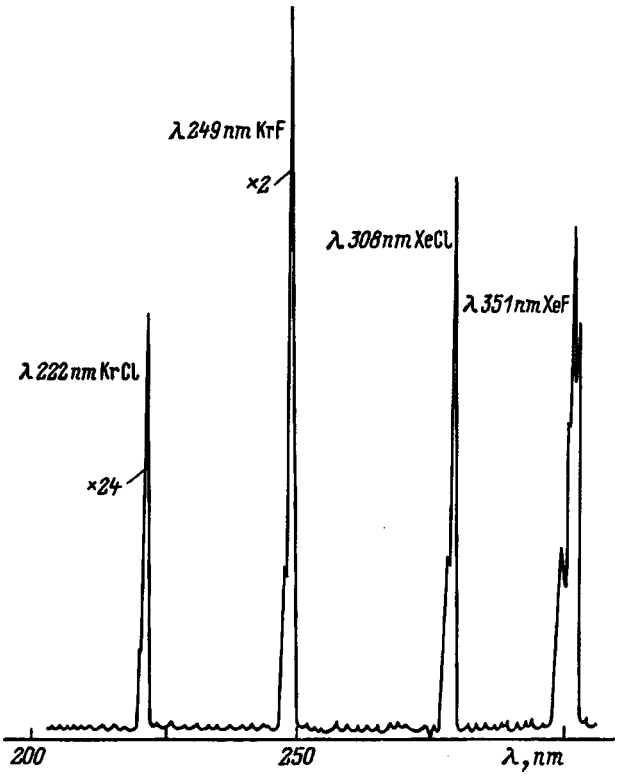


FIG. 2. Plasma radiation spectrum for a transverse discharge in a $He/Kr/Xe/SF_6/HCl=92/8/0.8/0.12/0.16$ kPa mixture in the ultraviolet for $U_{ch}=15$ kV.

$\lambda = 587.6$ nm HeI line was not observed. The predominant population of the 3^3D He atomic state, which is the upper state for this transition, is attributed in Ref. 18 to mixing of HeI states with discharge electrons since the rate constant of this reaction is negative and this level plays the role of acceptor.

The spectrum in Fig. 2 takes no account of the spectral sensitivity of the recording system. Under the conditions used to record this spectrum measurements were made of the relative brightness of the radiation bands for B–X transitions of KrCl, KrF, XeCl, and XeF with allowance made for the spectral sensitivity of the monochromator plus photomultiplier system. The brightness of the bands was defined as the area below the corresponding spectral curve on the chart paper. The brightness of the 222 nm KrCl, 249 nm KrF, 308 nm XeCl, and 351 nm XeF bands under these conditions was 3.8, 2.1, 6.1, and 4.1 a.u., respectively. An increase in the content of SF_6 molecules above 12 kPa resulted in the appearance of inhomogeneities in the transverse discharge. In the initial period of operation using freshly prepared mixture, the presence of highly electronegative molecules resulted in undercharging of the storage capacitor in the discharge supply system, as was observed with fluorine molecules.[19] After a certain time had elapsed, corresponding to the production of $(5-7)\times 10^3$ pulses, the SF_6 molecules in the discharge plasma dissociated, and the undercharging of the storage capacitor disappeared. The service life of the mixture was determined from this time at a pulse repetition frequency of 3–5 Hz. A study of the radiation brightness of the RX(B–X) bands as a function of the number of discharge current pulses

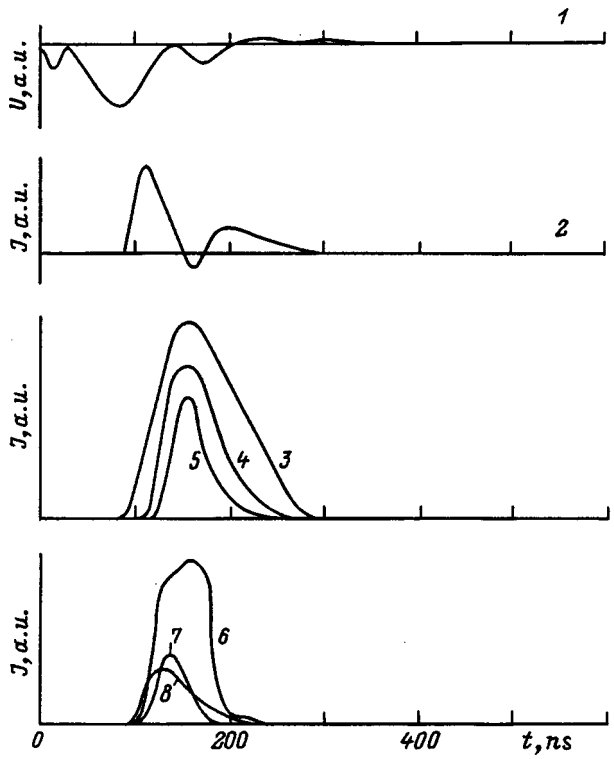


FIG. 3. Oscilloscope traces of voltage (*1*), current (*2*), and radiation (*3–8*) for a transverse discharge in a He/Kr/Xe/SF_6/HCl=92/8/0.8/0.12/0.04 kPa mixture: *3* — 351 nm XeF ($B-X$), *4* — 308 nm XeCl ($B-X$), *5* — 264 nm XeF ($D-X$), *6* — 249 nm KrF ($B-X$), *7* — 222 nm KrCl ($B-X$), and *8* — 587 nm HeI.

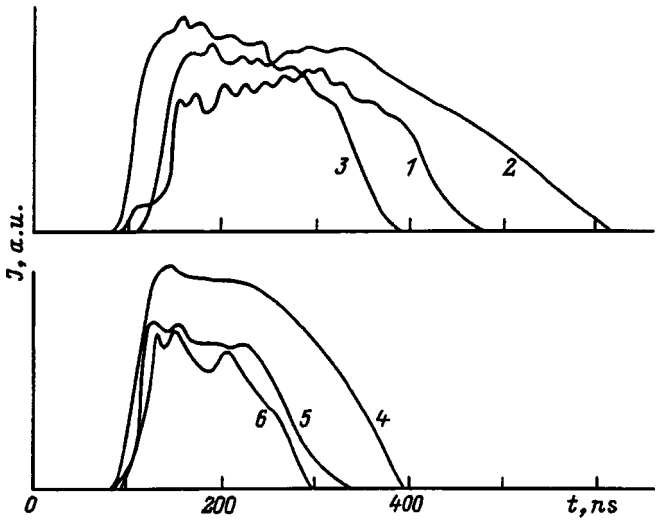


FIG. 4. Oscilloscope traces of plasma radiation for a transverse discharge in a He/Kr/Xe/SF_6/HCl=25/8/0.8/0.04/0.04 kPa mixture: *1* — XeF ($B-X$), *2* — XeCl ($B-X$), *3* — XeF ($D-X$), *4* — KrF ($B-X$), *5* — KrF ($D-X$), and *6* — KrCl ($B-X$).

in a freshly prepared mixture revealed that this decreases by a factor of 3–4 after $(1–0.51)\times10^5$ pulses.

When the density of HCl molecules increased above 0.1–0.2 kPa, radiation on the KrCl and XeCl transitions began to predominate. For approximately the same brightness of the Xe and Kr excimer radiation bands, the ratio of the krypton and xenon densities was 8/0.8 kPa, which approximately corresponded to the same lasing power as that observed for the system of 222 nm KrCl/308 nm XeCl bands in a He/Kr/Xe/HCl transverse discharge.[8,9]

Figures 3 and 4 show oscilloscope traces of the voltage, current, and radiation for RX molecules in a transverse discharge using He/Kr/Xe/SF_6/HCl mixtures. At charging voltages of 10–20 kV the pulse current was 5–15 kA. The discharge current pulse consisted of two peaks, the first having a base duration not exceeding 50 ns. The excimer molecular radiation reached a maximum in the near afterglow of the transverse discharge. The shortest radiation pulses on the RX ($B–X$) transition were observed in an atmospheric-pressure active mixture and had a base duration of 100–200 ns. In this case, the duration of the XeX molecular radiation was 1.5–2 times that of the KrX radiation. The radiation on the XeF ($D–X$) transition was delayed by approximately 20–25 ns compared with the XeF($B–X$) band and its duration was shorter. When the mixture pressure was reduced and the content of halogen-containing molecules (SF_6, HCl) in the discharge was lower, the duration of the RX radiation pulses showed a substantial (a factor of 2–3) increase (Fig. 4). This increase in the duration of the radiation may be attributed to a reduction in the efficiency of RX quenching by halogen-containing molecules or their decay products, and also to the increased lifetime of the homogeneous discharge stage.[20,21]

CONCLUSIONS

An investigation of the optical characteristics of a multiwave excimer lamp using the system of 222 nm KrCl/249 nm KrF/308 nm XeCl/351 nm XeF bands obtained by exciting a He/Kr/Xe/SF_6/HCl mixture using a nanosecond transverse discharge with ultraviolet spark preionization has shown that the relative brightnesses of all the excimer bands are equal (to within $\pm25\%$) when the content of SF_6 molecules is 0.04–0.12 kPa, the HCl content is 0.1–0.2 kPa, and the concentration ratio is [Kr]/[Xe]=10. The service life of this medium was 10^5 pulses. The duration of the radiation pulses for the RX($B–X$) transitions could be regulated between 100 and 400 ns by varying the content of SF_6 and HCl molecules in the mixture. As well as having applications in photochemistry, medicine, and biology, this source can also be used to calibrate spectral apparatus in the 222–351 nm wavelength range.

[1] Yu. G. Basov, *Pump Sources for Microsecond Lasers* [in Russian], Énergoatomizdat, Moscow (1990), 240 pp.
[2] T. Gerber, W. Lüthy, and P. Burkhard, Opt. Commun. **35**, 242 (1980).
[3] H. Kumagai and M. Obara, IEEE Trans. Plasma Sci. **16**, 453 (1988).
[4] B. A. Koval', V. S. Skakun, V. F. Tarasenko, E. A. Fomin *et al.*, Prib. Tekh. Eksp. No. 4, 224 (1992).
[5] A. M. Boĭchenko, V. S. Skakun, V. F. Tarasenko, E. A. Fomin *et al.*, Kvantovaya Élektron. (Moscow) **20**, 613 (1993) [Quantum Electron. **23**, 532 (1993)].
[6] A. N. Panchenko, É. A. Sosnin, and V. F. Tarasenko, Zh. Tekh. Fiz. **67**(1), 78 (1997) [Tech. Phys. **42**, 68 (1997)].
[7] A. M. Boĭchenko, V. S. Skakun, É. A. Sosnin *et al.*, Kvant. Elektron. (Moscow) **23**, 344 (1996).
[8] M. Shusen, Y. Yongbang, and Sh. Xinxin, Rev. Roum. Phys. **31**, 881 (1986).
[9] I. V. Chaltakov, N. I. Minkovski, T. P. Mirtshev, and I. V. Tomov, Opt. Quantum Electron. **20**, 215 (1988).

[10] V. V. Atezhev, V. S. Bukreev, S. K. Vartapetov *et al.*, Kratk. Soobshch. Fiz. No. 9, 19 (1987).
[11] V. N. Ishchenko, V. N. Lisitsyn, and A. M. Razhev, Pis'ma Zh. Tekh. Fiz. **2**, 839 (1976) [Sov. Tech. Phys. Lett. **2**, 327 (1976)].
[12] A. M. Razhev, in *Proceedings of the Twelfth Scientific Conference of Young Researchers*, Novosibirsk, 1978 [in Russian], pp. 123–128.
[13] A. K. Shuaibov, Yu. Yu. Neĭmet, A. I. Khodanich, and V. S. Shevera, Opt. Spektrosk. **75**, 713 (1993) [Opt. Spectrosc. **75**, 422 (1993)].
[14] A. K. Shuaibov and A. I. Minya, Zh. Prikl. Spektrosk. **64**, 523 (1997).
[15] A. K. Shuaibov, A. I. Minya, V. V. Zvenigorodskiĭ, and V. S. Shevera, Pis'ma Zh. Tekh. Fiz. **22**(13), 73 (1996) [Tech. Phys. Lett. **22**, 551 (1996)].
[16] *Excimer Lasers*, edited by C. K. Rhodes (Springer-Verlag, New York, 1979; Mir, Moscow, 1981, 245 pp.).
[17] *Applied Atomic Collision Physics*, Vol. 3, edited by E. W. McDaniel and W. Nigan (Academic Press, New York, 1982, Mir, Moscow, 1986, 548 pp.].
[18] Yu. A. Piotrovskiĭ and Yu. P. Tolmachev, Opt. Spektrosk. **52**, 33 (1982) [Opt. Spectrosc. **52**, 18 (1982)].
[19] V. Yu. Baranov, V. M. Borisov, and Yu. Yu. Stepanov, *Electric-Discharge Excimer Lasers using Rare-Gas Halides* [in Russian], Énergoatomizdat, Moscow (1988), 216 pp.
[20] S. V. Mel'nichenko, A. N. Panchenko, and V. F. Tarasenko, Kvantovaya Élektron. **11**, 1490 (1984) [Sov. J. Quantum Electron. **14**, 1009 (1984)].
[21] G. A. Mesyats, V. V. Osipov, and V. F. Tarasenko, *Pulsed Gas Lasers* (SPIE Press, Bellingham, WA, 1995; Nauka, Moscow, 1991, 272 pp.).

Translated by R. M. Durham

Measurement of the attenuation constant and the phase propagation constant of magnetostatic surface waves in a microstrip transducer in the 9–9.8 GHz range

M. A. Grigor'ev, A. P. Geĭ, A. I. Mishchenko, Yu. N. Navrotskaya, and A. V. Tolstikov

Saratov State University, 410071 Saratov, Russia
(Submitted October 15, 1996; resubmitted July 3, 1997)
Zh. Tekh. Fiz. **68**, 68–74 (December 1998)

An experimental investigation was made of an open-ended microstrip transducer which excites a magnetostatic surface wave in the frequency range 9–9.8 GHz in various magnetic fields. A description is given of a method of measuring the input impedance of the transducer in the microwave when an array of connecting waveguide elements with losses were present between the transducer and the resistance measuring system. Dependences of the radiation coefficient and the phase propagation constant of an electromagnetic wave in the transducer were obtained as a function of the magnetic field induction. Calculations were made of the efficiency of conversion of electromagnetic energy into a electromagnetic surface wave. © *1998 American Institute of Physics.* [S1063-7842(98)01312-9]

INTRODUCTION

A microstrip transducer is usually used to excite and receive magnetostatic waves. This transducer is formed by a section of narrow metal strip either deposited on the surface of a ferrite film in which the magnetostatic waves (MSWs) are excited or on a dielectric substrate clamped to the ferrite film. Together with the grounded base, this strip forms a short two-conductor line, one end of which is connected to an electromagnetic power source. The length of the strip is usually short compared with the length of the electromagnetic wave in this line and then the other end should be shorted (short-circuit regime). If this end is left open (''idle'' regime), the alternating current flowing in the strip will be almost zero and cannot excite magnetostatic waves.

The present paper reports an investigation of a microstrip transducer which excited magnetostatic surface waves (MSSWs) at frequencies around 10 GHz in a tangentially magnetized ferrite film provided that the strip length exceeds the electromagnetic wavelength. A similar situation is found, for example, when magnetooptic interaction takes place at these frequencies, for which the efficiency increases with increasing transducer length.

In an infinitely long microstrip transducer, a traveling electromagnetic wave would excite MSWs propagating in both directions away from the strip in the ferrite film regardless of the direction of magnetization. These waves have linear wavefronts and wave vectors whose moduli are $K_{\mathrm{MSW}}=\omega/V_{\mathrm{MSW}}$, where ω is the frequency and V_{MSW} are the velocities of those types of MSWs which are excited. The projections of the wave vectors of these waves on the direction of the microstrip transducer should be equal to the wave number of the electromagnetic wave in the transducer $K_{\mathrm{EM}}=\omega/V_{\mathrm{EM}}$, where V_{EM} is the electromagnetic wave velocity. From this it follows that the wave vectors of the excited MSWs in the plane parallel to the ferrite film should form the angle $\Theta=\arcsin(V_{\mathrm{MSW}}/V_{\mathrm{EM}})$ with the normal to the strip.

If the microstrip transducer has a finite length, regardless of whether the short-circuit or idle regime prevails, a standing electromagnetic wave will be formed which is the sum of the forward and reflected traveling waves. Each of these can be considered to be a source of excitation of partial MSWs. As a result of the diffraction divergence these will have quasilinear fronts and various angular spectra. In this case, however, the directions of the principal lobes of these spectra form the angles $\pm\Theta$ with the normal to the strip. The plus and minus signs refer to the forward and return electromagnetic waves, respectively.

For a tangentially magnetized ferrite film when only MSSWs will be excited, the forward and return electromagnetic waves forming a standing wave in the transducer create two pairs of partial MSSWs. The first propagates in one direction from the plane of symmetry of the transducer and the second propagates in the other direction.

As the electromagnetic wave propagates along the transducer, it undergoes attenuation not only as a result of dissipation in the metal conductors and the dielectric but also as a result of excitation of MSSWs. We can assume that the excitation of MSSWs is equivalent to introducing some additional linear impedance into the line which we shall call the radiation impedance $Z_{\mathrm{rad}}=R_{\mathrm{rad}}+jX_{\mathrm{rad}}$. Dissipation may occur in the line as a result of the finite Ohmic resistance of the conductors and because of the conductance of the medium filling the line. Assuming that this loss conductance is negligible, the dissipation will only be determined by the linear effective resistance of the conductors R_{los}. The fraction of the supplied electromagnetic power expended in exciting all the MSSWs is then $R_{\mathrm{rad}}/(R_{\mathrm{rad}}+R_{\mathrm{los}})$ and the remainder is converted into heat.

Since the electromagnetic wave is attenuated, the amplitude of the excited MSW does not remain constant over the beam cross section and will decrease along the transducer. In this case, an MSSW originating from a forward electromagnetic wave will evidently have a larger amplitude than one

originating from a return wave. Thus, the microstrip transducer should be oriented on the surface of the ferrite film such that the higher-intensity partial MSSW would propagate in the required direction and the direction of the tangential magnetization of the film would be selected so that the excitation of other types of MSWs could be neglected. It should also be borne in mind that in the near-field zone where the beams generated by the forward and return electromagnetic waves overlap, the distribution of the oscillation amplitude reveals a complex interference pattern. Quite clearly, at a sufficient distance from the transducer, each of the partial MSSWs may be detected by a receiving antenna if this is correctly positioned on the surface of the ferrite film. In the near-field zone, however, these waves can easily be detected (separated) in practice by means of magnetooptic interaction since the incident laser beam must be directed toward the front of the MSW at a certain angle, known as the Bragg angle, in order to obtain diffracted light.

All the excited waves together have the total power

$$P_{\Sigma\ \mathrm{MSW}}=P_{\mathrm{in}}R_{\mathrm{rad}}/(R_{\mathrm{rad}}+R_{\mathrm{los}}),$$

where P_{in} is the electromagnetic power coupled into the microstrip transducer.

In the waveguide used to supply energy from the generator to the microstrip transducer, a standing wave is established, this being formed by the forward and reflected electromagnetic waves carrying the powers P_+ and P_-, respectively. The latter is determined by the reflection coefficient Γ which itself depends on the input impedance $Z_{\mathrm{in}}=R_{\mathrm{in}}+jX_{\mathrm{in}}$ of the transducer:

$$P_{\mathrm{in}}=P_+-P_-=P_+(1-|\Gamma|^2)$$
$$=P_+(1-|Z_{\mathrm{in}}-Z_0|^2/|Z_{\mathrm{in}}+Z_0|^2),$$

where Z_0 is the wave impedance of the waveguide.

If we introduce the total conversion efficiency $\eta_\Sigma=P_{\Sigma\ \mathrm{MSW}}/P_+$, we will have

$$\eta_\Sigma=(1-|Z_{\mathrm{in}}-Z_0|^2/|Z_{\mathrm{in}}+Z_0|^2)R_{\mathrm{rad}}/(R_{\mathrm{rad}}+R_{\mathrm{los}}). \quad (1)$$

The power of the highest-intensity partial MSSW $P_{\mathrm{MSW}+}$ is less than half the power dissipated in the excitation of all four waves. Thus, the conversion efficiency for the partial wave $\eta=P_{\mathrm{MSW}+}/P_+$ cannot exceed $\eta_\Sigma/2$. Assuming that both partial MSSWs excited by the forward electromagnetic wave have the same power, we obtain

$$\eta=\eta_\Sigma/[2(1+\exp(-2\alpha l))], \quad (2)$$

where α is the attenuation constant of the electromagnetic wave in the transducer caused by the combined effect of dissipation and conversion, and l is the length of the microstrip transducer.

Neglecting the shunting loss conductance in the microstrip transducer and assuming that dissipation is determined only by the resistance R_{los} per unit length, we can show[1] that if $(R_{\mathrm{rad}}+R_{\mathrm{los}})^2/4\beta^2Z_{0\ \mathrm{MS}}^2\ll 1$, where β is the phase propagation constant of the electromagnetic wave in the transducer, $Z_{0\ \mathrm{MS}}$ is the wave impedance of the transducer calculated neglecting attenuation, the coefficient α is related to the resistance per unit length $(R_{\mathrm{rad}}+R_{\mathrm{los}})$ by

$$\alpha\approx(R_{\mathrm{rac}}+R_{\mathrm{los}})/2Z_{0\ \mathrm{MS}}, \quad (3)$$

$$\alpha=\alpha_{\mathrm{rad}}+\alpha_{\mathrm{los}}, \quad (4)$$

where $\alpha_{\mathrm{rad}}=R_{\mathrm{rad}}/2Z_{0\ \mathrm{MS}}$ will be called the radiation coefficient and $\alpha_{\mathrm{los}}=R_{\mathrm{los}}/2Z_{0\ \mathrm{MS}}$ is the loss factor.

It should be noted that for transducers whose length is comparable with or exceeds the electromagnetic wavelength, it is more logical to characterize the efficiency not by the radiation resistance R_{rad}, but by the radiation coefficient α_{rad}. Then, formula (1) has the form

$$\eta_\Sigma=(1-|Z_{\mathrm{in}}-Z_0|^2/|Z_{\mathrm{in}}+Z_0|^2)\alpha_{\mathrm{rad}}/\alpha. \quad (5)$$

It is known[1] that the input impedance Z_{in} of a lossy line in the short-circuit or idle state is given by

$$Z_{in}=\dot{Z}_{0\ \mathrm{MS}}[\tanh(pl)]^i, \quad (6)$$

where $p=\alpha+j\beta$, $i=1$ for short circuit, $i=-1$ for idle, $\dot{Z}_{0\ \mathrm{MS}}$ is the complex wave impedance of the microstrip transducer:

$$\dot{Z}_{0\ \mathrm{MS}}=|\dot{Z}_{0\ \mathrm{MS}}|\exp(i\varphi), \quad (7)$$

where for $G_{\mathrm{los}}=0$ we have

$$|Z_{0\ \mathrm{MS}}|=Z_{0\ \mathrm{MS}}\sqrt[4]{1+[2\alpha\beta/(\beta^2-\alpha^2)]^2},$$

$$\varphi=\{\arctan[(\beta^2-\alpha^2)/2\alpha\beta]-\pi/2\}/2.$$

Formulas (5)–(7) can be used to calculate the conversion efficiency and thus to find the fraction of the excited MSW power given the power of the forward electromagnetic wave in the supplying waveguide. For this purpose we need to know the complex propagation constant p and also the radiation coefficient α_{rad} and the loss factor α_{los}.

The aim of the present study is to find the values of α, β, and α_{rad} using the results of measurements of the input impedance of a microstrip transducer for the case where a magnetostatic surface wave is excited. The experiments were carried out at frequencies of 9.0, 9.4, and 9.8 GHz in different magnetic fields.

We know of four publications[2–5] which report results of measurements of the radiation impedance at frequencies of 1–5 GHz for short-circuited microstrip transducers. Two of these studied short transducers whose length was substantially shorter than the electromagnetic wavelength. Sorokin *et al.*[4] studied a microstrip transducer of varying length between 4 and 16 mm. However, the dimensions of the transducer expressed in electromagnetic wavelengths varied in the range between 0.075 and 0.6. Dmitriev *et al.*[5] investigated microstrip transducers 1, 2, and 4 mm long in the range 4–5 GHz, which corresponded to 0.05–0.2 wavelengths. In Refs. 2–4 measurements were made of the complex transmission coefficients using industrial meters and these authors neglected the influence of the array of elements in the transmitting channel which are unavoidably present in an experimental system. In order to eliminate the influence of transition channels, Dmitriev *et al.*[5] fabricated a special microstrip measuring line with a moving loop probe which directly supplies the transducer. This was used to measure the input im-

pedance of a closed-end microstrip transducer and these results were used to calculate the required radiation impedance.

It should be noted that the fabrication of a special measuring line to measure the input impedance of a microstrip transducer is a difficult task. A mechanism must be fabricated to move the probe which can operate in the immediate vicinity of the gap in the electromagnet where the magnetic field is already fairly strong. In addition, extremely accurate measurements of the probe displacements are required. We shall give an example to illustrate this last statement. At 5 GHz the wavelength in a Polykor microstrip line is 2 cm. If the transducer length is 4 mm, the phase of the reflection coefficient at the entrance to the transducer is 0.2π. In this case, a $\sim$0.1 mm error in measuring the position of the probe gives a relative phase measurement error of $\sim$20%.

Unlike these previous studies, we carried out experiments with microstrip transducers operating in the idle regime at higher frequencies (9–9.8 GHz) where the transducer length investigated was 13.5 mm which was 1.1–1.2 Λ (Λ is the wavelength in the waveguide). The measuring device was a standard waveguide measuring line. In view of this last circumstance, we paid considerable attention to the measuring technique to allow for a transitional impedance transformer with losses. It should also be noted that we measured a more natural quantity for traveling-wave transducers, i.e., the complex propagation constant $\alpha+j\beta$, rather than the radiation impedance per unit length. Finally, unlike other authors, we used the experimental result to estimate the efficiencies of conversion of electromagnetic energy into the energy of an excited surface MSW.

We also note that the experimental results are not compared with the theory. This is because in existing theoretical studies the radiation impedance was sought for short microstrip transducers[2,6,7] where the electromagnetic field along the transducer remains constant. In Ref. 8 equations were only derived to find the complex propagation constant $\alpha+j\beta$ and the characteristic impedance and not to make a numerical analysis. A considerable amount of computation work would be required to find the theoretical values of these quantities, which is outside the scope of our problem.

CONVERTER STUDIED

An MSSW was excited in an iron yttrium garnet (YIG) film grown in a gallium gadolinium parent substrate. This substrate was parallel to the (111) plane and had a thickness $h=0.5$ mm while the YIG layers deposited on its surface were each 1.56 μm thick. The microstrip transducer comprised a copper strip deposited in vacuum directly onto the YIG film, of length $l=13.5$ mm, width $w=50\,\mu$m, and thickness 2–3 μm. The entire structure was placed on a grounded brass base to which was clamped a coaxial strip junction with a wave impedance of 50 Ω. The strip together with the base formed a section of high-resistivity microstrip line open at one end and connected to a supply channel at the other. The final section of this channel, directly connected to the transducer, served as a microstrip line with a wave impedance of 50 Ω. This was formed on the same YIG–GGG–YIG structure by a copper strip 0.5 mm wide, 5 mm long, perpendicular to the transducer being studied. The entire structure was inserted in the gap of an electromagnet which generated a static magnetic field parallel to the transducer. With this orientation an MSSW was excited in the YIG layer.[9]

MEASUREMENT METHOD

Electromagnetic power was supplied to this microstrip transducer by the following elements of the transmission channel: a rectangular waveguide (23×10 mm), waveguide-coaxial and coaxial-strip junctions, a section of rigid coaxial line ($Z_0=50\,\Omega$), and a section of microstrip line ($Z_0=50\,\Omega$, $l=5$ mm). These elements together formed a quadrupole whose input ''terminals'' (*aa*) were positioned in a certain cross section of the rectangular waveguide and whose output terminals (*bb*) were positioned at the connecting point between the microstrip transducer and the microstrip line. The essence of this method, usually known as the quadrupole method, involved measuring the impedance at the terminals (*aa*) and then calculating the required impedance at terminals (*bb*). In order to make these calculations, it was first necessary to determine the parameters of the quadrupole positioned between these terminals.

It is known[10] that any passive quadrupole transforms the resistance as given by

$$Z_{\text{in}}=(AZ_{\text{out}}+B)/(CZ_{\text{out}}+D), \tag{8}$$

where *A*, *B*, *C*, and *D* are complex elements of the transfer matrix (*A*-parameters).

These parameters can be determined by carrying out additional experiments in which Z_{in} is measured with known resistances (standards) Z_{et} connected to the output. Three such experiments are sufficient to find the reduced *A*-parameters $A/D=a$, $B/D=b$, and $C/D=c$, which completely determine the law of transformation of the resistance.

In our case, technical difficulties prevent standards from being connected directly to the quadrupole output. Thus, we formed an auxiliary (new) quadrupole whose output ended in a standard coaxial connector. This ensured that a section of coaxial line with a moving plunger (reactance standard) could be connected to the output as standards. The impedances of these standards for various positions of the plunger were first measured using a coaxial measuring line. The new quadrupole was formed by replacing the microstrip line with a coaxial-strip junction by a section of coaxial line with a standard connector at the exit. Thus, the parameters *a*, *b*, and *c* of the new quadrupole were determined experimentally and this change was taken into account in the calculations by then calculating the required impedance of the transducer. Two additional quadrupoles were then connected, as it were, in series to the new quadrupole, to allow for a return to the initial state. The first allowed for removal of the coaxial line section and the second allowed for return of the microstrip line. The resistance transformation law was used in the form

$$Z_{\text{in}}=Z_0\frac{Z_{\text{out}}\cos\beta l+jZ_0\sin\beta l}{Z_0\cos\beta l+jZ_{\text{out}}\sin\beta l}, \tag{9}$$

where β is the phase propagation constant and l is the line length.

The first section was considered to be a line of negative length with a wave impedance of 50 Ω and the second was a line of positive length ($l=5$ mm) with the calculated[11] wave impedance and wavelength $\Lambda=\lambda/\sqrt{\varepsilon_{ef}}$, where ε_{ef} is the effective permittivity. It was assumed that the dielectric in the line is gadolinium gallium garnet with a relative permittivity $\varepsilon_r=14$. The length of the first auxiliary section was determined experimentally as the length of a standard line which gives the same phase shift as a real one. For this purpose the measuring line was shorted by a short-circuiting device and the standing-wave minimum was found. The junction under study, shorted at the output, was then connected to the measuring line and a shift of the minimum was observed, which was taken to be the equivalent length. This junction was shorted by means of a thin copper disk inserted in the gap in a plane matched with the beginning of the microstrip line.

DETERMINATION OF *A*-PARAMETERS

First we measured the impedances of the standards for which we used a system comprising a signal generator operated in a pulsed modulation regime with a repetition frequency of 1 kHz and inverse duty cycle 2, a ferrite isolator, a polarization attenuator, and a coaxial measuring line. A selective microvoltmeter was connected to the detector of this measuring line. The standards being studied were connected to the output of the measuring line. The measuring line was first loaded with the short-circuiting device and the standing-wave minimum was determined. It is known that the position of this minimum is the plane of the equivalent short-circuiting device. A reactance standard was then connected to the measuring line in place of the short-circuiting device and for all three positions of its plunger the impedances were measured in this equivalent plane. The standing-wave coefficients were measured using a calibrated attenuator method[12] and the coordinate d of the equivalent plane was determined as the distance from the nearest minimum (on the generator side). The required impedance was calculated using the measured values of the standing-wave coefficient and d (Ref. 12). Thus, three standard impedances $Z_{et,i}=R_{et,i}+jX_{et,i}$, $i=1,2,3$ were obtained for each of the three frequencies (9.0, 9.4, and 9.8 GHz). Then, the coaxial measuring line was replaced by a waveguide line. This was loaded with a ''new'' quadrupole to whose output we connected a reactance standard. The impedance was measured in the plane equivalent to the flange of the measuring system. It was determined as the position of the minimum with the short-circuiting device at this flange. For a given frequency this position was taken to be the terminals *aa*. These values of the standing-wave coefficient and the coordinate d were then used to calculate the resistance and the reactance at the quadrupole input, i.e., $R_{in,i}$ $X_{in,i}$. The propagation constant was calculated using the measured wavelength in the waveguide $\beta=2\pi/\Lambda$, and the wave impedance was calculated using the formula $Z_0=240\pi(b'/a')(\Lambda/\lambda)$ (Ref. 12) where a' and b' are the dimensions of the broad and narrow waveguide walls, respectively, and λ is the wavelength in free space. As a result, for each frequency we obtained three values of the impedance at the input to this quadrupole $Z_{in,i}=R_{in,i}+jX_{in,i}$, $i=1,2,3$. The required Λ-parameters $a=a_1+ja_2$, $b=b_1+jb_2$, and $c=c_1+jc_2$ were obtained as the solution of a system of three pairs of linear algebraic equations

$$\begin{aligned}
&a_1R_{et,i}-a_2X_{et,i}+b_1-c_1(R_{et,i}R_{in,i}-X_{in,i}X_{et,i})\\
&\quad+c_2(X_{et,i}R_{in,i}+X_{in,i}R_{et,i})=R_{in,i},\\
&a_1X_{et,i}+a_2R_{et,i}+b_2-c_1(X_{et,i}R_{in,i}+X_{in,i}R_{et,i})\\
&\quad-c_2(R_{et,i}R_{in,i}-X_{et,i}X_{in,i})=X_{in,i},
\end{aligned}\tag{10}$$

($i=1,2,3$), obtained by substituting $Z_{in,i}$ and $Z_{et,i}$ into Eq. (8).

For each frequency we obtained three complex values of the Λ-parameters by solving the system (10).

MEASUREMENTS OF THE INPUT IMPEDANCE OF THE MICROSTRIP TRANSDUCER

After restoring the quadrupole to the initial state, we measured the impedances at terminals *aa* using the waveguide measuring line. The experiments were carried out for four values of the magnetic field induction B, 0.136, 0.198, 0.233, and 0.308 T. These results and formulas (8) and (9) were used to calculate the impedances at terminals *bb*, i.e., at the input to this microstrip transducer. Figure 1 gives experimental dependences of the input impedance R_{in} (a) and the input reactance X_{in} (b) of the microstrip transducer as a function of the induction B for constant frequencies. The values of R_{in} and X_{in} vary appreciably with B. The impedance R_{in} passes through a maximum, reaching the highest value in different magnetic fields depending on the frequency. At $f=9$ GHz the maximum $R_{in}\approx 130\ \Omega$ is observed for $\beta\approx 0.18$ T. As the frequency increases, this maximum shifts toward higher B and at the same time, its value decreases. At frequencies of 9.4 and 9.8 GHz this maximum is 57 and 40 Ω, respectively. The value of R_{in} clearly depends on the radiation resistance and a decrease implies a reduction in the excited MSSW power. The lowest value $R_{in}\approx 16\ \Omega$ was obtained at $B=0.308$ T and a frequency of 9.8 GHz. We can assume that in this case, the microstrip transducer barely excites any MSSWs and the measured value of R_{in} is attributed to dissipation, mainly caused by the Ohmic resistance of its conductors. Variation of the reactive component of the input impedance X_{in} may be caused by changes in the linear radiation resistance and changes in the phase propagation constant β of the electromagnetic wave in the microstrip transducer. The largest fluctuations in X_{in} as a function of B are observed at a frequency of 9 GHz where R_{in} also exhibits the strongest fluctuations.

It should be noted that these experimental results by themselves provide valuable information on the input impedance of a long open microstrip transducer and its dependence on the magnetic field and frequency. This is required to match the transducer with the transmitting channel.

We shall use these results for a transducer with traveling electromagnetic waves to determine how the attenuation con-

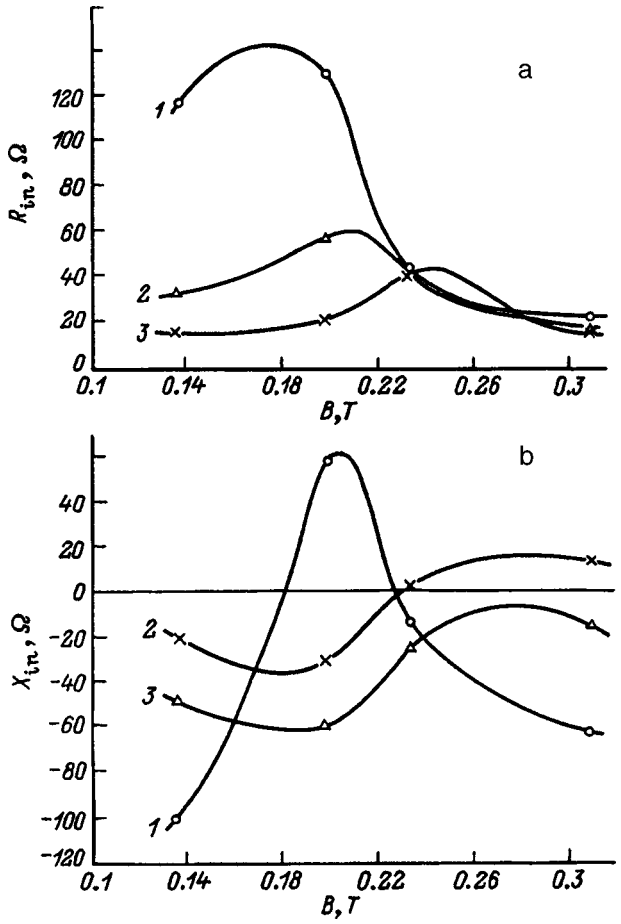


FIG. 1. Dependences of the components of the input impedance of an open-ended microstrip MSSW transducer 13.5 mm long on the magnetic field induction: Frequency, GHz: *1* — 9.0, *2* — 9.4, and *3* — 9.8.

stant $\alpha=\alpha_{rad}+\alpha_{los}$ and the phase propagation constant β of these waves behave at selected frequencies and magnetic fields.

DETERMINATION OF THE ATTENUATION CONSTANT AND PROPAGATION CONSTANT

Neglecting the shunting loss conductance, the complex wave impedance of a microstrip transducer appearing in Eq. (6) may be expressed in terms of α and β as follows:[1]

$$Z_{0\,MS}=(\beta-j\alpha)/\omega C_l, \tag{11}$$

where C_1 is the linear capacitance of the microstrip transducer, which is determined by the geometric dimensions of the transducer and the effective permittivity of the medium used to form the microstrip line.

Neglecting the influence of the ferrite films, this medium in our case is a gadolinium gallium garnet plate. Substituting formula (11) into (6), we have

$$Z_{in}=R_{in}+jX_{in}=[(\beta-j\alpha)\omega C_l]\cot(pl).$$

Introducing the notation $\cot[(\alpha+j\beta)l]=\xi+j\psi$, we can easily obtain the following system of equations in a form convenient for numerical solution:

$$\alpha=R_{in}\omega C_l/\psi-\beta\xi/\psi, \tag{12}$$

$$\beta=X_{in}\omega C_l/\psi+\alpha\xi/\psi, \tag{13}$$

where $\xi=(AC+BD)/(C^2+D^2)$, $\psi=(BC-AD)/(C^2+D^2)$, $A=\cosh(\alpha l)\cos(\beta l)$, $B=\sinh(\alpha l)\sin(\beta l)$, $C=\sinh(\alpha l)\cos(\beta l)$, $D=\cosh(\alpha l)\cos(\beta l)$.

In addition to the measured values of R_{in} and X_{in}, the transducer length $l=13.5$ mm and the linear capacitance C_l, determined from formula (11) for $\alpha=0$ as $C_l=\sqrt{\varepsilon_{ef}}/cZ_{0MS}$, where c is the velocity of light in vacuum, were set in these formulas. The values of ε_{ef} and $Z_{0,MS}$ were calculated using formulas taken from Ref. 11 for a microstrip line with zero strip thickness and $w/h<1$

$$\varepsilon_{ef}=\{[(f/f_\omega)^2+1]/[(f/f_\omega)^2+\sqrt{\varepsilon_r/\varepsilon_{ef\,0}}]\}^2\varepsilon_r,$$

$$f_\omega[\mathrm{GHz}]=3.5+(16.2\varepsilon_r^{0.25})/(1+0.12w\varepsilon_r^{0.35}/h),$$

$$\varepsilon_{ef\,0}=(1/2)[\varepsilon_r+1+(\varepsilon_r-1)(1+10h/w)^{-0.5}],$$

$$Z_{0,MS}=[120/\sqrt{2(\varepsilon_r+1)}]\{\ln(8h/w)+(w/h)^2/32-(1/2)\times(\varepsilon_r-1)[\ln(\pi/2)+\ln(4/\pi)/\varepsilon_r]/(\varepsilon_r+1)\}.$$

As a result of using these formulas, we obtained the following numerical values of $\varepsilon_{ef}, Z_{0,MS}, \Omega, C_l, 10^{-10}$, and F/m, respectively, for 9.0 GHz 8.407, 91.56, 1.056, for 9.4 GHz 8.429, 91.56, 1.058, and for 9.8 GHz 8.452, 91.56, 1.059.

In order to find α and β from the many possible solutions of the system (12) and (13), we selected that for which the propagation constant β was closest to $\beta=2\pi\sqrt{\varepsilon_{ef}}f/c$. The dependences are plotted in Fig. 2.

DISCUSSION OF RESULTS

Attention is drawn to the external similarity between the curves $\alpha(B)$ and $R_{in}(B)$, and also $\beta(B)$ and $X_{in}(B)$. The similarity between the first two is physically understandable since the presence of attenuation in an open or closed line leads to the establishment of an impedance at the input. The similarity noted between the reactive component X_{in} and the phase constant β is determined by the mathematical relationship between them. The value of β appears in the argument of the function (6) whose imaginary part is approximately described by $X_{in}\approx-\cot(\beta l)$. Over a small range of βl the latter can be approximated by a linear dependence. This leads to the similarity between the curves $X_{in}(B)$ and $\beta(B)$.

The values of α for the maxima of the curves in Fig. 2 can be interpreted as follows. As an electromagnetic wave propagates along a microstrip transducer, the power transferred by it decreases as a result of losses and the excitation of MSSWs. For each centimeter of the transducer length at frequencies of 9.0, 9.4, and 9.8 GHz this decrease is a factor of 2.3, 1.9, and 1.6, respectively.

The conversion efficiencies η_Σ and η can easily be calculated using formulas (1) and (2). Assuming that the microstrip transducer is a transmission line load with the wave impedance $Z_0=50\,\Omega$,and the coefficients α_{los} are equal to the minimum values of α on the curves in Fig. 2a, we obtain the results presented in Table I.

It is particular interesting to compare these values of α_{rad} with the results of other authors. The closest in terms of

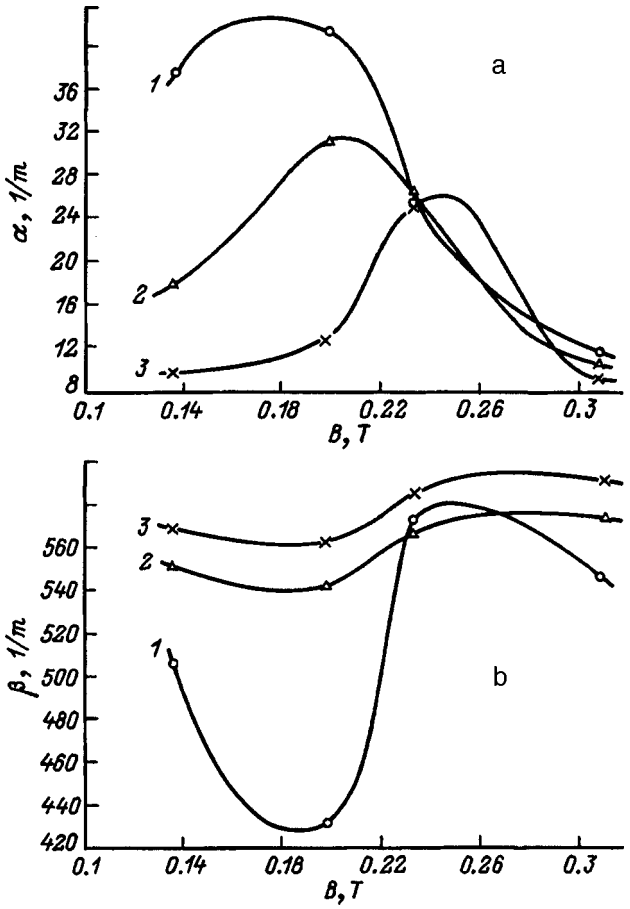


FIG. 2. Dependences of the attenuation factor (a) and the phase propagation constant (b) of an electromagnetic wave in a microstrip MSSW transducer as a function of the magnetic field induction (*1–3* as in Fig. 1).

frequency and experimental method is Ref. 5. Converting our values of α_{rad} given in Table I to the values of R_{rad} in order to make a comparison with this study yielded values of 5.3, 3.8, and 3.1 Ω/mm for frequencies of 9.0, 9.4, and 9.8 GHz, respectively. The radiation resistance from Ref. 5 for excitation of MSSWs in a tangentially magnetized ferrite film is of similar order of magnitude to our value and has a similar frequency dependence (for $f=4.1 \ldots 5$ GHz we have $R_{\text{rad}}=20 \ldots 4\ \Omega$/mm). Accurate agreement should not be expected because in our case, the frequency was almost twice as high and the YIG film was three times as thick.

It can be seen from Table I that although α varies with frequency, the efficiency η remains almost constant for the partial MSSW used. The given values of η_Σ imply that all four partial MSSWs account for 56–63% of the transferred electromagnetic wave power. Moreover, only 17.4–21.4% is converted into the partial MSSW used. Note that the intensity of the excited MSSW varies along the microstrip transducer by a factor of 1.6–2.3.

TABLE I.

f_0, GHz	R_{in}, Ω	X_{in}, Ω	α, 1/m	α_{rad}, 1/m	$-\eta_\Sigma$, dB	$-\eta$, dB	$Z_{0,\text{opt}}$, Ω	$-\eta_{\text{opt}}$, dB
9.0	131	59	42	30	2.9	7.1	143	5.9
9.4	57	60	32	21	3.0	7.6	83	7.2
9.8	41	3.5	26	17	2.0	6.7	41	6.7

We know that when a transmission line is loaded with a certain complex impedance, the power required by the load reaches a maximum if the wave impedance of the line is optimized and the line is matched with the generator. The optimum value of Z_0 is equal to the modulus of the load impedance. The values of the efficiencies η_Σ and η given in Table I were calculated for constant Z_0. The latter value differs most strongly from the optimum at a frequency of 9.0 GHz. Table I gives the optimum wave impedances $Z_{0,\text{opt}}$ and the corresponding conversion efficiencies η_{opt}. It is easy to see that using optimum wave impedances could only give an appreciable increase in η at 9.0 GHz.

We shall now discuss the behavior of the phase propagation constant β. The experimental curves plotted in Fig. 2b show that β fluctuates as the magnetic field varies. The amplitude of these fluctuations is small, being 3% of the average at frequencies of 9.4 and 9.8 GHz and reaching 15% at 9.0 GHz. Moreover, an increase in α is accompanied by a decrease in β. This implies that the excitation of MSSWs increases the phase velocity of the electromagnetic wave in the microstrip transducer. It is known[1] that when the attenuation in the line can be neglected, the phase velocity is related to the linear parameters L_1 and C_1 by the formula $\nu_f=1/\sqrt{L_1C_1}$. We can postulate that the increase in the phase velocity is caused by a decrease in the linear inductance when MSSWs are excited.

CONCLUSIONS

A method of making microwave measurements of the input impedance of a microstrip MSSW transducer has been developed experimentally which can allow for the influence of a resistance transformer formed by the elements of the transmission channel. As a result of using this method in the short-wave part of the microwave range, we obtained experimental dependences of the attenuation constant and phase propagation constant in a long open-ended microstrip MSSW transducer as a function of the magnetic field induction, and we also calculated the conversion efficiencies. These curves can be used to check the theory and also to develop efficient MSSW transducers.

[1] V. I. Kalinin and G. M. Gershteĭn, *Introduction to Radio Physics* [in Russian], GITTL, Moscow (1957), 660 pp.
[2] A. K. Ganguly and D. C. Webb, IEEE Trans. Microwave Theory Tech. **23**, 998 (1975).
[3] A. K. Ganguly, D. C. Webb, and C. Banks, IEEE Trans. Microwave Theory Tech. **26**, 444 (1978).
[4] V. G. Sorokin, P. V. Bogun, and P. E. Kandyba, Zh. Tekh. Fiz. **56**, 2377 (1986) [Sov. Phys. Tech. Phys. **31**, 1421 (1986)].
[5] V. F. Dmitriev, B. A. Kalinikos, and N. G. Kovshikov, Zh. Tekh. Fiz. **56**, 2169 (1986) [Sov. Phys. Tech. Phys. **31**, 1300 (1986)].
[6] P. R. Emtage, J. Appl. Phys. **49**, 4475 (1978).
[7] G. A. Vugal'ter and V. N. Makhalin, Zh. Tekh. Fiz. **55**, 497 (1985) [Sov. Phys. Tech. Phys. **30**, 296 (1985)].
[8] I. A. Gilinskiĭ and I. M. Shcheglov, Preprint No. 3-84 [in Russian], Institute of Semiconductor Physics, Siberian Branch of the Academy of Sciences of the USSR, Novosibirsk (1984), 21 pp.

[9] A. V. Vashkovskiĭ, V. S. Stal'makhov, and Yu. P. Sharaevskiĭ, *Magnetostatic Waves in Microwave Electronics* [in Russian], Saratov State University Press (1993), 311 pp.
[10] K. Shimoni, *Theoretical Electronics*, Mir, Moscow (1964), 760 pp.
[11] *Handbook of the Design and Construction of Microwave Strip Devices*, edited by V. I. Vol'man [in Russian], Radio i Svyaz', Moscow (1982), 328 pp.
[12] E. L. Ginzton, *Microwave Measurements* (McGraw-Hill, New York, 1957; IL, Moscow, 1960, 620 pp.).

Translated by R. M. Durham

Intrinsic and mutual inductance of apertures in parallel superconducting films

O. G. Vendik and I. S. Danilov

St. Petersburg State Electrotechnical University, 197376 St. Petersburg, Russia
(Submitted December 8, 1997)
Zh. Tekh. Fiz. **68**, 75–79 (December 1998)

Analytic expressions are derived to calculate the inductance of apertures in parallel superconducting films using the current distribution along the edge of the aperture in the film. The relations obtained show good agreement with the experimental results and can be used to model rf squids. © *1998 American Institute of Physics.* [S1063-7842(98)01412-3]

INTRODUCTION

Modeling the equivalent-circuit parameters of planar implementations of high-temperature superconducting squids requires accurate calculations of the inductance of the aperture in the superconducting film and also of the mutual inductance of the apertures in parallel superconducting films.[1] The equivalent parameters must be determined accurately to match the active element with an external circuit, for example, using a narrow slotted line.[2] At present, the inductance of an aperture in a superconducting screen is determined using approximate relations obtained by analyzing experimental results.[1] However, no theoretically substantiated formulas are available to model the equivalent parameters of an aperture in a superconducting screen which could be used at the initial design stage of a device.

Here we describe the derivation and present final relations to calculate the intrinsic and mutual inductances of apertures in parallel superconducting films.

MAGNETIC FIELD STRENGTH

An important part of the task of modeling the inductance of a squid loop is to determine the inductance of an aperture of radius a in a superconducting screen shown schematically in Fig. 1. For this particular problem the following condition is satisfied

$$\lambda_\perp < a \ll b, \tag{1}$$

where

$$\lambda_\perp = 2\lambda_L^2/t \tag{2}$$

is Pearl depth of penetration of the magnetic field in the film,[3] λ_L is the London depth of penetration of the magnetic field in the material, t is the film thickness, b is the external dimension, and a is the aperture radius.

If inequality (1) is satisfied, the dimensions of the film can be considered to be infinite. A single screen is shown in Fig. 2a. At the initial stage of deriving the calculation formulas, the problems of searching for the intrinsic and mutual inductances are combined.

The superconducting screens are located in planes parallel to the XY plane, separated by the distance d. The origin is positioned at the center of the aperture in one of the screens. It is assumed that the radii of the apertures are the same and equal to a (Fig. 2b). The current density distribution $j_{\text{surf}}(r')$ in the superconducting film at $T=77$ K along the r' coordinate is given by:[3]

$$j_{\text{surf}}(r') = j_0 \exp(-(r'-a)/\lambda_\perp), \tag{3}$$

where r' is the coordinate on which the current density is defined.

Expression (3) is valid when $t<\lambda_\perp \ll a$. The fundamental relation for the magnetic field strength may be written in a general form for $J_{\text{surf}}(r')$ (Ref. 4)

$$\mathbf{H}(\mathbf{r}) = \frac{1}{\pi}\int j_{\text{surf}}(r')\times(\mathbf{r}-\mathbf{r}')\cdot|\mathbf{r}-\mathbf{r}'|^{-3}d^3r', \tag{4}$$

where $\mathbf{r}$ is the coordinate of the point at which the magnetic field strength is determined.

The flux through the aperture in the screen is integrated over $\mathbf{r}$. When searching for the mutual inductance, $\mathbf{r}$ is the coordinate of the point inside the aperture in the second screen. Expression (4) can be conveniently written in cylindrical coordinates. For the various components of expression (3) we then have

$$|\mathbf{r}-\mathbf{r}'| = (R^2+R'^2-2RR'\cos(\varphi-\varphi')+(z-z')^2)^{-\frac{3}{2}}, \tag{5}$$

$$\mathbf{r} = \mathbf{e}_x R\sin(\varphi)+\mathbf{e}_y R\cos(\varphi)+\mathbf{e}_z\cdot(z-z'), \tag{6}$$

$$\mathbf{r}' = \mathbf{e}_x R'\sin(\varphi')+\mathbf{e}_y R'\cos(\varphi'), \tag{7}$$

$$\mathbf{j}_{\text{surf}}(\mathbf{r}') = j_{\text{surf}}(r')(\mathbf{e}_z\cos(\varphi')-\mathbf{e}_y\sin(\varphi')), \tag{8}$$

$$\mathbf{j}_{\text{surf}}(\mathbf{r}')\times\mathbf{r} = j_{\text{surf}}(r')\cdot\mathbf{e}_z(R\cos(\varphi)\cos(\varphi') + R\sin(\varphi)\sin(\varphi'))+\mathbf{e}_y(z-z')\cos(\varphi') - \mathbf{e}_x(z-z')\sin(\varphi'), \tag{8a}$$

$$\mathbf{j}_{\text{surf}}(\mathbf{r}')\times\mathbf{r}' = j_{\text{surf}}(r')\cdot\mathbf{e}_z R, \tag{8b}$$

$$d^3r' = R'\,dR'\,d\varphi'\,dz', \tag{9}$$

where $R=\sqrt{|\mathbf{r}|^2+d^2}$, $R'=\sqrt{|\mathbf{r}'|^2+d^2}$, d is the distance between the parallel films; when searching for the intrinsic inductance, d is assumed to be zero.

Since $t<\lambda_\perp$, integration over the film thickness may be replaced by multiplication by t. Then we find $z'=0$ and expression (9) has the form

$$d^3r' = t\,R'\,dR'\,d\varphi'. \tag{10}$$

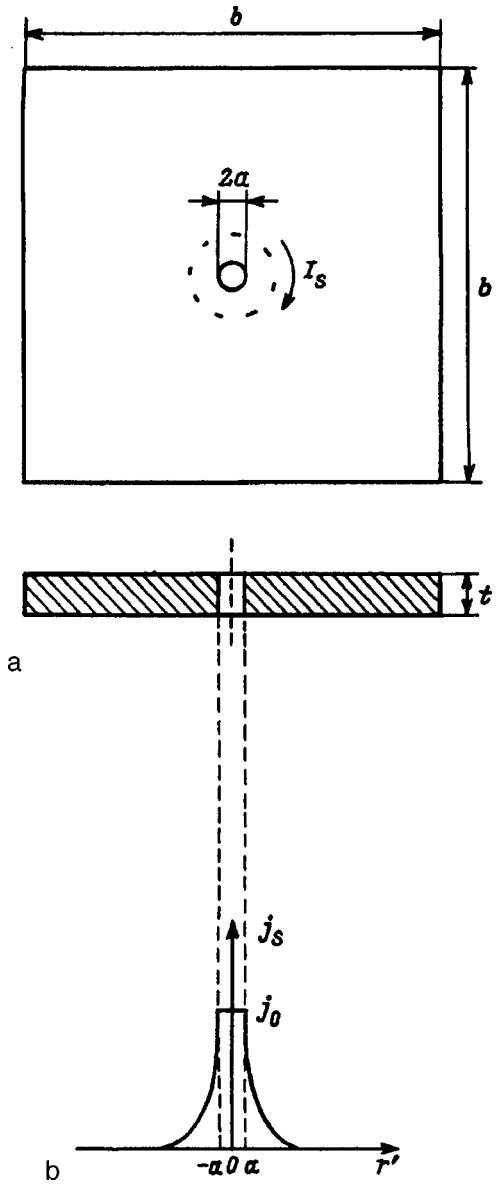


FIG. 1. Current density distribution along the edge of an aperture in a superconducting film: a — schematic of squid, b — current distribution $j(r')$ in high-temperature superconducting film along the coordinate r'.

Using expressions (5)–(10) and assuming $\varphi=0$, for the magnetic field strength given by expression (4) we have

$$\mathbf{H}(R,z)=\mathbf{e}_z\frac{j_0 t}{4\pi}\times\int_0^{2\pi}\int_a^{+\infty}\frac{(R'-R\cos(\varphi'))\exp(-(R'-a)/\lambda_\perp)}{(R^2+R'^2-2RR'\cos(\varphi')+z^2)^{3/2}}\times R'dR'd\varphi'. \quad (11)$$

MUTUAL INDUCTANCE OF APERTURES IN PARALLEL SUPERCONDUCTING FILMS

In order to find the magnetic flux through the second aperture in a parallel screen, we need to integrate expression (11) over the surface of the second aperture. If we bear in mind that the distance between the parallel films

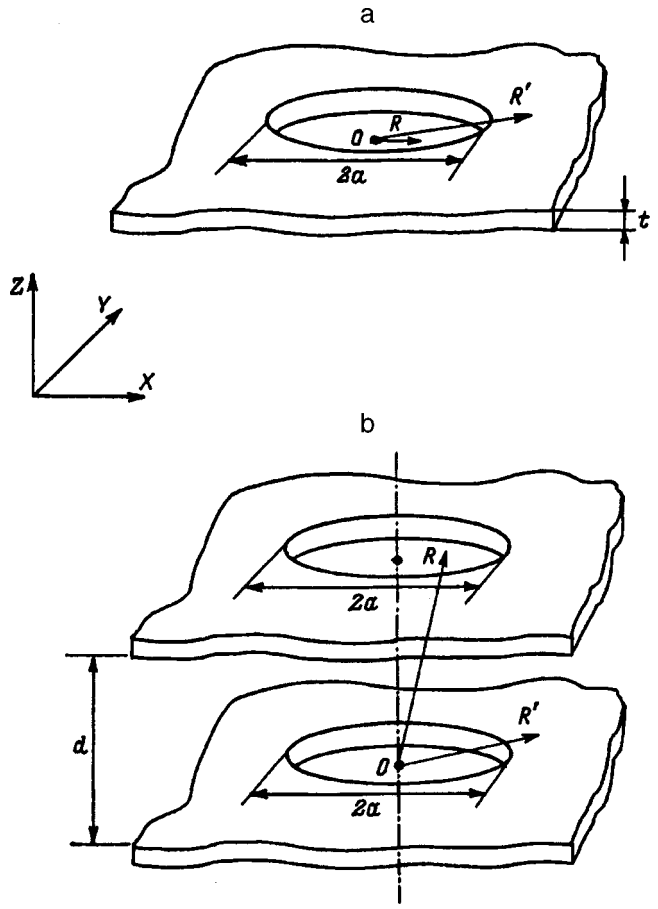


FIG. 2. Schematic of superconducting screens: a — infinite screen with aperture of radius a, b — two parallel superconducting screens with apertures of radius a; d is the distance between the screens.

$d=10\ldots200\,\mu$m is much greater than the film thickness, we can set $z=d$ in expression (11). The formula for the flux then has the form

$$\Phi=\mu_0 j_0 t\times\int_0^{\pi}\int_0^{a}\int_a^{+\infty}\frac{(R'-R\cos(\varphi'))\exp(-(R'-a)/\lambda_\perp)}{(R^2+R'^2-2RR'\cos(\varphi')+d^2)^{3/2}}\times R'RdR'dRd\varphi'. \quad (12)$$

The total current in the first screen is obtained by integrating expression (2) for the current density distribution over R'. As a result, we obtain

$$I_1=tj_0\lambda_\perp. \quad (13)$$

Assuming that the flux through the aperture in the second screen and the current in the first screen are related by $\Phi_2=MI_1$, where M is the mutual inductance of the apertures in the screens, the expression for M can be written as

$$M=\frac{\mu_0}{\lambda_\perp}\times\int_0^\pi\int_0^a\int_a^{+\infty}\frac{(R'-R\cos(\varphi'))\exp(-(R'-a)/\lambda_\perp)}{(R^2+R'^2-2RR'\cos(\varphi')+d^2)^{3/2}}\times R'RdR'dRd\varphi'. \quad (14)$$

We introduce the normalized parameters

$$x=R/a,\quad y=R'/a,\quad \delta=\lambda_\perp/a,\quad d=p/a. \quad (15)$$

In terms of the normalized values, the expression for M then has the form

$$M=\mu_0 a G_M, \quad (16)$$

where G_M is a geometric factor given by

$$G_M(p,\delta)=\frac{1}{\delta}\times\int_0^\pi\int_0^1\int_1^{+\infty}\frac{(y-x\cos(\varphi'))\exp(-(y-1)/\delta)}{(x^2+y^2-2xy\cos(\varphi')+p^2)^{3/2}}\times xy\,dy\,dx\,d\varphi'. \quad (17)$$

The method used for the subsequent transformation and calculation of the integral (17) is described in the Appendix.

INTRINSIC INDUCTANCE OF AN APERTURE IN A SUPERCONDUCTING SCREEN

The value of z is assumed to be zero when searching for the intrinsic inductance of an aperture. In order to find the magnetic flux through an aperture in a single screen, expression (11) must be integrated over the surface of the aperture. For the magnetic flux we then have

$$\Phi=\mu_0 j_0 t\times\int_0^\pi\int_0^a\int_a^{+\infty}\frac{(R'-R\cos(\varphi'))\exp(-(R'-a)/\lambda_\perp)}{(R^2+R'^2-2RR'\cos(\varphi'))^{3/2}}\times R'RdR'dRd\varphi'. \quad (18)$$

Using Eqs. (13) and (18), the normalization (15), and also assuming that $\Phi=LI_L$, where L is the intrinsic inductance of the aperture in the screen, the expression for L may be written as

$$L=\mu_0 a G_L, \quad (19)$$

where G_L is a geometric factor

$$G_L(\delta)=1/\delta\times\int_0^\pi\int_0^1\int_1^{+\infty}\frac{(y-x\cos(\varphi'))\exp(-(y-1)/\delta)}{(x^2+y^2-2xy\cos(\varphi'))^{3/2}}\times xy\,dy\,dx\,d\varphi'. \quad (20)$$

The subsequent transformation of the integral (20) is described in the Appendix. The values of the geometric factor G_L are calculated for the typical values $a=10\ldots100\,\mu$m, $\lambda_\perp=0.25\ldots1\,\mu$m ($\delta=0.003\ldots0.1$) and are given in Table I.

TABLE I.

Values of δ	Values of the geometric factor G_L
0.001	7.563
0.01	5.288
0.1	3.114

It should be noted that the dependence $G_L(\delta)$ is logarithmic and for $\delta=0.001\ldots0.1$ is well approximated by the expression

$$G_L(\delta)=2.280\log(1/\delta)+0.728. \quad (21)$$

Figure 3 gives the dependence $G_L(\delta)$ calculated using formulas (20) and (21). Essentially, formula (17) is universal for searching for both the intrinsic and the mutual inductance of the films. In the first case, p in formula (17) must be set to zero. On the basis of the results of calculations using formulas (16) and (19). Fig. 4 gives the coupling coefficient $k=M/L$ plotted as a function of p for various δ.

COMPARISON OF RESULTS

Unfortunately no reliable information is available on the experimental determination of the intrinsic inductance of a circular aperture in a superconducting screen. Thus, the comparison is made using the well-tried relations to calculate the inductance of a ring made of circular cross-section wire[5]

$$L=\mu_0 f(\ln(16f/h)-2), \quad (22)$$

where f is the radius of the ring and h is the cross-sectional diameter.

The aperture radius a in the screen is equivalent to the radius of a wire ring f. We shall find the correspondence between the cross-sectional diameter of the wire and the geometric parameters characterizing the region of current concentration in the screen $\lambda_\perp$ and t. Thus, for figures having the same perimeter[5] this relation will be

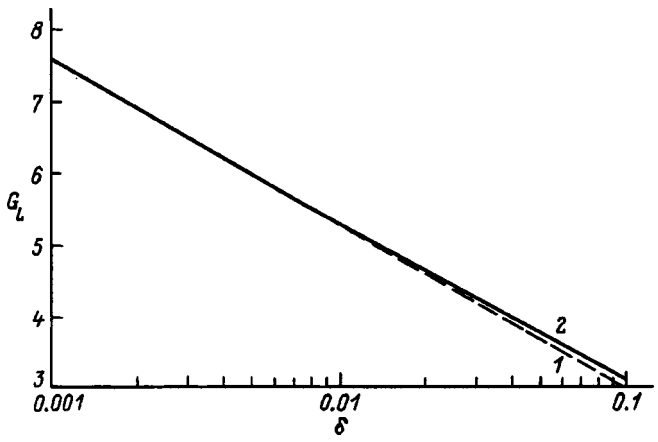


FIG. 3. Geometric factor G_L as a function of the parameter δ: *1* — calculated using Eq. (20), *2* — calculated using Eq. (21).

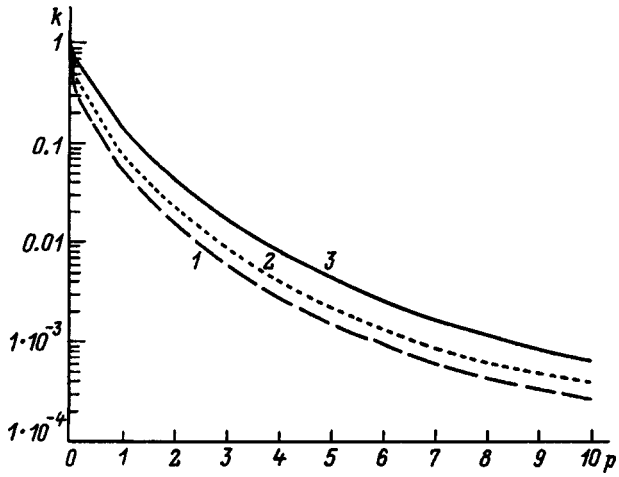


FIG. 4. Coupling coefficient $k=L/M$ as a function of the normalized distance p plotted for various values of δ: *1* — 0.001, *2* — 0.01, and *3* — 0.1.

$$h=(2/\pi)(\lambda_{\perp}+t). \tag{23}$$

For a superconducting YBCO film we find $t=0.3\,\mu$m and $\lambda_D(77\,\mathrm{K})=0.27\,\mu$m. From this it follows that $\lambda_{\perp}\approx 0.5\,\mu$m. For comparison Table II gives the inductances determined using Eqs. (19) and (22) for various values of a.

It can be seen from Table II that the calculated results show good agreement with the data from Ref. 5.

CONCLUSIONS

These analytic expressions are convenient for modeling the equivalent parameters of a circuit containing apertures in a superconducting screen. An advantage of these formulas is that they take into account the ratio of the depth of magnetic field penetration in the film to the aperture radius. The results of calculations using formula (16) for typical t and a show good agreement with the experimental results.

The authors are grateful to I. A. Nazarov for calculating the integrals.

APPENDIX

The problem involves tabulating the functions of two variables

$$G(p,\delta)=\frac{1}{\delta}\times\int_1^{+\infty}dy\int_0^1 dx\int_0^{\pi}\frac{y-x\cos(\varphi)}{(x^2+y^2+p^2-2xy\cos(\varphi))^{3/2}}xy\times\exp\left(\frac{1-y}{\delta}\right)d\varphi.$$

TABLE II.

Aperture radius a (radius of ring $f=a$), μm	Inductance, of circular cross-section wire ring Eq. (22), H	Inductance of aperture in screen Eq. (19), H
5	1.93×10^{-11}	1.96×10^{-11}
50	3.38×10^{-10}	3.32×10^{-10}
500	4.83×10^{-9}	$4.75\cdot10^{-9}$

We use the following identity transformations to calculate the internal integral

$$\frac{(y-x\cos(\varphi))y}{(x^2+y^2+p^2-2xy\cos(\varphi))^{3/2}}=0.5\frac{2y^2-2xy\cos(\varphi)+x^2+p^2-x^2-p^2}{(x^2+y^2+p^2-2xy\cos(\varphi))^{3/2}}$$

$$=0.5\frac{1}{(x^2+y^2+p^2-2xy\cos(\varphi))^{3/2}}+0.5\frac{y^2-x^2-p^2}{(x^2+y^2+p^2-2xy\cos(\varphi))^{3/2}}.$$

Using formulas 2.571.5 and 2.575.4 from Ref. 6, we obtain

$$\int_0^{\pi}\frac{d\varphi}{(x^2+y^2+p^2-2xy\cos(\varphi))^{1/2}}=\frac{2}{\sqrt{(x+y)^2+p^2}}K\left(\sqrt{\frac{4xy}{(x+y)^2+p^2}}\right),$$

$$\int_0^{\pi}\frac{d\varphi}{(x^2+y^2+p^2-2xy\cos(\varphi))^{3/2}}=\frac{2}{((x+y)^2+p^2)\sqrt{(x+y)^2+p^2}}\times E\left(\sqrt{\frac{4xy}{(x+y)^2+p^2}}\right).$$

The problem was reduced to calculating the repeated integral

$$G(p,\delta)=\frac{1}{\delta}\int_1^{+\infty}\exp\left(\frac{1-y}{\delta}\right)dy\int_0^1\frac{x}{\sqrt{(x+y)^2+p^2}}\times\left(K\left(\sqrt{\frac{4xy}{(x+y)^2+p^2}}\right)\frac{y^2-x^2-p^2}{(x-y)^2+p^2}\times E\left(\sqrt{\frac{4xy}{(x+y)^2+p^2}}\right)\right)dx.$$

For $p=0$ the internal integral can be calculated explicitly

$$\int_0^1\frac{x}{x+y}\left(K\left(\frac{2\sqrt{xy}}{x+y}\right)+\frac{y^2-x^2}{(y-x)^2}E\left(\frac{2\sqrt{xy}}{x+y}\right)\right)dx=\int_0^1\frac{x/y}{1+x/y}\left(K\left(\frac{2\sqrt{x/y}}{1+x/y}\right)+\frac{1+x/y}{1-x/y}E\left(\frac{2\sqrt{xy}}{1+x/y}\right)\right)dx$$

$$=y\int_0^{1/y}t\left(\frac{1}{1+t}K\left(\frac{2\sqrt{t}}{1+t}\right)+\frac{1}{1-t^2}(1+t)E\left(\frac{2\sqrt{t}}{1+t}\right)\right)dt.$$

Using formulas 8.126.3 and 8.126.4 from Ref. 6, we obtain

$$\left(\frac{1}{1+t}K\left(\frac{2\sqrt{t}}{1+t}\right)+\frac{1}{1-t^2}(1+t)E\left(\frac{2\sqrt{t}}{1+t}\right)\right)=\frac{2E(t)}{1-t^2},$$

and thus with allowance for formula 5.112.13 from Ref. 6

$$\int_0^{1/y} t\left(\frac{1}{1+t}K\left(\frac{2\sqrt{t}}{1+t}\right)+\frac{1}{1-t^2}(1+t)E\left(\frac{2\sqrt{t}}{1+t}\right)\right)dt$$

$$=\int_0^{1/y}\frac{tE(t)}{1-t^2}dt=2\left(K\left(\frac{1}{y}\right)-E\left(\frac{1}{y}\right)\right).$$

Finally we obtain

$$G(0,\delta)=\frac{2}{\delta}\int_1^{+\infty} y\left(K\left(\frac{1}{y}\right)-E\left(\frac{1}{y}\right)\right)\exp\left(\frac{1-y}{\delta}\right)dy.$$

This integral can also be determined numerically.

[1] Y. Zhang, N. Wolters *et al.*, in *Proceedings of the Sixth International Superconductive Electronics Conference*, Berlin, 1997, Vol. 1, p. 51.
[2] O. G. Vendik, I. S. Danilov, and S. P. Zubko, Zh. Tekh. Fiz. **67**(9), 94 (1997) [Tech. Phys. **42**, 1068 (1997)].
[3] J. Pearl, Appl. Phys. Lett. **5**(4), 65 (1964).
[4] I. E. Tamm, *Principles of the Theory of Electricity* [in Russian], Nauka, Moscow (1996), 624 pp.
[5] B. A. Smirenin, *Handbook of Radio Engineering* [in Russian], Goséner-goizdat, Moscow (1950).
[6] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, transl. of 4th Russ. ed. (Academic Press, New York, 1980; Fizmatgiz, Moscow 1962, 1100 pp.).

Translated by R. M. Durham

Device for visualizing the atomic structure of surface layers based on an electron focusing effect

I. I. Pronin, D. A. Valdaĭtsev, M. V. Gomoyunova, N. S. Faradzhev, and A. G. Banshchikov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted October 6, 1997)
Zh. Tekh. Fiz. **68**, 80–84 (December 1998)

A description is given of a simple device for visualizing the atomic structure of surface layers by recording the focusing maxima in the spatial distribution of 1–3 keV electrons reflected from a sample with energy losses up to ≈300 eV. This is based on a wide-angle retarding-field energy analyzer with a microchannel electron flux analyzer and video system for data acquisition and processing, which can obtain data at a rate of up to 50 distributions per second, allowing the dynamics of structural rearrangements to be studied in a surface region approximately 15 Å thick. The low primary electron beam current (0.1μA) minimizes the electron-stimulated effects on the object being studied. © *1998 American Institute of Physics.* [S1063-7842(98)01512-8]

The atomic structure of a surface is one of the most important characteristics which must be known to understand the specific features of its properties and the wide range of physicochemical processes taking place therein, and also to solve a broad spectrum of applied problems. Thus, over the last few decades particular attention has been focused on the development of methods of structural analysis for surface regions of the order of a few angstrom thick. As a result, a comprehensive range of highly informative methods have been developed of which the most important are high-energy (E_p = 20–50 keV) and low-energy (up to 300 eV) electron diffraction. Until recently, intermediate-energy electrons ($E_p \approx 1$ keV) were not used in structural analysis. However, evidence has recently appeared to suggest that the diffraction patterns (Kikuchi patterns) of moderate-energy reflected electrons are also highly informative and even have some advantages over the conventional methods. In this energy range a key role in the mechanism of electron scattering by a solid is played by a forward-focusing effect.[1] This has the result that the dominant characteristics of the diffraction patterns are maxima oriented along chains of atoms. Thus, simple measurements of the Kikuchi pattern over a fairly wide solid angle can directly visualize the crystal structure of a thin surface layer of an object in real space.[2–6]

Here we use two methods of recording the Kikuchi patterns. The simplest involves using a standard electron diffraction camera widely used in low-energy electron diffraction. The main elements of this device are an electron gun, a grid system to produce a retarding electric field, and a luminescent screen. In order to observe the diffraction pattern the sample is irradiated by a fairly high-intensity primary electron beam of approximately 2 keV (the beam current is a few microamperes) and inelastically reflected electrons with losses not exceeding 10% of the primary electron energy are detected. When fitted with a video camera, this device can record the diffraction patterns at fairly high speed (within ≈1 s) in a solid angle bounded by a cone with a polar expansion half-angle of 55° (Ref. 2).

Precision measurements of the Kikuchi patterns over a wider range of electron emission angles can be made using a different type of device. In this case, data are recorded ''point-by-point'' by scanning the complete hemisphere of electron reflection using a narrow-aperture energy analyzer.[4] However, this method of obtaining information is very slow (around 1.5 h is required to obtain a complete pattern) and is not efficient for practical applications.

Here we describe a new device specially designed for visualizing the structure of surface layers by recording the moderate-energy electron diffraction patterns. This device combines the possibility of enhancing the brightness of the Kikuchi pattern measured over a fairly wide solid angle with structural analysis exhibiting high surface sensitivity, which is achieved by using a primary electron beam incident on the surface of the sample at a glancing angle.[7]

The electron-optical system of this device is shown in Fig. 1. As in the electron diffraction camera considered above, the sample *1* is probed by an electron beam from an electron source *2* for which we use the standard electron gun of a 09IOS-3 Auger electron spectrometer (NITI, Ryazan). This was mounted on a separate Dy60 flange positioned at an angle of 10° to the surface of the target. The diameter of the 1.5–3 keV electron beam does not exceed 0.1 mm. The gun was fitted with a scanning system which is used to tune the device by matching the point of incidence of the beam on the target with the center of the electron-optical system.

The electrons incident on the sample are intensively scattered in the solid and are partially reflected, undergoing some energy losses. The production of a contrast interference pattern formed only by the fast reflected electrons is ensured by two concentric hemispherical grids *3* and *4* between which the retarding field creates a barrier for the insufficiently fast electrons. The first grid *3* at the same potential as the target is used to create a field-free space around the

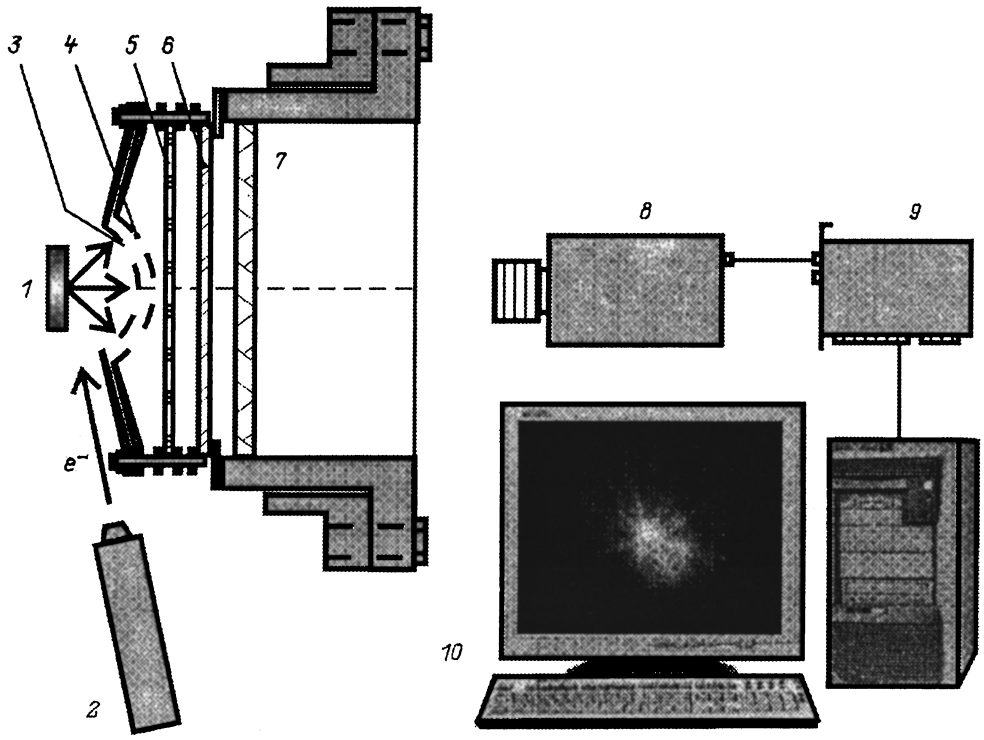


FIG. 1. Schematic of device.

sample (thereby creating conditions for stopping electrons from a point source over all velocities) while a retarding potential close to the cathode voltage of the electron gun is applied to the second grid *4*. Since the electron flux capable of overcoming the applied retarding field is small (the total coefficient of reflection of these electrons from the surface of the sample does not exceed a few percent), the recorded signal must be amplified. This problem can be solved by using a multichannel brightness amplifier *5* placed behind the retarding grid. The MKP 56-15 microchannel plate used has a diameter of 56 mm and a gain of 10^4 at a supply voltage not exceeding 1100 V. The amplified electron flux is accelerated by a 4 kV voltage applied between the output of the microchannel plate and the luminescent screen *6* and by bombarding the screen, it causes the phosphor to luminesce so that the diffraction pattern can be observed via a window *7* in the vacuum chamber.

The comparatively small size of the microchannel plate imposes stringent constraints on the grid system whose overall dimensions must be reduced (compared with standard electron-diffraction cameras) to maintain the possibility of wide-angle recording of the Kikuchi patterns. Thus, particular attention was paid to the development of the grid system. The optimum design consisted of grids of radii 12.5 and 15 mm which allowed the required patterns to be observed in a range of polar electron emission angles up to 57°. It should also be noted that in this case, unlike the devices mentioned above, there is no dead zone at the center of the screen caused by its shading by the electron gun. Miniature grids were fabricated of 85% transparent molybdenum mesh with a mesh size of 0.17 mm. To fix their shape, these grids were clamped between special mandrels and annealed in a rare gas atmosphere. This method reduced the deviations of the grids from sphericity to 0.05 mm. The entire electron-optical system was made of materials suitable for use under high-vacuum conditions. It is fairly small and mounted on a Dy100 flange, which has a 60 mm optical window at the center and microflanges with electrical voltage supply leads around the perimeter.

For testing purposes the vacuum systems were mounted in the all-metal chamber of an electron spectrometer used for complex investigations of the surface of solids. The spectrometer is fitted with an oil-free pumping system which can attain an ultrahigh vacuum (1×10^{-8} Pa) and has a transfer chamber for rapid sample loading. A transfer device moves the sample into the analysis chamber where it is fixed on a precision manipulator with three translational and one rotational degree of freedom. The manipulator can place the sample in various working positions, where its surface can be cleaned by high-temperature heating and ion etching, and the elemental composition and atomic structure of the surface can be monitored by Auger spectroscopy and low-energy electron diffraction. This manipulator can also position the sample at the focus of the electron-optical system to visualize the structure of the surface layers.

The pattern observed on the luminescence screen when the sample is bombarded with electrons is recorded by a video camera *8* mounted near the window of the vacuum system. A Sharp IS101 B/W/B uncased CCD camera was used, which can produce a television-standard black and white image with a spatial resolution of 512×582 pixels. It can also be operated in a data acquisition mode by increasing the exposure time 2^n times, where $n<8$, which helps to increase the brightness of weak patterns without the need to

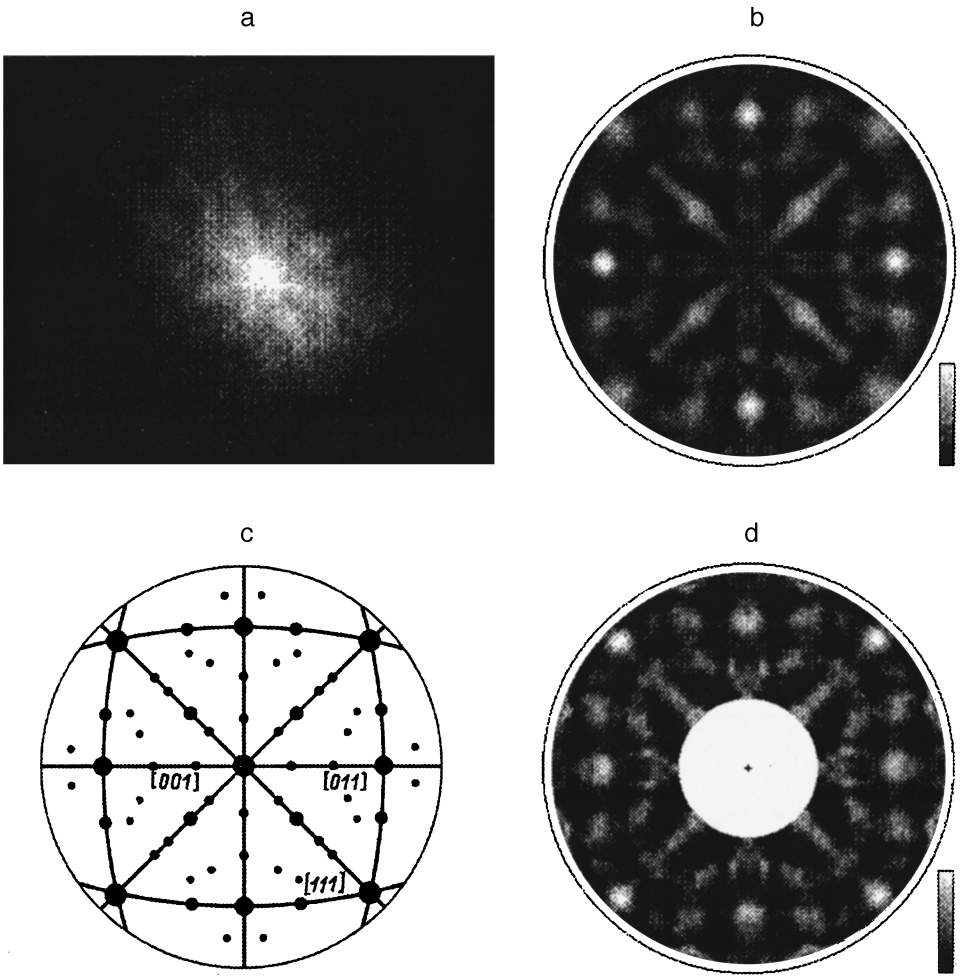


FIG. 2. a — Video frame showing Kikuchi pattern of Mo(100) single crystal obtained for a primary electron energy $E_p=2$ keV; b — the same pattern represented as a two-dimensional intensity distribution pattern of the electron reflection over the polar and azimuthal angles of electron emission (the light spots in the distribution show the orientation of close-packed chains of molybdenum atoms in real space); c — stereographic projection of the Mo(100) face showing the close-packed {100} and {110} crystal planes; d — two-dimensional distribution of the electron reflection from Mo(100) obtained by the usual method with a narrow-aperture energy analyzer.[8]

increase the primary electron current. The signal from the video camera is fed to a Cirrus Logic PV-CL5446P+ video image decoder *9* which transfers it in real time to the monitor *10* of an IBM PC. It is also possible to write individual frames of the Kikuchi patterns to hard disk using a 24 bit bitmap (.bmp format). In this graphical format, these images can be processed using standard application programs for IBM-compatible computers.

A special program, OPR-2M, was developed to extract structural information from the Kikuchi patterns. The program is written for Windows 95 and can perform the following operations: it can convert visual images of the patterns in graphical format to numerical files of two-dimensional distributions of the electron reflection intensity over polar and azimuthal angles of emission; it can produce computer displays of these distributions in graphical form using various crystallographic projections (in this case an image of the intensity at different points of the pattern on the screen is produced using a linear grayscale) and determine the spatial orientation of the focusing maxima; it can transform these two-dimensional distributions in various ways, including normalizing them to different functions of the polar angle and azimuthally averaged intensities for different polar angles, to obtain difference Kikuchi patterns, and it can rotate the patterns about arbitrary directions; and it can quantitatively compare patterns from different objects, and also compare these with results of model calculations using reliability factors and other means.

An Mo(100) single crystal was chosen as a test object to check out this device since the atomic structure of the surface of this material is well-known and the Kikuchi patterns observed for it have been analyzed in detail in our earlier studies. Data obtained for a single crystal with an atomically pure surface are illustrated in Fig. 2. Figure 2a shows a typical video frame with a Kikuchi pattern (in the form in which it appears on the screen). The pattern was recorded under conditions where the primary electron energy was 2 keV, the beam current was 0.1μA, and the voltages supplied to the retarding grid and the microchannel plate were 1.8 and 0.9 kV, respectively. The diffraction pattern has the form of an ordered system of light spots against a dark background. Their relative position reflects the spatial orientation of different close-packed rows of atoms in the sample. Thus, the symmetry of the Kikuchi pattern corresponds to the symmetry of the (100) face of a body-centered molybdenum crystal. This can be seen clearly in the central part of the image, created by electrons leaving the sample at small polar angles ($\theta<30°$). In particular, we can observe a characteristic cross-shaped region of intensified electron reflection along two pairs of intersecting {100} and {110} crystal planes. Here the light spot at their point of intersection visualizes chains of molybdenum atoms in the ⟨100⟩ direction perpendicular to

the surface. The slight displacement of this direction from the center of the screen occurred because when the sample was inserted in the device, its surface was not perpendicular to the electron-optic axis of the system,

In order to identify the characteristics of the pattern associated with other close-packed rows of atoms, we need to know the relative orientation of all the observed maxima. To obtain this information, the data given in Fig. 2a must be transformed, representing the Kikuchi pattern as a two-dimensional map of the spatial distribution of the electron reflection intensity over the polar and azimuthal angles of emission. As has been noted,the OPR-2M program has been specially developed for this purpose. It is also advisable to eliminate the appreciable decay of the luminescence intensity of the screen from the center of the image toward the edge which interferes with the overall perception of the pattern and arises because the microchannel plate has a planar rather than a spherical surface, thus symmetrizing the curves. This is usually accomplished by normalizing each azimuthal jump in the distribution to its average intensity.[5] The results of analyzing the data in Fig. 2a is shown in Fig. 2b where the required distribution is shown in a stereographic projection. The electron reflection intensity is shown using a linear gray-scale, given on the right of the figure, where the highest-intensity part of the distribution is shown white and region of lowest electron reflection is shown black. The center of the picture corresponds to the emission of electrons along the $\langle 100 \rangle$ crystal axis (in Fig. 2b the focusing maximum at $\theta = 0$ cannot be seen because the data are normalized) and the outer circle indicates the edge of the distribution with a polar emission angle of 60°.

This two-dimensional picture can now be compared with the standard stereographic projection of the (100) face of a cubic crystal which shows the relative orientation of the main crystallographic directions in single-crystal molybdenum. Note that since the surface of this material only undergoes slight relaxation, the atomic structure of the surface layers is the same as the bulk structure. To make this comparison clearer, this projection is shown in Fig. 2c. It can be seen that the data given in Figs. 2b and 2c correspond very well. In particular, the highest-intensity maxima on the observed pattern coincide with the projections of the $\langle 110 \rangle$ and $\langle 111 \rangle$ directions along which the closest-packed chains of atoms in a body-centered molybdenum crystal are oriented. Also clearly visible are regions of enhanced intensity of the pattern along the projections of the {100} and {110} planes in which lie many other, less frequent chains of atoms, the most important of which are shown by the black points in Fig. 2c (their size is proportional to the packing density of the atoms). Conversely, the minima on the pattern correspond to directions of electron emission furthest from these orientations. Thus, the pattern shown in Fig.2b actually visualizes the atomic structure of the surface region of a molybdenum single crystal in real space.

The layer thickness studied with this device depends on the energy of the incident and recorded electrons. The higher the primary electron energy and the greater the energy losses sustained by the recorded electrons, the greater is the depth of analysis. Thus, the probed depth can be controlled by varying the retarding voltage and the voltage used to accelerate the primary electrons. The thinnest surface layer is analyzed when quasi-elastically reflected electrons are recorded.[6] Its thickness is $d \approx \lambda \cos\theta_i/(1+\cos\theta_i)$, where θ_i is the polar angle of incidence of the primary beam and λ is the electron mean free path before inelastic scattering which is approximately 20–30Å at $E_p = 2$ keV. The thickness of the layer responsible for forming the Kikuchi pattern shown in Fig. 2b is approximately 15 Å.

The reliability of the data obtained using this device is confirmed by the fact that all the characteristics of this distribution show good agreement with the Kikuchi pattern obtained earlier for the same sample using a narrow-aperture energy analyzer. This can be seen clearly from a comparison of Figs. 2b and 2d, where the latter gives the Kikuchi pattern for Mo(100) from Ref. 8, also measured at $E_p = 2$ keV. An advantage of this device is that it is possible to observe the important region near the normal to the surface where Fig. 2d has a dead zone. This means that the symmetry of the pattern can be uniquely determined even for such complex objects as adsorption coatings. Thus, this device is particularly convenient for studying surface processes accompanied by changes in the symmetry of the surface layers, including recrystallization of the surface and phase transitions. Moreover, the fast recording system, capable of obtaining data at a rate of 50 images per second, allows the dynamics of the process to be studied in many cases, while the low primary electron beam current (0.1μA) minimizes the distorting influence of the electron irradiation on the objects.

This work was carried out as part of project No. 96-02-16909 of the Russian Fund for Fundamental Research supported by the Ministry of Science of the Russian Federation (under the program ''Surface Atomic Structures,'' Project No. 95-1.21).

[1] W. F. Egelhoff, Jr., Crit. Rev. Solid State Mater. Sci. **16**(3), 213 (1990).
[2] M. Erbudak, M. Hochstrasser, and E. Wetli, Mod. Phys. Lett. B **8**, 1759 (1994).
[3] M. H. Hochstrasser, M. Erbudak, A. Atrei, and E. Wetli, Phys. Low-Dimens. Semicond. Struct. **10/11**, 325 (1995).
[4] M. V. Gomoyunova, I. I. Pronin, N. S. Faradzhev *et al.*, Fiz. Tverd. Tela (St. Petersburg) **36**, 2295 (1994) [Phys. Solid State **36**, 1250 (1994)].
[5] M. V. Gomoyunova, I. I. Pronin, and N. S. Faradzhev, Zh. Eksp. Teor. Fiz. **110**, 311 (1996) [JETP **83**, 168 (1996)].
[6] N. S. Faradzhev, M. V. Gomoyunova, and I. I. Pronin, Phys. Low-Dimens. Semicond. Struct. **3/4**, 93 (1997).
[7] S. Kono, Surf. Sci. **298**, 362 (1993).
[8] N. S. Faradzhev, M. V. Gomoyunova, and I. I. Pronin, Poverkhnost'. No. 6, 117 (1997).

Translated by R. M. Durham

Mass spectrometric investigation of low-temperature diffusion and solubility of helium in lead fluoride crystals

A. G. Dudorov and A. Ya. Kupryazhkin

Urals State Technical University, 620002 Ekaterinburg, Russia
(Submitted August 11, 1997)
Zh. Tekh. Fiz. **68**, 85–89 (December 1998)

Thermal desorption of helium from presaturated crystals was used for mass spectrometric investigations of the diffusion and solubility of helium in lead fluoride crystals in the temperature range 606–714 K which precedes the transition of the crystal to the superionic state. The experimental apparatus and measurement method are described and mechanisms for the diffusion and solubility of helium in PbF_2 are discussed. © *1998 American Institute of Physics.* [S1063-7842(98)01612-2]

INTRODUCTION

Investigations of helium diffusion in crystals of rare-earth metal fluorides, which are structural analogs of the UO_2 oxide fuel in nuclear reactors, have revealed the main mechanisms for the diffusion and solubility of helium which predominate at relatively low temperatures in this type of crystal.[1–4] However, it is known[5] that as the melting point is approached in these crystals, including UO_2 (Ref. 6), melting of the anionic sublattice is observed prior to melting of the entire crystal (a phase transition to the superionic state[5]). The behavior of helium in crystals in the pretransition range has not yet been studied.

For this purpose we used lead fluoride crystals which have the lowest transition temperature to the superionic state. The smallness of the measurable effects at low temperatures necessitated improving the sensitivity and modernizing the experimental apparatus and measurement method.[1–4]

EXPERIMENTAL APPARATUS AND MEASUREMENT METHOD

The experimental investigations were carried out using thermal desorption of helium from crystals presaturated with helium under given conditions. The apparatus is shown schematically in Fig. 1. The apparatus includes a diffusion cell *2*, *3*, an admission and evacuation system *9*, *10*, *13*, *14*, a system for measuring small helium fluxes based on an updated MI-1201B mass spectrometer *1*, *6*, and a calibration system *7*, *8*, *11*. A vacuum of $\sim 10^{-7}$ Pa was provided in the system by a diffusion pump *15*.

The working cell consisted of two nickel chambers (saturation and degassing) connected by a heated high-vacuum valve. Unlike Ref. 7, the sample was moved between the chambers by means of magnetic pushers located inside each chamber. This cell allows the sample to be moved at high temperatures without precooling. The recording system of the mass spectrometer also included a secondary electron multiplier (VÉU-2) and in the pumping system, one of the oil-vapor pumps was replaced by a high-temperature getter pump to achieve a quasistatic helium regime (Fig. 1b).

Unlike Ref. 7, where liquid-nitrogen-cooled silica gel was used to achieve quasistatic pumping, and Ref. 8 where a set of metal powders, each absorbing a particular gas, was used as getters, here we utilized the capacity of titanium to absorb different chemically active gases under heating. Titanium shavings placed in a temperature gradient field ensure that all residual gases (including water vapor) are absorbed without absorbing helium, by maintaining the pressure in the measuring system better than 5×10^{-5} Pa for a long time during degassing of the sample. When the getter volume was $\sim 20\ \mathrm{cm}^3$, no deterioration in the operation of the pump was recorded after 3000 h.

Unlike Ref. 9, the mass spectrometer was calibrated by double expansion of a known quantity of gas in a calibrated volume V_1 (Fig. 1, curve *7*) in which the pressure was measured using an optical manometer with a measurement range of 0–1 Torr (*11* in Fig. 1). Helium from the volume V_1 was preliminarily expanded into the calibrated volume V_2 (*8* in Fig. 1) ($V_1/V_2\sim10^{-3}$). The gas from the volume V_2 was then fed in successive portions, using the volume V_1, back into the measuring system whose volume V_3 was also measured ($V_1/V_3\sim10^{-3}$). The resultant pressure was calculated using the formula

$$P=P_0(V_1/V_2)^2(V_2/V_3). \tag{1}$$

The calibration characteristic is linear over a wide range of pressures. The sensitivity coefficient $\gamma=P/I$, where I is the recorded mass spectrometer signal, for one series of calibrations is

$$\gamma=(8.7\pm0.8)\times10^{-8}\ \ \mathrm{Pa/mV}. \tag{2}$$

The sensitivity of the measuring system allowed degassing experiments to be carried out using crystals with a volume of $\sim 1\times10^{-3}\ \mathrm{cm}^{-3}$.

The experimental method involved saturating the crystal under study in a saturation chamber at a given temperature T and pressure P, transferring the saturated sample in vacuum into the measuring chamber, and investigating the helium degassing at temperature T using the recording system of the mass spectrometer *1*.

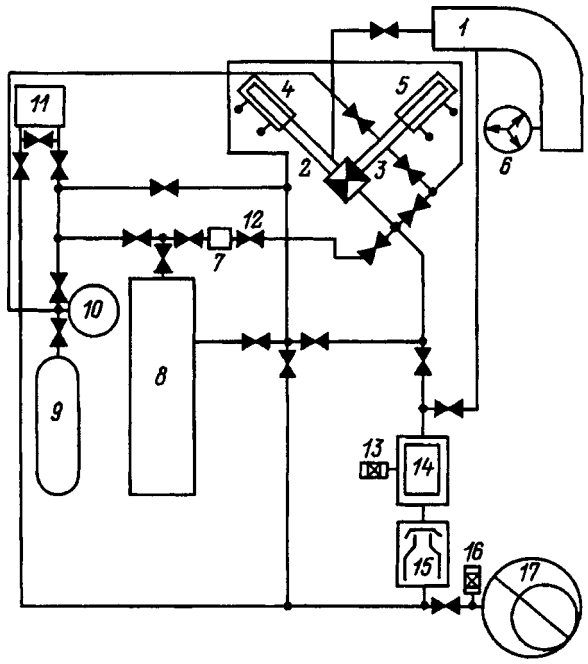


FIG. 1. Schematic of apparatus: *1* — MI-201B mass spectrometer, *2* — degassing chamber, *3* — saturation chamber, *4, 5* — heating furnaces, *6* — getter pump, *7* — calibrated volume V_1, *8* — calibrated volume V_2, *9* — gas cylinder, *10* — mechanical pressure gauge, *11* — optical pressure gauge, *12* — vacuum valves, *13* — magnetoionization pressure gauge (MID), *14* — liquid nitrogen trap, *15* — diffusion pump, *16* — thermocouple gauge (PMT-M), and *17* — roughing pump.

For the experiments we used PbF_2 single crystals grown by the Stockbarger method. The impurity content in the initial samples was less than 1×10^{-2}%. Samples of different sizes between 0.343 and 8×10^{-3} cm^3 were used to identify the influence of the surface and the presence of different-rate mechanisms of helium transport in the crystals. The saturation time was determined experimentally. In order to eliminate any influence of sample degradation on the results of the measurements, a new sample was used in each measurement.

Figure 2 shows typical experimental data for limiting temperatures and different crystal sizes. The solid curves in Figs. 2–5 are obtained by approximations to the data points. Over the entire range studied, the logarithm of the flux as a function of time is described by straight lines within experimental error, which indicates that a single helium transport mechanism is achieved experimentally.

The desorption curves were analyzed using a normal solution of the second Fick equation for a cubic sample of dimensions $a\times a\times a$ with the effective diffusion coefficient D_{eff} and solubility C_{eff} under zero boundary conditions. For the total helium desorption flux at a given temperature T from a sample presaturated in helium at pressure P and temperature T, which determine D_{eff} and C_{eff}, we obtain

$$J(t)=\frac{512\,D_{\text{eff}}\,C_{\text{eff}}\,a}{\pi^4}\sum_{n,m,k}\frac{n^2+m^2+k^2}{n^2m^2k^2}\times\exp\left(-\frac{D_{\text{eff}}\,\pi^2(n^2+m^2+k^2)t}{a^2}\right). \qquad (3)$$

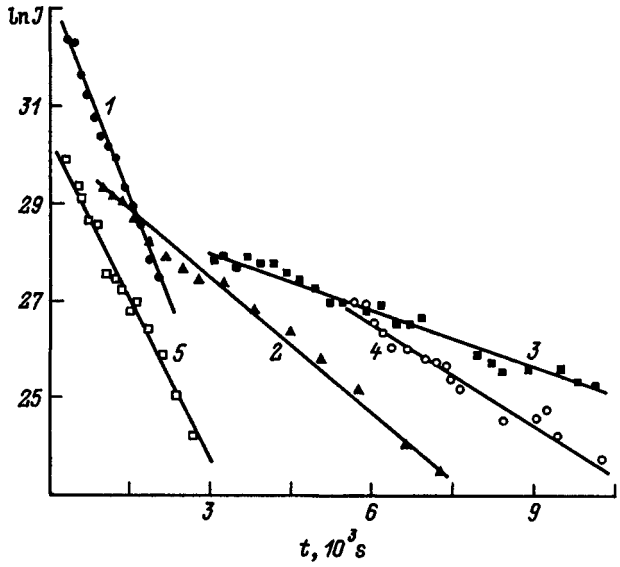


FIG. 2. Time dependence of helium flux from PbF_2 sample at saturation pressure P = 1 MPa; T=740 (*1*), 689 (*2*), 666 (*3*), 714 K (*4, 5*); a=0.1 (*1, 2, 3, 5*) and 0.2 cm (*4*).

Here n, m, and k are odd positive numbers. According to Eq. (3), for long desorption times $t>a^2/(\pi^2 D_{\text{eff}})$ the desorption flux is described by a single exponential function. Thus, a single exponential function was usually used to analyze the experimental curves (3) (Fig. 2) to obtain D_{eff} and C_{eff}.

It was established experimentally that, within experimental error, mechanical treatment of the crystal surface to give it a suitable cubic shape does not influence the values of D_{eff} obtained using Eq. (3) but does lead to a systematic overestimate of the helium solubility compared with that in cut crystals because of defects introduced in the surface layer of the sample during the mechanical grinding process. Thus, in the experiments the solubility was determined independently from the total degassing of arbitrarily shaped crystals

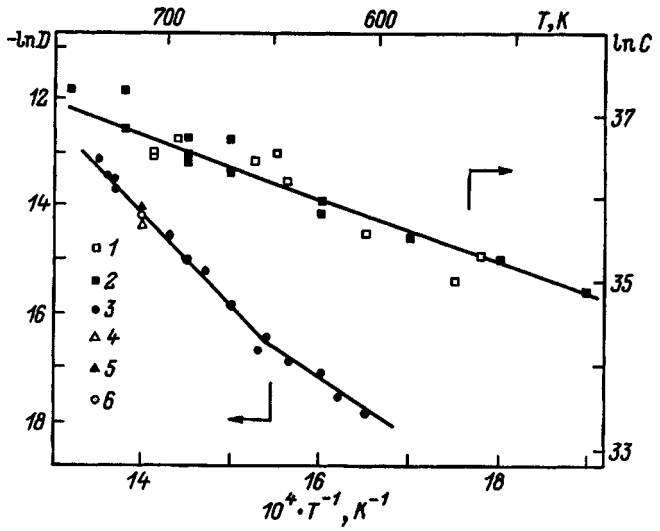


FIG. 3. Temperature dependence of the helium diffusion coefficient and solubility in PbF_2: *1* — PbF_2+0.5% HoF_3, *2–6* — PbF_2+less than 1×10^{-3}% HoF_3 ; a=0.15 (*4*), 0.2 (*5*), and 0.1 cm (*6*).

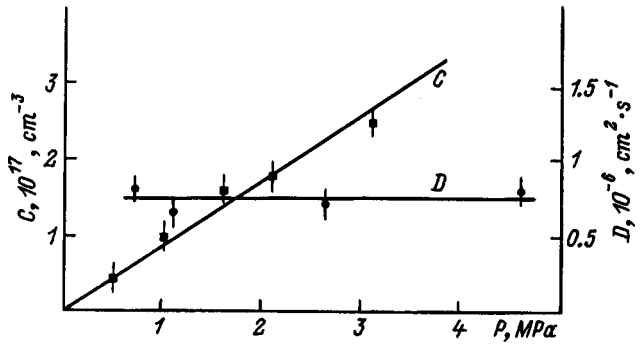


FIG. 4. Pressure dependence of helium diffusion coefficient and solubility in PbF_2 ($T=714$ K).

cut from the main single crystal. The fact that within experimental error, the surface (size) of the sample has no influence on the diffusion coefficients obtained (Figs. 2–4) indicates that it is correct to select zero boundary conditions to solve the diffusion problem.

EXPERIMENTAL RESULTS AND DISCUSSION

Figure 3 gives results of experimental investigations of the helium diffusion coefficient and solubility in PbF_2 at temperatures between 606 and 741 K before the crystal is transferred to the superionic state. Over the entire range of temperatures studied the effective solubility is described by a simple exponential dependence with a single dissolution energy ($E^p=0.36\pm0.04$ eV)

$$C_{\text{eff}}=(3.5^{+1.7}_{-1.2})\times10^{18}\exp\left(-\frac{0.36\pm0.03eV}{kT}\right),\ \text{cm}^{-3}. \quad (4)$$

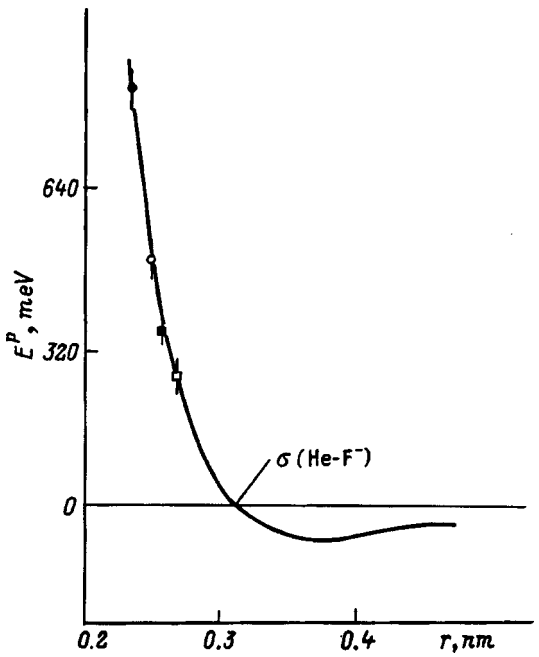


FIG. 5. Energy of interstitial solubility of helium in calcium, strontium, lead, and barium fluorides as a function of the distance between the helium atom and fluorine ions: ● — CaF_2, ○ — SrF_2, ■ — PbF_2, □ — BaF_2.

In order to determine the mechanism responsible for the solubility of helium in lead fluoride, we also investigated the influence of Ho^{3+} impurities and the helium saturation pressure on the solubility. The experiments indicate that within experimental error, the solubility in samples containing 0.5% HoF_3 in the initial charge is the same as that for pure samples (Fig. 3). The solubility of helium in these crystals at $T=714$ K and $P=0$–3 MPa (Fig. 4) obeys Henry's law and is described by

$$C_{\text{eff}}=(8.6\pm0.4)\times10^{10}P,\ \text{cm}^{-3}. \quad (5)$$

The fact that within experimental error, the helium solubility does not depend on the influence of impurities or the dislocation density (the samples were cut from different parts of the single crystal having different dislocation densities) and the linear dependence of the solubility on pressure indicate that the helium solubility in lead fluoride is predominantly interstitial. A helium atom is in the neighborhood of the nearest eight fluorine anions which are situated at the distance $r=d\sqrt{3}/2$, where $d=0.595$ nm is the distance between the fluorine anions in the fluorine sublattice of the crystal.[10] The approximation of an undeformed lattice is justified for crystals of this type because of the fairly large sizes of the interstitial sites compared with the size of the helium atom.[2] Moreover, in fluorite-structure crystals the helium dissolution energy is preferentially determined by its interaction with the nearest fluorine anions,[1] while the dependence of the helium dissolution energy in different crystals on the distance between the helium atoms and the nearest fluorine ions, which is determined by the lattice constant, can be used to reconstruct the parameters of the interaction potential of an He–F pair in the pair interaction approximation. In this approximation the dissolution energy is equal to the energy of interaction between a helium atom and fluorine anions $E^P=8U(d\sqrt{3}/2)$, where $U(r)$ is the potential of a pair interaction between a helium atom and a fluorine anion.

In Refs. 11 and 12, the interaction potential of an He–F pair was reconstructed from the results of gas-phase measurements (diffusion of fluorine ions in helium in electric fields) and data on the interstitial solubility of helium in CaF_2, SrF_2, and BaF_2 (Ref. 2). The result of approximating the combined experimental data is given in Ref. 13 as

$$U(r)=(69.7\pm2.8)\exp\left\{-\frac{r}{0.0408\pm0.007}\right\}-\frac{(3.30\pm0.31)\times10^{-4}}{r^4},\ \text{eV}. \quad (6)$$

For comparison Fig. 5 gives data on the interstitial solubility of helium in calcium, strontium, lead, and barium fluorides and an approximation of the dependence $E^p(r)=8U(r)$ using relation (6). This comparison shows that the data obtained here for the energy of helium dissolution in lead fluoride (4) are fully consistent with the mechanism of interstitial solubility.

The temperature dependence of the effective diffusion coefficient reveals two temperature regions with different diffusion activation energies E^D. The data were analyzed

TABLE I. Comparison of data for helium diffusion coefficients in strontium, lead, and barium fluorides.

System No.	System	d, nm	ΔT, K	D_0, cm^2/s	E^0, eV	References
1	$He–SrF_2$	0.580 Ref. 14	930–1243	$(0.74^{+0.19}_{-0.15})\times10^{-2}$	0.75 ± 0.02	Ref. 2
2	$He–PbF_2$	0.595 Ref. 10	606–654	$(1.5^{+12.3}_{-1.3})$	0.95 ± 0.11	Present study
3	$He–PbF_2$	0.595 Ref. 10	654–741	$(1.2^{+2.3}_{-0.8})\times10^{4}$	1.58 ± 0.07	Present study
4	$He–BaF_2$	0.620 Ref. 14	903–1233	$(0.45^{+0.07}_{-0.06})\times10^{-2}$	0.69 ± 0.02	Ref. 2

independently for the low-temperature range $T = 606–654$ K and the high-temperature range $T = 654–741$ K using the simple exponential dependence

$$D_{\text{eff}} = D_0 \exp\left\{-\frac{E^D}{kT}\right\}. \quad (7)$$

The values of D_0 and E^D are given in Table I.

Investigations of the influence of the saturation pressure on the coefficient of helium diffusion showed (Fig. 4) that within experimental error, this is constant. At $T = 714$ K we find $D(P) = 7.3\pm0.4\times10^{-7}$ cm^2/s.

The mechanisms of helium diffusion in PbF_2 were analyzed by comparing our results with those obtained in studies of interstitial helium diffusion in the structural analogs SrF_2 and BaF_2 (see Table I), whose crystal lattices have larger or smaller periods than PbF_2, respectively.[14]

The diffusion activation energy for PbF_2 in the low-temperature range (No. 2 in Table I) is higher than the activation energy for interstitial diffusion for BaF_2 and for SrF_2. The preexponential factor of the diffusion coefficient for PbF_2 in this range (No. 2 in Table I) is also higher than that for BaF_2 and SrF_2. This indicates that these values evidently correspond to an interstitial diffusion mechanism with ''traps'' (see, for example, Ref. 15) which may comprise anionic vacancies.

An increase in the diffusion activation energy and the preexponential factor of the diffusion coefficient in the high-temperature range (No. 3 in Table I) for the same interstitial dissolution mechanism can evidently be attributed to a more complex transport mechanism, such as an expulsion mechanism (see, for example Ref. 16), where an interstitial helium atom is transferred to a neighboring free interstitial site via an anionic vacancy formed by the preliminary expulsion of a fluorine anion into this or a neighboring free interstitial site. A large number of particles may take part in this diffusive displacement, causing local disordering in the crystal. This mechanism can explain the high values of the preexponential factor and the effective diffusion activation energy.

To sum up, this system has been used to investigate helium diffusion and solubility in single-crystal PbF_2 samples with a volume of $\sim1\times10^{-3}$ cm^3 at temperature between 606 and 741 K. Temperature dependences of the helium transport coefficients have been obtained in the range preceding the transition to the superionic state. It has been shown that the helium solubility in this range corresponds to predominantly interstitial sites. The low-temperature part of the diffusion coefficient corresponds to an interstitial mechanism with traps while the high-temperature section corresponds to a complex diffusion mechanism of the expulsion type.

This work was partly supported by the Russian Fund for Fundamental Research (Project No. 95-02-03649).

[1] A. Ya. Kupryazhkin, A. U. Kurkin, and O. V. Semenov, J. Nucl. Mater. **204**, 180 (1994).
[2] A. Ya. Kupryazhkin and E. V. Popov, Fiz. Tverd. Tela (Leningrad) **26**, 160 (1984) [Sov. Phys. Solid State **26**, 94 (1984)].
[3] A. Ya. Kupryazhkin and A. Yu. Kurkin, Fiz. Tverd. Tela (Leningrad) **32**, 2349 (1990) [Sov. Phys. Solid State **32**, 1363 (1990)].
[4] A. Ya. Kupryazhkin, A. Yu. Kurkin, and A. G. Dudorov, Fiz. Tverd. Tela (St. Petersburg) **38**, 1272 (1996) [Phys. Solid State **38**, 703 (1996)].
[5] *Physics of Superionic Conductors*, edited by M. B. Salamon (Springer-Verlag, New York, 1979; Zinatne, Riga, 1982, 315 pp.).
[6] L. V. Matveev and M. S. Veshchunov, Zh. Éksp. Teor. Fiz. **111**, 585 (1997) [JETP **84**, 322 (1997)].
[7] E. V. Popov and A. Ya. Kupryazhkin, Zh. Tekh. Fiz. **53**, 365 (1983) [Sov. Phys. Tech. Phys. **28**, 224 (1983)].
[8] M. G. Pan'yan, K. A. Petrzhak, and V. F. Teplykh, Prib. Tekh. Éksp. No. 4, 250 (1971).
[9] E. V. Popov and A. Ya. Kupryazhkin, Prib. Tekh. Éksp. No. 4, 194 (1982).
[10] B. H. Schulz, E. Perenthaler, and U. H. Zuker, Acta Crystallogr. **38**, 729 (1982).
[11] A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Springer-Verlag, Berlin, 1985; Atomizdat, Moscow 1986, pp. 319–326.).
[12] V. L. Bychkov, A. A. Radtsig, and B. M. Smirnov, Teplofiz. Vys. Temp. **16**, 713 (1978).
[13] A. Ya. Kupryazhkin and A. Yu. Kurkin, Fiz. Tverd. Tela (St. Petersburg) **35**, 3003 (1993) [Phys. Solid State **35**, 1475 (1993)].
[14] W. Hayes, *Crystals with Fluorite Structure* (Clarendon Press, Oxford, 1974, 450 pp.).
[15] A. Ya. Kupryazhkin, I. R. Shein, and E. V. Popov, Zh. Tekh. Fiz. **53**, 1578 (1983) [Sov. Phys. Tech. Phys. **28**, 970 (1983)].
[16] *Atomic Diffusion in Semiconductors*, edited by D. Shaw (Plenum Press, London, 1973; Mir, Moscow, 1975, 682 pp.).

Translated by R. M. Durham

BRIEF COMMUNICATIONS

Numerical analysis of the influence of metallization nonuniformity on the thermal regime of microcircuits

E. P. Taran and V. V. Starostenko

Simferopol State University, 333022 Simferopol, Ukraine
(Submitted January 13, 1998; resubmitted June 23, 1998)
Zh. Tekh. Fiz. **68**, 90–92 (December 1998)

A numerical analysis of the thermalized regime of metallized sections has revealed that local overheating occurs at sites of nonuniformity together with short circuiting, which results in failure of the microcircuits. © *1998 American Institute of Physics.* [S1063-7842(98)01712-7]

The current state of the art in semiconductor electronics is leading to more stringent requirements for the reliability of semiconductor elements, especially integrated microcircuits. This has given rise to the problem of modeling the physical processes which cause microcircuit failure. One failure classification, reported in Ref. 1, shows that the vast majority of failures are caused by metallization and contact defects, defective fitting, and vaporization of internal leads. The Department of Radio Physics at Simferopol State University has developed a thermal model to calculate the processes leading to local burnout of the metallization when current flows.

The numerical calculations of the thermal regime were based on solving the nonlinear heat conduction equation. Metallization nonuniformity caused by nonuniform diffusion of silicon into the aluminum was only assumed along the length of the metallization. As a result, the dynamics of the thermal processes were described by the one-dimensional heat conduction equation[2]

$$c\rho\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(K(x)\frac{\partial T}{\partial x}\right)+Q, \tag{1}$$

where c is the specific heat of the metallization mass, ρ is the density, K is the thermal conductivity, T is the temperature field of the metallization, and Q is the density of heat sources (sinks).

A voltage U produced by an electric field E was applied to the metal track

$$U=E\cdot d, \tag{2}$$

where d is the distance between the microcircuit outputs.

In order to analyze the thermal processes in the integrated microcircuit, the metallization was simulated by series-connected resistances R_i. As a result, the power of the thermal sources was given by

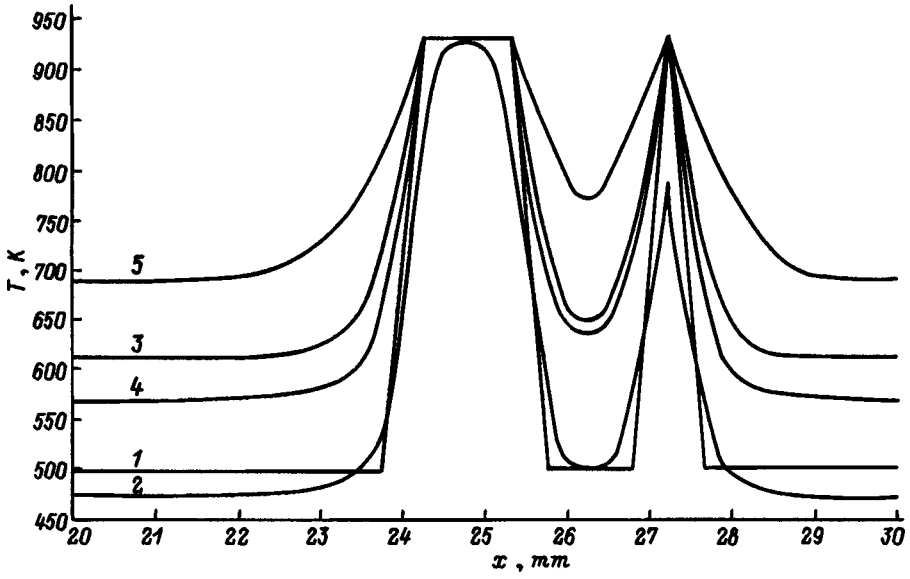


FIG. 1. Temperature field of metallization at various times: *1* — 1 μs, *2–5* — 1, 1.001, 2, and 2.001 μs, respectively.

$$Q_i = \frac{U^2}{h^2 \sigma_i \left(\sum_{i=1}^{N} I/\sigma_i \right)^2}, \tag{3}$$

where σ_i is the electrical conductivity of the ith section of the metallization, N is the number of metallized sections, and h is the length of the homogeneous section.

The electrical conductivity σ_i is a function of temperature and is given by the following expression:[3]

$$\sigma_i = \frac{\sigma_{0i}}{I + \alpha \cdot T_i}, \tag{4}$$

where σ_{0i} is the electrical conductivity of the ith metallized section at 0 and $\alpha \cong I/273$ is the temperature coefficient of resistance of the metallization.

The density of effective emission sinks Q_s was introduced to allow for heat exchange between the microcircuit and the ambient medium. It was assumed in the model that heat exchange with the ambient medium obeys Newton's law,[3]

$$Q_s = h_t (T - \theta), \tag{5}$$

where h_t is the coefficient of heat exchange, $T(t,x)$ is the temperature field of the metallization, and θ is the temperature of the ambient medium.

Heat exchange between the metallization boundaries and the ambient medium was taken into account using Newton's law[3]

$$\left(\frac{\partial T}{\partial x} \right)_L = \lambda \left[T|_L - \theta \right], \tag{6}$$

where $\lambda = h_t/K$ is the normalized coefficient of heat exchange, K is the thermal conductivity, $\theta(t)$ is the temperature of the ambient medium, and L is the metallization boundary.

The temperature dependence of the thermal conductivity is given by the Wiedemann–Franz formula[4]

$$\frac{K(x,t)}{\sigma(x,t)} = \alpha \cdot T, \tag{7}$$

where α is a proportionality factor which does not depend on the type of metal.

The nonlinear heat conduction equation (1) was solved by an implicit scheme (advancing scheme).[2] At the melting point, calculations were made of the mass of molten metal

$$m_p = \frac{Q_p}{\lambda_p}, \tag{8}$$

where Q_p is the quantity of heat expended in melting the metal and λ_p is the specific heat of fusion.

A periodic electric field of amplitude 50 kV/m, duration 1 μs, and inverse duty cycle of 1000 was applied to the metallization to study the influence of nonuniform sections of metallization on the thermal regime of an integrated microcircuit.

The presence of nonuniform sections in the metallization resulted in these being locally heated (Fig. 1, curves *1–5*). The rates of the heat conduction and heat exchange processes were much lower than the rate of heat buildup and thus the metallization became heated within a single electric field pulse (curve *2* in Fig. 1).

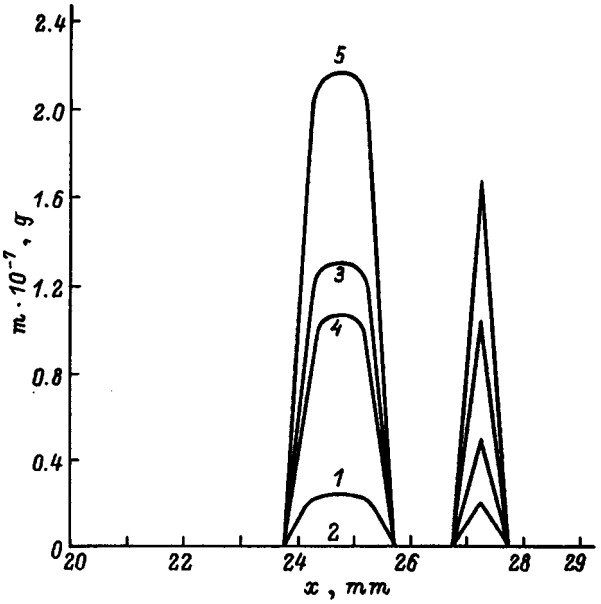


FIG. 2. Distribution of molten sections of metallization as a function of time: *1–5* as in Fig. 1.

Local overheating of sections of the metal track caused some of the metal in these sections to melt (Fig. 2). The buildup of heat caused further melting of the aluminum and after the action of the second and third pulses on sections containing inhomogeneities, most of the metal was already molten (Fig. 2, curves *1–4*).

The rate of the heat conduction process was so low that it only began to influence the thermal processes after the second pulse. In this case, however, the processes of ''heat spreading'' over the track were fairly bounded.

A numerical analysis of the influence of metallization nonuniformity has shown that when a current flows along the metal track, local heating occurs at sites of nonuniformity. The presence of local sections of overheating and nonlinear thermal conductivity causes a buildup of heat in a pulsed regime. This buildup of heat leads to melting of the crystalline aluminum at sites of local overheating and may cause short circuiting of the metallization through the oxide film, resulting in failure of the microcircuits.

[1] A. A. Chernyshev, *Principles of the Reliability of Semiconductor Devices and Integrated Microcircuits* [in Russian], Radio i Svyaz', Moscow (1988), 256 pp.
[2] A. A. Samarskiĭ, *Theory of Difference Schemes* [in Russian], Nauka, Moscow (1977), 656 pp.
[3] N. M. Tugov, B. A. Glebov, and N. A. Charykov, *Semiconductor Devices* [in Russian], Énergoatomizdat, Moscow (1990), 576 pp.
[4] A. G. Shashkov, V. A. Bubnov, and S. Yu. Yanovskiĭ, *Wave Effects of Heat Conduction: System-Structural Approach* [in Russian], Nauka i Tekhnika, Minsk (1993), 279 pp.

Translated by R. M. Durham

Model of current oscillations in a metal–thin insulator–semiconductor structure

I. I. Abramov and A. L. Danilyuk

Belarus State University of Informatics and Radioelectronics, 220027 Minsk, Belarus
(Submitted February 2, 1998)
Zh. Tekh. Fiz. **68**, 93–94 (December 1998)

A model is proposed for current oscillations in a metal–thin insulator–p-type semiconductor structure which are caused by the self-organization of carrier transport processes via insulator states. [S1063-7842(98)01812-1]

We shall analyze electron transport in a metal–thin insulator–p-type semiconductor structure under conditions where the barrier at the metal–insulator interface is low, the surface state density is low, and the insulator contains monoenergetic donor and acceptor states at a distance s from the interface with the metal. The current in the structure is determined by tunneling of carriers via acceptor and donor states in the insulator and the semiconductor bands. The charge distribution at states in the insulator is simulated by a charged plane. The energy of the acceptor states E_{nt} under application of a constant external potential corresponds to the band gap and the energy of the donor states E_{pt} corresponds to the valence band.

Making these assumptions and using the model of electron capture by tunneling to insulator states at the semiconductor–insulator interface,[1] we write a system of equations for the rates of change of the electron and hole densities at the insulator states

$$\frac{dn_t}{dt}=\alpha_{na}N_{nt}T_{na}n_m[1-(n_t/N_{nt})(N_m/n_m)]-\beta_n n_t-k_R n_t p_t, \tag{1}$$

$$\frac{dp_t}{dt}=\alpha_p N_{pt}T_{pd}[p_{S0}\exp(-\delta Y_S)-(p_t/N_{pt})N_V] - \beta_p p_t-k_R n_t p_t-\alpha_{nd}k_R p_t T_{nd}n_m, \tag{2}$$

where α_{na} and α_{nd} are the coefficients of electron capture by acceptor and donor insulator states, α_p is the coefficient of hole capture by donor insulator states, N_{nt} and N_{pt} are the densities of acceptor and donor insulator states, β_n and β_p are the coefficients of thermal emission of electrons and holes from insulator states, T_{na} and T_{nd} are the transparencies of the tunnel barriers to transition of electrons from the metal to acceptor and donor insulator states, T_{pd} is the transparency of the tunnel barrier to transition of holes from the valence band to donor insulator states, k_R is the coefficient of recombination of electrons and holes at insulator states, N_m and n_m are the density of states and the electron density in the metal corresponding to the energy of the acceptor insulator states, p_{S0} is the density of holes at the surface of the semiconductor in the absence of any charge oscillations of the insulator state, δY_S is the change in the surface potential of the semiconductor when the charge at the insulator states varies ($\delta Y_S<0$), and N_V is the density of states in the valence band of the semiconductor.

The transparencies of the tunnel barriers are determined using the Wentzel–Kramers–Brillouin approximation. When the insulator has a charged plane, we write this as

$$T_n=\exp\left[-\frac{2}{\hbar}\int_a^s\sqrt{2m\{U(x)-E-q^2|a-s|(n_t-p_t)\times(1/2\varepsilon_0\varepsilon_d)[1-\psi(x)]\}}\,dx\right], \tag{3}$$

where $\hbar$ is Planck's constant, E and $U(x)$ are the tunneling energy and the potential profile of the tunnel barrier in the absence of charge at the insulator states, m is the effective electron mass, ε_0 and ε_d are the absolute and relative permittivities of vacuum and the insulator, q is the electron charge, ψ is the distribution function of the potential energy of the insulator state charge ($0<\psi(x)\leqslant 1$), $a=0$ for tunneling of electrons from the metal to the insulator state, $a=d$ for tunneling from the semiconductor, and d is the thickness of the insulator.

We expand the integrand (3) as a series. To within first-order terms we have

$$T_n=\exp\left[-\frac{2}{\hbar}\int_a^s\sqrt{2m[U(x)-E]}\,dx\right]\times\exp\left[\frac{q^2m|a-s|(n_t-p_t)}{\varepsilon_0\varepsilon_d\hbar}\int_a^s\frac{[1-\psi(x)]dx}{\sqrt{2m[U(x)-E]}}\right]. \tag{4}$$

From Eq. (4) we obtain

$$T_{na}=T_n \quad \text{for} \quad m=m_e \quad \text{and} \quad E=E_{nt},$$
$$T_{nd}=T_n \quad \text{for} \quad m=m_e \quad \text{and} \quad E=E_{pt},$$
$$T_{pd}=T_n \quad \text{for} \quad m=m_p \quad \text{and} \quad E=E_{pt}, \tag{5}$$

where m_e and m_p are the effective masses of electrons in the metal and holes in the semiconductor.

We determine the change in the surface potential from the condition

$$q(n_t-p_t)=(\sqrt{2}\varepsilon_0\varepsilon_S kT/qL_D)F(\delta Y_S), \tag{6}$$

where ε_S is the relative permittivity of the semiconductor, k is the Boltzmann constant, T is the temperature, L_D is the

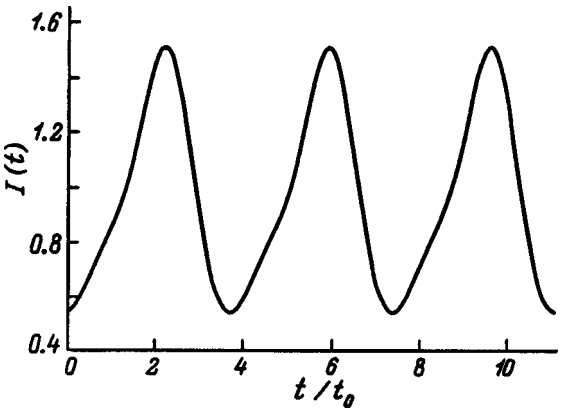


FIG. 1. Oscillations of the relative current through the insulator states in an Al–SiO_2–p-Si structure ($t_0=0.045$ s).

Debye length of holes in the semiconductor, and $F(\delta Y_S)$ is the band bending function for a nondegenerate semiconductor.

The current through the insulator states is proportional to the rate of capture of electrons from the metal to the insulator states or the rate of hole capture from the semiconductor to these insulator states. In this case, the onset of current oscillations is attributed to the existence of stable periodic oscillations of the insulator state charge $Q_D=q(n_t-p_t)$. In the mathematical description the establishment of self-oscillations from steady states corresponds to a cycle creation bifurcation (or a Hopf bifurcation).[2] These calculations revealed that the system of equations (1)–(6) can have stable and unstable periodic solutions for various values of the parameters and the relative magnitude of the current oscillations $I(t)$ through the insulator states is 5–100%.

The mechanism for the formation of these oscillations is as follows. Initially, the amplitude of the electron charge oscillations at the insulator states increases since the rate of electron capture is higher than the rate of hole capture. As the electron density in these states increases, these rates increase and at a certain time it is found that $(dp_t/dt) > (dn_t/dt)$ and the amplitude of the electron charge oscillations at the insulator states begins to decrease. This is accompanied by an increase in the recombination velocity of the electrons and holes at the insulator states as their densities increase. Recombination leads to a drop in the electron and hole densities at the insulator states and conditions are reestablished where the rate of electron capture exceeds the rate of hole capture. By way of example we consider an Al–SiO_2–p-Si structure at room temperature. The hole concentration in the bulk of the silicon is $2\times10^{21}\,\mathrm{m}^{-3}$ and the insulator thickness is 2×10^{-9} m. The density of insulator states is $(1-6)\times10^{16}\,\mathrm{m}^{-2}$, the barrier height at the metal–insulator interface is 3.2 eV, and the forward fixed bias is 0.4–0.6 V. Depending on the parameters of the structure, the oscillations have a period of 0.01–1.0 s and $I(t)$ varies by as much as 50% (see Fig. 1).

As confirmation of these calculations, we note the experimental results reported by Farmer *et al.*,[3] who studied the temperature dependence of low-frequency fluctuations of the resistance of a similar Al–SiO_2–p-Si structure and showed that this fluctuation reaches 10%. These results[3] can be explained in terms of the proposed model by the fact that an increase in temperature results in self-organization of carrier transport processes which is characterized by the onset of oscillations of the charge and current through the insulator states. The contribution of these oscillations to the total current depends on the ratios between the current through the insulator states and the currents through the valence band and the conduction band, which also determine the amplitude of the low-frequency fluctuations of the structure resistance. Note that the calculated data obtained for specific values of the parameters agree with these experimental results. It is impossible to make a more detailed comparison because the ratios between the currents and the values of various parameters are not given in Ref. 3.

To sum up, we have developed a model of current oscillations in a metal–thin insulator–p-type semiconductor structure which are caused by self-organization of carrier transport processes via insulator states.

This work was partially financed by the Republican Inter-University Program for Fundamental Research in the Natural Sciences ''Physical Principles of the Production, Diagnostics, Functioning, and Application of Low-Dimension Elements and Systems,'' Grant No. 04.04.

[1] *Properties of Metal–Insulator–Semiconductor Structures*, edited by A. V. Rzhanov [in Russian], Nauka, Moscow (1976).
[2] B. D. Hassard, N. D. Kazarinoff, and Y.-H. Wan, *Theory and Applications of Hopf Bifurcation* (Cambridge University Press, Cambridge 1981).
[3] K. R. Farmer, C. T. Rogers, and R. A. Buhrman, Phys. Rev. Lett. **58**, 2255 (1987).

Translated by R. M. Durham

Propagation velocity of the normal zone in a cooled composite superconductor with an external coating

V. K. Ozhogina and V. R. Romanovskiĭ

''Kurchatov Institute'' Russian Science Center, 123182 Moscow, Russia
(Submitted February 2, 1998)
Zh. Tekh. Fiz. **68**, 95–97 (December 1998)

An investigation was made of the influence of a nonconducting shell on the velocity of irreversible propagation of the normal zone along a cooled composite superconductor. It is shown that even a coating of negligible thickness may lead to a substantial deterioration in the thermal stabilization conditions of intensively cooled current-carrying elements. © *1998 American Institute of Physics.* [S1063-7842(98)01912-6]

An investigation of the processes taking place in superconducting magnet systems undergoing an irreversible transition to the normal state is one of the most important problems in technical superconductivity. This problem is topical because of the widespread use of various superconducting windings having a high level of stored energy. It is well known that the thermal instability observed in a superconductor as a result of the action of an external perturbation whose energy exceeds the so-called critical value, is accompanied not only by its propagation through the superconductor. Within a fairly short time a thermal wave forms, propagating at constant velocity, and against this background the superconducting properties of the current-carrying element sustain irreversible damage. The critical energies and velocities of the thermal wave are usually investigated neglecting the insulating coating on the lateral surface of a composite superconductor.[1,2] However, the results reported in Refs. 3 and 4 indicate that such a coating can alter the kinetics of the normal zone, not only quantitatively but also qualitatively. This means that the theory developed so far has limited application.

We shall analyze a extended cooled composite (a superconductor in a normally conducting matrix) whose outer surface is surrounded by a nonconducting shell. To simplify the analysis, we shall assume that the boundary conditions of this composite are planar, the transverse dimensions of the composite d_k and the coating d_i are small, there is no contact thermal resistance between the coating and the composite, and the thermal and electrical parameters do not depend on temperature. The change in the temperature field of the entire composite in the longitudinal direction may be described in terms of the model of a continuous medium using the one-dimensional time-dependent heat conduction equation

$$C_e\frac{\partial T}{\partial t}=\lambda_e\frac{\partial^2 T}{\partial x^2}-\frac{h_e p}{S}(T-T_0)+\frac{I^2}{S^2}\rho_e(T) \tag{1}$$

with the averaged coefficients $C_e=C_k d_k/d+C_i d_i/d$, $\lambda_e=\lambda_k d_k/d+\lambda_i d_i/d$, $h_e^{-1}=h^{-1}+d_i/\lambda_i$, and $\rho_e^{-1}=d_k/\rho_k d+d_i/\rho_i d\cong d_k/\rho_k d(\rho_i\gg\rho_k)$. Here C_k, C_i are the volume specific heats of the composite and the coating, λ_k and λ_i are their thermal conductivities, S is the cross-sectional area, p is the cooled perimeter, h is the coefficient of heat transfer, I is the transport current, T_0 is the temperature of the ambient medium, $\rho(T)$ is the effective electrical resistivity of the superconducting composite[1,2]

$$\rho(T)=\rho_0(T)\begin{cases}1, & T>T_{CB},\\ (T-T_C)/(T_{CB}-T_C), & T_c\leqslant T\leqslant T_{CB},\\ 0, & T<T_C=T_{CB}-(T_{CB}-T_0)I/I_C,\end{cases}$$

where ρ_0 is the electrical resistivity of the matrix, I_C and T_{CB} are the critical parameters of the superconductor.

Assuming that $L_k=[\lambda_k S^2(T_{CB}-T_0)/I_C^3\rho_0]^{1/2}$, we introduce the dimensionless variables

$$X=x/L_k,\quad i=I/I_C,\quad \tau=\lambda_k t/(C_k L_k^2),$$

$$\theta=(T-T_0)/(T_{CB}-T_0),$$

for which Eq. (1) has the form

$$\left(1+C\frac{\Delta_i}{\Delta_k}\right)\frac{\partial\theta}{\partial\tau}=\left(1+\Lambda\frac{\Delta_i}{\Delta_k}\right)\frac{\partial^2\theta}{\partial X^2}-\frac{\theta}{\alpha+\Delta_i\Delta_k/\Lambda}+i^2 r(\theta). \tag{2}$$

Here we have $C=C_i/C_k$, $\Lambda=\lambda_i/\lambda_k$, $\Delta_{i,k}=d_{i,k}/L_k$,

$$\alpha=I_C^2\rho_0/hpS(T_{CB}-T_0),$$

$$r(\theta)=i^2\begin{cases}1, & \theta>1,\\ (\theta-1+i)/i, & 1-i\leqslant\theta\leqslant 1,\\ 0, & \theta<1-i.\end{cases}$$

Even at the stage of formulating the problem, this generalized description of the thermal processes in a composite superconductor in a shell can identify clearly and simply characteristic dimensionless complexes which may have a substantial influence on the kinetics of the normal zone. These qualitative results can easily be compared with the conclusions formulated earlier using thermal stabilization theory,[1,2] since the model (2) not only uses generally accepted terminology but also comprehensively describes possible limiting cases. According to the model (2), the shell has a negligible influence on the temperature field of the com-

posite if $C\Delta_i/\Delta_k \ll 1$, $\Lambda\Delta_i/\Delta_k \ll 1$, and $\alpha \gg \Delta_k\Delta_i/\Lambda$. It is also easy to understand the part played by the thermophysical parameters of the coating and its thickness in the kinetics of the normal zone. It is easily shown that in the case $\Delta_i/\Delta_k \ll 1$, which is satisfied by the parameters of real current-carrying elements, the specific characteristics of the evolution of the thermal processes in these elements will be determined by the value of the dimensionless complex $\Delta_k\Delta_i/\Lambda$, which is inversely proportional to the dimensionless thermal resistance of the shell. Thus, it can be affirmed that for a wide range of practical applications, allowance for the specific heat in calculations of the normal zone velocity is less important than taking into account the possible influence that the thermal conductivity and thickness of the coating may have on its value. As will become obvious from the results presented below, this conclusion becomes even more accurate, the smaller the current and heat transfer coefficient.

We shall estimate the possible range of variation of the dimensionless parameters used. Assuming that for a niobium titanium superconductor in a copper matrix with an organic coating, the average values of the initial coefficients are $C_k \sim 5\times10^{-3}$ J/cm$^3\cdot$K, $\lambda_k \sim 1$ W/cm$\cdot$K, $d_k \sim 0.1$ cm; $C_i \sim 10^{-2}$ J/cm$^3\cdot$K, $\lambda_i \sim 10^{-3}$ W/cm$\cdot$K, $d_i \sim 5\times10^{-3}$ cm; $\rho_0 \sim 5\times10^{-8}\,\Omega\cdot$cm, $T_{CB}-T_0 \sim 5$ K, and $I_C \sim 10^3$ A, we find $C \sim 2$, $\Lambda \sim 10^{-3}$, $\Delta_k \sim 1$, and $\Delta_i \sim 5\times10^{-2}$.

We shall use the model (2) to determine the velocity of the normal zone propagating along the composite after a thermal wave has formed. According to Truck,[5] this can be calculated using the formula

$$V = \frac{\sqrt{1+\Lambda\Delta_i/\Delta_k}}{1+C\Delta_i/\Delta_k}\left[\frac{\alpha_e i^2+i-2}{\alpha_e(\alpha_e i^2-1)} + \sqrt{\frac{\alpha_e-1}{\alpha_e}}\left(2-2\sqrt{\frac{1-i}{1-i_s}}-\frac{i-i_s}{1-i_s}\right)\right],$$

where $\alpha_e = \alpha + \Delta_k\Delta_i/\Lambda$ and $i_s = (\sqrt{1+8\alpha_e}-1)/2\alpha_e$.

Figure 1 gives the normal zone velocities as a function of the current for various values of the coating thickness calculated for $\alpha=2$, $C=1$, $\Lambda=10^{-3}$, $\Delta_k=1$ (dashed curves). Also plotted for comparison are the curves $V(i)$ for to a composite superconductor without an insulating coating, calculated for three typical values of the stabilization parameter $\alpha=2$ — ''good'' cooling, $\alpha=100$ — ''poor'' cooling, and $\alpha\to\infty$ — thermally insulated conductor (solid curves). On comparing these results, we can easily see that even a coating of negligible thickness can substantially alter the conditions of thermal stabilization of well-cooled current-carrying elements. Of particular note is the fact that the coating thickness has a substantial influence on the range of currents for which the stability of the superconducting state is conserved. For example, for $\Delta_i=0$ and $\alpha=2$ the superconductor sustains irreversible damage, i.e., $V>0$, in the range of currents $0.7808=i_s<i<1$. However, when $\Delta_i=10^{-2}$ this range is almost doubled and when $\Delta_i=10^{-1}$, the thermal state of a rapidly cooled composite differs very little from adiabatic conditions.

Attention should also be drawn to the particular behavior of the normal zone velocity of well-cooled composite superconductors at near-critical currents which is observed as the coating thickness increases. The results of the calculations are plotted in Fig. 2. In this case, the curves $V(\Delta_i)$ have a maximum which, as can be seen from Fig. 2, is accompanied not only by a drop in the velocity of the normal zone but also by a shift of the maximum toward lower values of Δ_i in the presence of a coating with a higher heat capacity.

To sum up, when current-carrying elements fabricated using a composite superconductor with an external insulating coating are used in a real winding, particular attention must

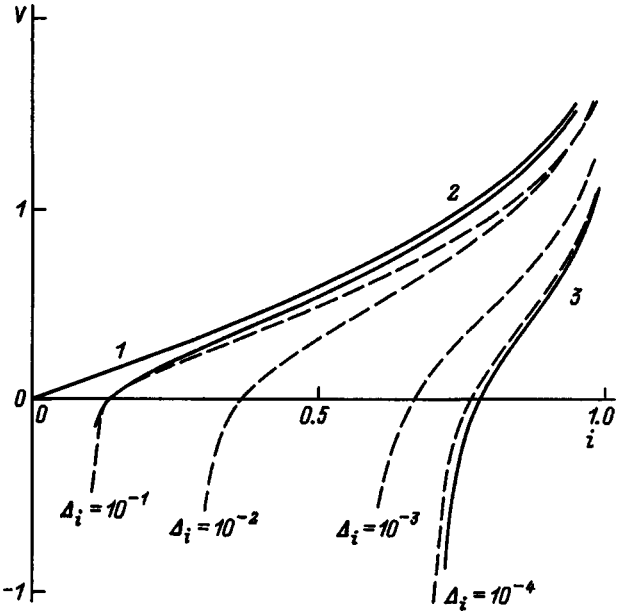


FIG. 1. Dependences of the normal zone velocity on current: solid curves — $\Delta_i=0$: *1* — $\alpha\to\infty$, *2* — $\alpha=100$, *3* — $\alpha=2$; dashed curves — $\Delta_i\neq0$: $\alpha=2$.

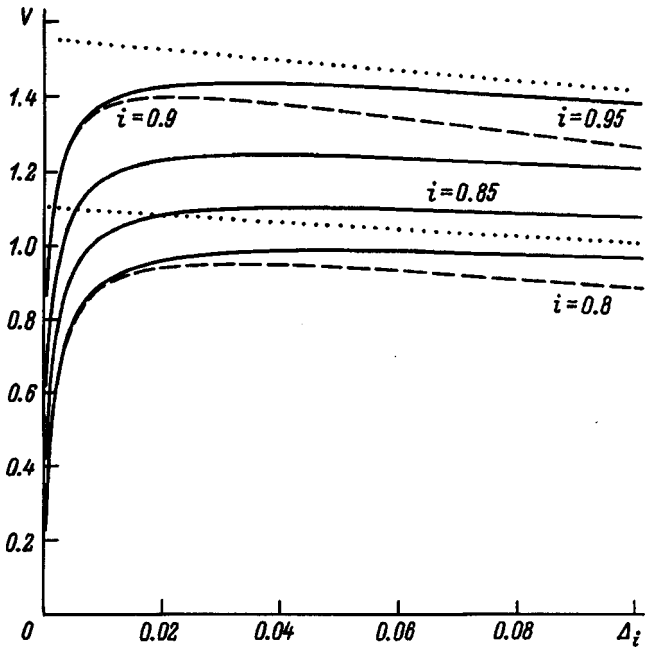


FIG. 2. Dependence of the normal zone velocity on the coating thickness for various currents: solid curves — $C=1$, $\alpha=2$, dotted curves — $C=2$, $\alpha=2$, and dashed curves — $C=1$, $\alpha\to\infty$.

be paid to the presence of a thermal barrier between them. The results discussed above indicate that the thermal resistance of the coating may almost completely negate the effectiveness of measures apparently designed to achieve intensive cooling of the superconducting winding. This must be taken into account when designing large magnet systems for which intensive cooling of the current-carrying elements is a prerequisite for stable performance.

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-16122a).

[1] V. A. Al'tov, V. B. Zenkevich, M. G. Kremlev, and V. V. Sychev, *Stabilization of Superconducting Magnet Systems* [in Russian], Énergoatomizdat, Moscow (1984), 312 pp.
[2] M. N. Wilson, *Superconducting Magnets* (Oxford University Press, London, 1983; Mir, Moscow, 1985, 407 pp.).
[3] A. V. Gurevich, A. A. Pukhov, and A. L. Rakhmanov, Preprint No. 4-285 [in Russian], Institute of High Temperatures, Academy of Sciences of the USSR, Moscow (1990), 27 pp.
[4] A. L. Rusinov, IEEE Trans. Magn. **30**, 2681 (1994).
[5] B. Truck, Cryogenics **20**(3), 146 (1980).

Translated by R. M. Durham

Experimental investigation of the dielectric strength of the H_{011} and E_{111} oscillation modes in a cylindrical resonator using compressed gas

V. A. Avgustinovich, S. N. Artemenko, and P. Yu. Chumerin

Research Institute of Nuclear Physics at Tomsk Polytechnical University, 634050 Tomsk, Russia
(Submitted May 25, 1996; resubmitted November 20, 1997)
Zh. Tekh. Fiz. **68**, 98–100 (December 1998)

It is shown experimentally that in order to increase the stored energy in storage resonators using compressed gas, it is preferable to have oscillation modes in which the electrical components of the fields do not interact with the inner surfaces of the resonator. © *1998 American Institute of Physics.* [S1063-7842(98)02012-1]

Problems of dielectric strength encountered in high-voltage devices using compressed gas can be attributed to sources of local amplification of the field by microinhomogeneities at the surface of the electrodes and impurity particles in the gas in a real discharge gap.[1,2] The influence of sources of local field amplification can be reduced by using different methods of mechanical and electric treatment of the electrode surface, gas purification techniques, conditioning of the electrodes and the dielectric medium. Electric fields may be excited in the discharge gaps of microwave sources which do not interact with the internal surfaces of the electrodynamic system. This method eliminates any influence of microinhomogeneities on the dielectric strength of the gas insulation.

The aim of the present paper is to demonstrate experimentally that the stored energy in a compressed-gas storage resonator may be increased by exciting oscillations for which the electric components of the fields do not interact with the internal surfaces. The investigations were carried out by comparing the dielectric strength of a resonator filled with nitrogen to pressures of $10^5-1.6\times10^5$ Pa when the H_{011} and E_{111} oscillation modes were alternately excited in it.

The apparatus is shown schematically in Fig. 1. We used a cylindrical copper standing-wave resonator *10* of diameter 4.38×10^{-2} m and length 3.49×10^{-2} m. Two types of modes were excited alternately in the resonator. The H_{011} mode was excited at a frequency of 9.36×10^9 Hz and the E_{111} mode was excited at 9.34×10^9 Hz. The required oscillation modes were excited through a circular coupling aperture positioned on one of the end covers of the resonator at the center of its radius, as shown in Fig. 1. The discharge volume was illuminated via a second similar aperture in the same cover, symmetric to the first about the center of the cover. The ultraviolet radiation source was an electric discharge in a beyond-limit circular waveguide connected coaxially to the illuminating aperture. The elements of the spark gap *11* were positioned at some distance from the illuminating aperture to avoid these having any influence on the Q factor and the operating frequency of the oscillations.

The degeneracy of the H_{011} and E_{111} oscillation modes was lifted by the exciting and illuminating apertures because the operating frequencies of these oscillations exhibited different sensitivity to the perturbing effect of these apertures.

The no-load Q factors of the H_{011} and E_{111} oscillation modes were 5750 and 11 700, respectively.

The resonator was excited using a tunable magnetron generator *1* with an output power up to 2×10^5 W and a pulse length of 10^{-6} s. The generator operated at a pulse repetition frequency of 50 Hz. The generator operating regime was matched to the resonant load by using a circulator *3* and a phase rotator *5*. The power reflected from the resonator coupling aperture was extracted to the load *4*. The power level entering the resonator was regulated by a variable attenuator *6*. Pressure decoupling between the resonator and the supply channel was provided by a radio-transparent, hermetically sealed window *8*. The resonator was filled with nitrogen using a gas admission system *9*. Before filling with gas the resonator was evacuated using a roughing pump.

In order to eliminate any spread of the breakdown field strengths, triggering of the spark gap was by a high-voltage modulator *12* synchronized with the beginning of the resonator excitation process. All parts of the system were synchronized using the unit *2*. The power levels of the incident and reflected waves were recorded experimentally using a directional coupler *7*, calibrated detector sections, and an S1-75 oscilloscope. The breakdown field in the resonator was determined by the breakdown power method.[3]

The nitrogen was passed through a gas filter before filling in order to separate the sources of local field amplification inside the resonator and at its inner surface. Figure 2 gives the breakdown field strength as a function of the nitrogen pressure obtained in the experiments for the H_{011} and E_{111} oscillation modes and also gives the results of calculations made using the electron balance equation from data given in Refs. 4 and 5 neglecting the influence of impurity particles and microinhomogeneities. The component of the electric field strength at the end cover of the resonator is indicated for the E_{111} oscillation mode. In the calculations it was assumed that the field in the resonator increases exponentially

$$E=[E_{\max}/\sqrt{2}\,][1-\exp(-t/\tau)],$$

where $E_{\max}$ is the maximum field in the resonator, τ is the resonator time constant, and t is the time.

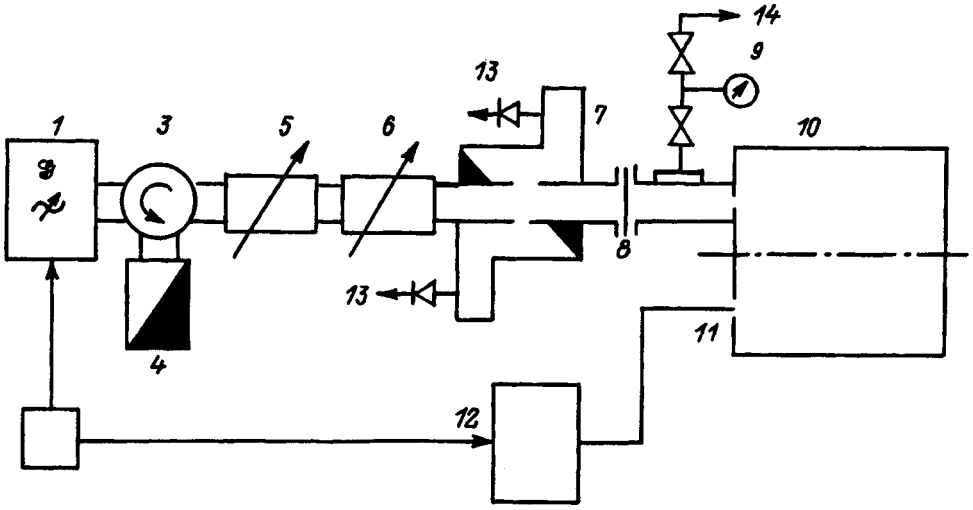


FIG. 1. Schematic of apparatus: *1* — microwave generator, *2* — synchronizing unit, *3* — circulator, *4* — absorbing load, *5* — phase rotator, *6* — attenuator, *7* — directional coupler, *8* — radio-transparent window, *9* — gas admission system, *10* — resonator, *11* — spark gap, *12* — high-voltage modulator, *13* — feed to oscilloscope, and *14* — line to gas cylinder.

The calculated dependences in the pressure range studied are linear. The slight difference between the breakdown fields for the H_{011} and E_{111} oscillation modes can be attributed to differences in the law governing the field rise in the resonator. When the resonator is filled with gas without preliminary purification, the levels of the breakdown field for the H_{011} and E_{111} oscillation modes differ slightly and from $\sim 10^6$ Pa onward, they cease to depend on pressure. Investigations of the influence of conditioning of the surface of the resonator walls and the dielectric medium carried out at 1.6×10^6 Pa show that after breakdown aging for 30 min, the breakdown field for the H_{011} and E_{111} oscillation modes was 3.5×10^7 and 2.5×10^7 V/m, respectively.

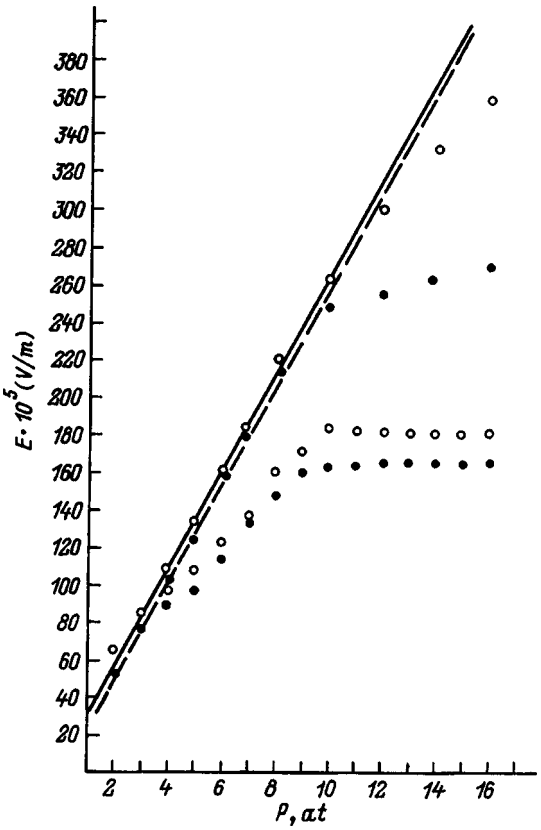


FIG. 2. Dependences of the breakdown field strength (E) on the nitrogen pressure (P): ○ — H_{011} oscillation mode, ● — E_{111} oscillation mode, solid curve — calculated dependence for H_{011}-oscillation mode, and dashed curve — for E_{111} oscillation mode.

In the experiments in which the gas was admitted via a filter, the breakdown field for the H_{011} oscillation mode becomes substantially greater than that for the E_{111} oscillation mode at pressures higher than 10^6 Pa and at 1.6×10^6 Pa this difference is 9×10^6 V/m.

The observed deviations of the experimental dependences of the breakdown field from linear are attributed to the influence of different sources of local field amplification. In the experiments in which the gas was admitted without preliminary purification, impurity particles inside the resonator were obviously the main source responsible for reducing the breakdown field. Thus, the dielectric strength of the H_{011} and E_{111} oscillation modes differs by an amount close to the calculated value. Conditioning the dielectric medium and purifying the gas reduced the influence of impurity particles inside the resonator on its dielectric strength. This increased the breakdown field for the H_{011} oscillation mode relative to that for the E_{111} mode.

To conclude, it has been shown experimentally that when specific technological operations are performed on the gaseous medium filling a storage resonator, oscillations modes for which the electric components of the rf field do not short onto the resonator walls are preferable for increasing the stored energy.

[1] V. Ya. Ushakov, *Isolation of High-Voltage Sources* [in Russian], Énergoatomizdat, Moscow (1994), 496 pp.
[2] D. V. Razevig and M. V. Sokolova, *Calculations of the Initial and Discharge Voltages of Gas Gaps* [in Russian], Énergiya, Moscow (1977), 200 pp.
[3] D. G. Raĭtsyn, *Dielectric Strength of Microwave Sources* [in Russian], Sovet-skoe Radio, Moscow (1977), 168 pp.
[4] A. D. MacDonald, *Microwave Breakdown in Gases* (Wiley, New York, 1966; Mir, Moscow, 1969, 206 pp.).
[5] Yu. A. Lupan, Zh. Tekh. Fiz. **46**, 2321 (1976) [Sov. Phys. Tech. Phys. **21**, 1370 (1976)].

Translated by R. M. Durham

Fabrication of piezo- and pyrosensitive films of polyvinylidene fluoride by vacuum evaporation in an electric field

A. I. Vaĭtenkov and O. E. Kovalenko

Institute of Applied Optics, Academy of Sciences of Belarus, 212793 Mogilev, Belarus
(Submitted August 6, 1997)
Zh. Tekh. Fiz. **68**, 101–103 (December 1998)

Films of polyvinylidene fluoride with tetrafluoroethylene several micron thick were prepared by vacuum electron-beam evaporation in a static electric field. It is shown that a polymer deposited on a negative electrode possesses a structure similar to that of uniaxially stretched and polarized films and has a high (up to 85%) content of the piezoactive β phase. The electromechanical coupling coefficient k_{33} and the pyrosensitivity of these samples are substantially higher than those of layers deposited under the same conditions in the absence of a field. © *1998 American Institute of Physics.* [S1063-7842(98)02112-6]

Continuing interest is being shown in studies of films of polyvinylidene fluoride (PVDF) and its copolymers with trifluoroethylene or tetrafluoroethylene (TFE) because of their good piezo- and pyroelectric characteristics.[1,2] Free PVDF films become piezosensitive after uniaxial stretching and polarization in a corona discharge.[1] In some cases, films on conducting substrates pulled from solution become polarized, but samples with inferior and unstable parameters are usually obtained.[3] It was shown in Refs. 4–6 that thin piezo- and pyrosensitive layers of PVDF may be produced by laser or thermal evaporation of the polymer in vacuum, followed by polarization. Promising results were also obtained by thermal deposition of films in the presence of an electric field.[7] However, because of the low rate of deposition, only very thin layers of active polymer were obtained which were barely suitable for studying the structure and physical properties.

Here we demonstrate the possibility of forming relatively thick piezoactive PVDF/TFE films by vacuum electron-beam evaporation in an electric field. The initial material was an F2-MÉ copolymer of polyvinylidene fluoride with 5% tetrafluoroethylene. A 0.5 mm thick polymer strip prepared by hot pressing was mounted on a movable water-cooled copper plate using a tungsten grid. This strip was evaporated by an electron beam scanned at a frequency of 1 Hz with a current density of around 5 mA/cm^2 and giving a partial pressure of the polymer dissociation products of $(5-7)\times 10^{-2}$ Pa. At lower beam currents the fraction condensing on the substrate was negligible and under harsher conditions the deposited product exhibited clear signs of destruction. The films were deposited on silicon and glass substrates at 300 K at a rate of 4–10 nm/s, which is an order of magnitude faster than in Ref. 7. This was achieved by combining pulsed surface heating of the target with partial ionization of the evaporation products. The positive role of the latter in increasing the rate of deposition of polymer films was noted in Refs. 8 and 9. A static electric field longitudinal or transverse to the vapor flux was created in the polymer condensation zone (Fig. 1). The distance between the electrodes was 25 mm and the supplied voltage was ± 7 kV relative to ground. The interelectrode current flowing during the evaporation process was a few mA/cm^2.

The crystal structure of the films was investigated by infrared spectroscopy using a UR-20 spectrometer. The transmission spectra of PVDF/TFE layers between 2 and 3.5 μm thick deposited under different conditions are shown in Fig. 2. For any electrode configuration films with signs of severe destruction and broadened absorption bands typical of molten PVDF (Ref. 10) were deposited on the anode. The agreement between the bands in the spectra of samples deposited without a field, on a cathode, and pulled from a dimethylacetamide solution, and also their agreement with the published data[10,11] serves to confirm that the chemical composition of the films and the initial material is identical. Only the weak bands at 1600–1620 cm^{-1} may indicate the presence of conjugated C—C bonds in the polymer evaporated in vacuum. Its refractive index n_s for *s*-polarized light ($\lambda = 633$ nm), measured simultaneously with the layer thickness using a waveguide refractometer, is 1.425–1.430 and also shows good agreement with the published data.[12] The refractive index of the films deposited on the anode was not less than 1.50.

The spectra in Fig. 2 indicate that the α- and β-phases of the polymer are present in the films in different proportions depending on the deposition conditions. The 535, 612, 765, and 805 cm^{-1} absorption bands are typical of the α-phase, while the 510, 840, and 1273 cm^{-1} bands are typical of the piezo- and pyroactive β-phase.[10,11] The content of the β phase in films deposited without a field, calculated from a comparison of the intensities of the 510 and 535 cm^{-1} absorption bands,[13] did not exceed 25%. For samples deposited on a cathode in a longitudinal or transverse field, this fraction was between 70 and 85%. The result was improved further by applying to the grid *6* a positive potential relative to the target of approximately the same order of magnitude. In this case, the grid clearly shields the growing film more effectively from bombardment by scattered electrons and also reduces the energy of the CF_3^+, $C_3F_5^+$, and other ions depos-

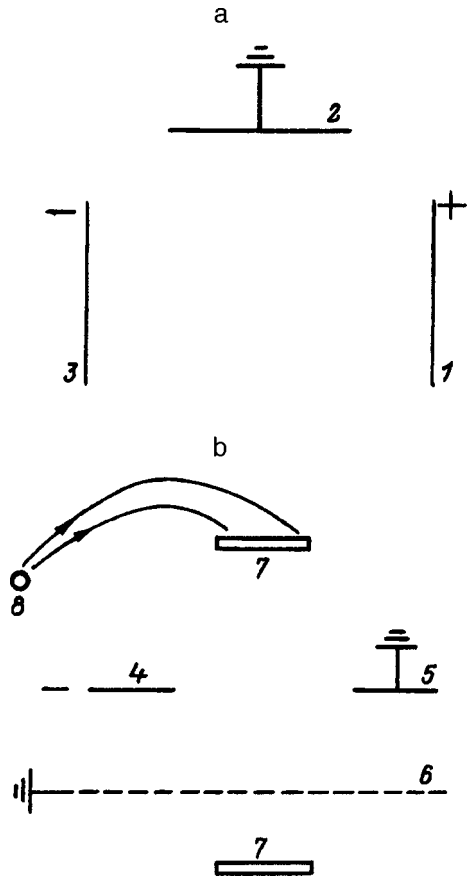


FIG. 1. Schematic of electrode configuration for the depositing films: *1–6* — electrodes, *7* — target, and *8* — electron gun.

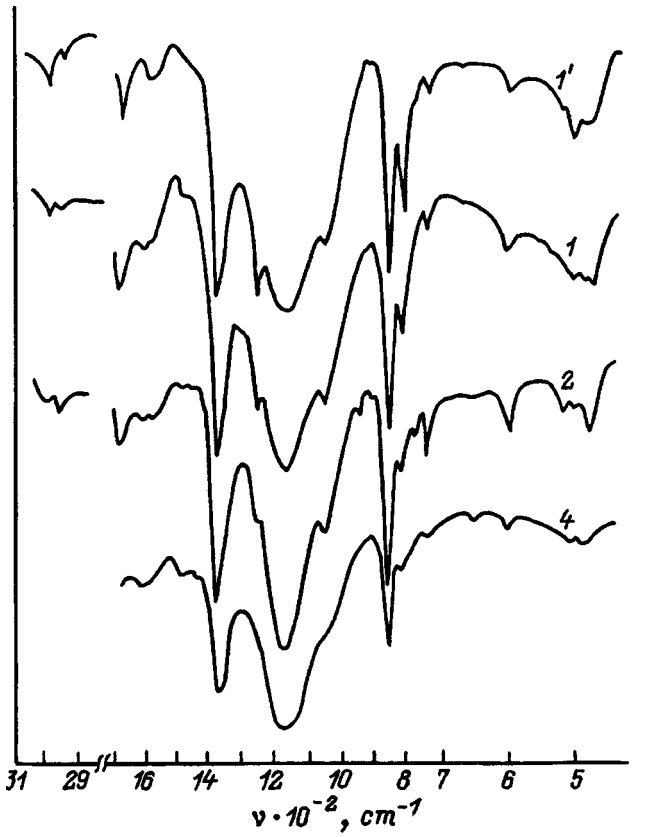


FIG. 2. Transmission spectra of deposited PVDF/TFE films. The numbers on the curves correspond to the numbering of the electrodes on which the Si substrates were mounted. Spectrum 1′ was recorded in *p*-polarized light at an angle of incidence of 45°.

ited on the substrate, which are present in the PVDF destruction products.[8] The weakly defined 668 cm^{-1} absorption band in spectra *1* and *2* indicates that the number of ''head to head'' and ''tail to tail'' structural defects in PVDF/TFE is lower than that obtained with a transverse field. However, bearing in mind the higher intensity of the 840 cm^{-1} band compared with the 1273 cm^{-1} band and on the basis of data given in Refs. 11 and 13, we can conclude that the crystalline phase of the polymer contains a relatively small fraction of long *trans* sequences with $m \geqslant 4$.

When the light is perpendicularly incident, our films exhibit nondichroic spectra. When the film is positioned obliquely in *p*-polarized light, the intensity of the 510, 840, 1273, 3020, and 2980 cm^{-1} bands assigned to different vibration modes of the CF_2 and CH_2 bonds, increases substantially in relation to the weakly dichroic 880 cm^{-1} band. At the same time the intensity of 1080 cm^{-1} band decreases (antisymmetric C—C vibrations). The anisotropy of the refractive index $\Delta n = n_s - n_p$ for films deposited in a field is 0.015–0.020, whereas without a field, it is an order of magnitude lower. The limiting value of Δn for uniaxially oriented PVDF/TFE films polarized by the usual methods is 0.4 (Ref. 14).

These results indicate that the molecular chains of a polymer deposited in a field typically lie in the substrate plane while the CF_2 and CH_2 dipoles are preferentially oriented perpendicular to the substrate. This is observed for both longitudinal and transverse electrode configurations. Uniaxially stretched and polarized PVDF films have the same conformational characteristics.[1,11,13] Thus, the processes of orientation and polarization of the polymer are combined in the electron-beam deposition of PVDF/TFE films in an electric field and this creates conditions for the appearance of piezo- and pyroelectric properties.

The electromechanical coupling coefficient k_{33} of the films was measured using a piezotransformer method[1] whereby a sandwich of two deposited polymer layers of approximately the same thickness, separated by a thin aluminum electrode 12 mm in diameter, is mounted on a glass substrate. These electrodes were placed on two different surfaces of the structure. The measurements were made at a frequency of 1 kHz several days after depositing the films when their characteristics had essentially stabilized. The permittivity of the films at this frequency was 9.8. The value of k_{33} for films deposited without a field did not exceed 0.025, whereas for films deposited in a longitudinal field, it fluctuated between 0.11 and 0.14, i.e., approached the value for industrial polarized PVDF/TFE samples.[1,2]

The pyroresponse of the films was measured by a photodynamic method. A sample with 0.05 μm thick aluminum films deposited on opposite sides was irradiated by $\lambda = 633$ nm rectangular laser pulses of 40 ms duration. The percent modulation of the temperature was estimated as ap-

proximately 2 K. The variable component of the pyrosignal was measured using a phase-sensitive voltmeter with an input resistance of 1 GΩ. The response of films $1-3$ μm thick deposited without a field did not exceed 0.2 mV under these conditions and was probably caused by the presence of electret properties. The signal from films prepared in longitudinal and transverse fields was substantially higher, 4–8 mV. These data can be used to obtain an approximate estimate of the pyroelectric coefficient. This gives 0.5 μC/cm$^2\cdot$K, which is an order of magnitude lower than that for standard PVDF samples.[1]

To conclude, the physical characteristics of these polymer layers are comparable with those of films of this material prepared by the usual methods. This method is fully compatible with microelectronics technology. Thus, in addition to the conventional ranges of application, these piezo- and pyrosensitive polymer structures deposited directly on a substrate may also of interest as active elements in integrated optics and fast matrix infrared detectors.

The authors are grateful to É. P. Kalutska for measuring the infrared spectra of the films.

[1] G. A. Lushcheĭkin, *Polymer Piezoelectrics* [in Russian], Khimiya, Moscow (1990), 176 pp.
[2] R. Gerchard-Multhaupt, Ferroelectrics , **75**, 385 (1987).
[3] T. Furukawa, J. X. Wen, K. Suzuki *et al.*, J. Appl. Phys. **56**, 829 (1994).
[4] A. W. Stephens, A. W. Levine, J. Fech *et al.*, Thin Solid Films **24**, 361 (1974).
[5] K. Maki, H. Terashima, and K. Kikuma, Jpn. J. Appl. Phys. **29**, 991 (1990).
[6] A. Takeno, N. Okui, T. Kitoh *et al.*, Thin Solid Films **202**, 205 (1991).
[7] Y. Yoshida, T. Horiuchi, and K. Matsushige, Jpn. J. Appl. Phys. **32**, 1248 (1993).
[8] K. P. Gritsenko, Ukr. Khim. Zh. **57**, 782 (1991).
[9] A. M. Krasovskiĭ and E. M. Tolstopyatov, *Preparation of Thin Films by Deposition of Polymers in Vacuum* [in Russian], Nauka i Tekhnika, Minsk (1989), 181 pp.
[10] M. Kobayashi, K. Tashiro, and H. Tadokoro, Macromolecules **8**, 158 (1975).
[11] K. Tashiro, M. Kobayashi, and H. Tadokoro, Macromolecules **14**, 1757 (1981).
[12] T. Yamada and T. Kurokawa, Appl. Phys. Lett. **40**, 364 (1982).
[13] V. V. Kochervinskiĭ, V. A. Glukhov, V. G. Sokolov *et al.*, Vysokomol. Soedin., Ser. A **31**, 2590 (1989).
[14] V. V. Kochervinskiĭ, V. F. Romadin, V. A. Glukhov *et al.*, Vysokomol. Soedin., Ser. A **31**, 1382 (1989).

Translated by R. M. Durham

Electrostatic problem for a planar eccentric ring

T. V. Denisova and V. S. Protsenko

Kharkov Aviation Institute, 310070 Kharkov, Ukraine
(Submitted November 11, 1996; resubmitted August 4, 1997)
Zh. Tekh. Fiz. **68**, 104–106 (December 1998)

The electrostatics problem for a concentric annular disk has been studied fairly comprehensively (see Ref. 1 and the literature cited). An eccentric ring is considered for the first time using an assumption known in the literature as the Kelvin transformation.[2] This transformation allowed the initial problem to be reduced to that for a concentric ring with a point charge lying outside the ring in the same plane. This problem was solved using triple integral equations. An approximate solution for the charge density distribution at an eccentric annular disk is given as an expansion in terms of a small parameter.
[S1063-7842(98)02212-0]

1. For a disk charged to the potential u_0, the boundary-value problem is formulated as finding the harmonic function u which decreases to zero at infinity in a space having the permittivity ε^*, if its value at the disk u_0 is known.

Let us assume that a disk is situated in the $z=0$ plane in a Cartesian coordinate system (see Fig. 1a). On account of the symmetry of the field relative to the plane $z=0$, we shall have the additional condition $u'_z(x,y,0)=0$ outside the disk. We shall assume that an eccentric disk occupies the region (S)

$$\{x^2+y^2<R^2, \quad (x+h)^2+y^2>R_1^2\}, \tag{1.1}$$

where $h>0$, $R_1<R$, $h<R$, $R>R_1+h$.

The following lemma, which we shall put forward without proof because of its simplicity, holds.

Whatever the parameters of an eccentric ring (S) satisfying the system of inequalities (1.1), the equation

$$k^2h+k(R_1^2-h^2-R^2)+hR^2=0 \tag{1.2}$$

invariably has real positive roots k_1 and k_2 such that the inequalities $k_1>R$, $h<k_2<h+R_1$, and $k_1>k_2$ are satisfied.

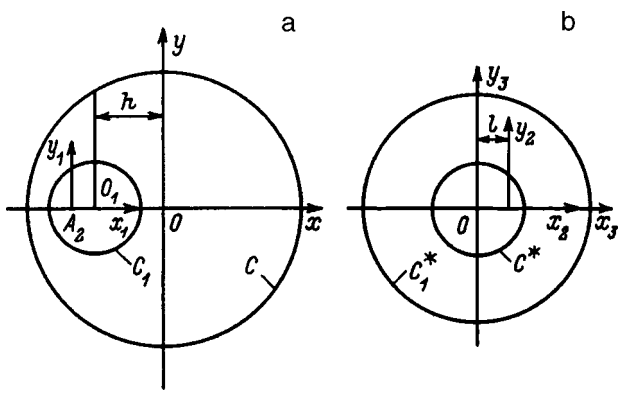


FIG. 1.

Geometrically this implies that the points $A_1(-k_1,0)$ and $A_2(-k_2,0)$ lie outside the ring (S), where the point A_1 lies to the left of the major circumference C and the point A_2 lies inside the minor circumference C_1.

Theorem. If the point $A_p(p=1,2)$ is related to a new Cartesian coordinate system (x_1,y_1), placing the origin at this point and directing the axis of this system in the direction of the axes of the (x,y) system, the inversion transformation with the center at point A_p

$$x_1=x_2/\rho_2^2, \quad y_1=y_2/\rho_2^2, \quad \rho_2^2=x_2^2+y_2^2,$$

$$y=y_1, \quad x_1=x+k_p$$

converts the eccentric ring S into a concentric ring S_1 on the plane (x_2,y_2) (see Fig. 1b).

To be specific, we shall take the point A_2 as the center of inversion. The ring (S_1) will then be determined by the inequalities

$$(x_2+l)^2+y_2^2>a_1^2, \quad (x_2+l)^2+y_2^2<a^2,$$

$$l=k_2(R^2-k_2^2)^{-1}, \quad a_1=R(R^2-k_2^2)^{-1},$$

$$a=R_1[R_1^2-(k_2-h)^2]^{-1}. \tag{1.3}$$

With this choice of center of inversion, it is convenient to go from the limit of an eccentric ring to a concentric one. If the point A_1 is taken as the center of inversion, it will be convenient to go from the limit of a concentric ring to a circular disk.

We shall apply a Kelvin transformation[2] with the center of inversion at point A_2 to the initial boundary-value problem for an eccentric ring

$$x_1=x_2/r_2^2, \quad y_1=y_2/r_2^2, \quad z_1=z_2/r_2^2,$$

$$r_2^2=x_2^2+y_2^2+z_2^2. \tag{1.4}$$

The function $\psi(x_2,y_2,z_2)=r_2^{-1}u(x_1,y_1,z_1)$ will then satisfy the Laplace equation with the following boundary conditions in the half-space $z_2<0$

$$\psi(x_2,y_2,0)u_0\rho_2^{-1} \quad (x_2,y_2)\in(S_1), \tag{1.5}$$

$$\psi'_{z_2}(x_2,y_2,0)=0 \quad (x_2,y_2)\notin(S_1), \tag{1.6}$$

$$\psi(x_2,y_2,z_2)\to 0 \quad \text{for } r_2\to\infty. \tag{1.7}$$

2. Using the shear transformation $x_3=l+x_2$, $y_3=y_2$, $z_3=z_2$ the center of the ring (S_1) moves to the origin of the system (x_3,y_3,z_3) (see Fig. 1b). The boundary conditions (1.5) and (1.6) are transformed to give the conditions

$$\psi(x_3,y_3,0)=u_0[(x_3-l)^2+y_3^2]^{-1/2}, \quad a_1<\rho_3<a, \tag{2.1}$$

$$\psi'_{z_3}(x_3,y_3,0)=0, \quad 0\leqslant\rho_3<a_1, \quad \rho_3>a. \tag{2.2}$$

For simplicity we shall omit the indices of the variables x, y, and z for the time being.

We shall solve the problem for a concentric ring using the method proposed in Ref. 3. We express the function $\psi(x,y,z)$ in a cylindrical coordinate system as a Fourier–Bessel expansion

$$\psi=\sum_{m=0}^{\infty}\cos m\varphi\int_0^{\infty}A_m(\lambda)\lambda^{-1}e^{\lambda z}J_m(\lambda\rho)d\lambda \quad (z<0). \tag{2.3}$$

We write the boundary function (2.1) in the form of a similar expansion

$$\frac{1}{\sqrt{(x-l)^2+y^2}}=\frac{1}{\sqrt{\rho^2+l^2-2l\rho\cos\varphi}}$$

$$=\sum_{m=0}^{\infty}\delta_m\cos m\varphi\int_0^{\infty}J_m(\lambda\rho)J_m(\lambda l)d\lambda,$$

$$\delta_0=1, \quad \delta_m=2, \quad m\geqslant 1, \quad l<a_1<a, \tag{2.4}$$

where (ρ,φ) are polar coordinates with the origin at the point O^* (see Fig. 1b).

From the boundary conditions (2.1) and (2.2) using the expansions (2.3) and (2.4) we obtain a system of triple equations to determine the function $A_m(\lambda)$

$$\int_0^{\infty}A_m(\lambda)\lambda^{-1}J_m(\lambda\rho)d\lambda=\delta_m\int_0^{\infty}J_m(\lambda\rho)J_m(\lambda l)d\lambda$$

$$\times(a_1<\rho<a),$$

$$\int_0^{\infty}A_m(\lambda)J_m(\lambda\rho)d\lambda=0 \quad (0\leqslant\rho<a_1, \quad \rho>a). \tag{2.5}$$

Equations (2.5) in accordance with the method described in Ref. 4 are reduced to a system of integral equations for the auxiliary functions $f_m^{(k)}(t)(k=1,2)$

$$\rho^{1-m}f_m^{(1)}(\rho)=\frac{2}{\pi}\rho(a^2-\rho^2)^{-1/2}\int_a^{\infty}\frac{t^{1-m}(t^2-a^2)^{1/2}}{t^2-\rho^2}f_m^{(2)}$$

$$\times(t)\,dt\,(0\leqslant\rho<a),$$

$$\rho^{1+m}f_m^{(2)}(\rho)=\frac{2}{\pi}\rho(\rho^2-a_1^2)^{-1/2}\int_0^{a_1}\frac{t^{1+m}(a_1^2-t^2)^{1/2}}{t^2-\rho^2}f_m^{(1)}$$

$$\times(t)\,dt+g_m(\rho) \quad (\rho>a),$$

$$g_m(\rho)=\frac{2}{\pi}l^m\delta_m\frac{\rho\sqrt{a_1^2-l^2}u_0}{(\rho^2-l^2)\sqrt{\rho^2-a_1^2}}. \tag{2.6}$$

Note that the function of interest to us

$$\sigma_m(\rho)=\int_0^{\infty}A_m(\lambda)J_m(\lambda\rho)d\lambda \quad (a_1<\rho<a),$$

in terms of which the density of the charge distribution at the ring is expressed, is given by

$$\sigma_m(\rho)=f_m^{(1)}(\rho)+f_m^{(2)}(\rho) \quad (a_1<\rho<a). \tag{2.7}$$

In Ref. 3 the system (2.6) is replaced by a set of two infinite systems of algebraic equations. For this we set

$$f_m^{(1)}(\rho)=\frac{2}{\pi}\varepsilon\frac{\rho^m a^{-(m+1)}}{\sqrt{a^2-\rho^2}}\sum_{k=0}^{\infty}a_{2k}^{(m)}\left(\frac{\rho}{a}\right)^{2k},$$

$$f_m^{(2)}(\rho)=\frac{2}{\pi}\varepsilon\frac{\rho^{-m}a_1^{m-1}}{\sqrt{\rho^2-a_1^2}}\sum_{k=0}^{\infty}b_{2k}^{(m)}\left(\frac{a_1}{\rho}\right)^{2k+2},$$

$$\varepsilon=a_1a^{-1}. \tag{2.8}$$

As a result of expanding the functions $(t^2-\rho^2)^{-1}$ and $(\rho^2-l^2)^{-1}$ as power series, calculating the integrals, and equating the coefficients with the same powers of ρ, we obtain

$$a_{2k}^{(m)}=\sum_{p=0}^{\infty}b_{2p}^{(m)}\varepsilon^{2p+m+1}M_p^k(m,\varepsilon^2), \quad k\geqslant 0,$$

$$b_{2k}^{(m)}=\gamma_{2k}^{(m)}+\sum_{p=0}^{\infty}a_{2p}^{(m)}\varepsilon^{2p+m+2}M_p^k(m,\varepsilon^2),$$

$$M_p^k(m,\varepsilon^2)=\frac{(m+k+p)!}{2\sqrt{\pi}\Gamma(m+k+p+5/2)}$$

$$\times F(1/2,m+k+p+1;m+k+p+5/2;\varepsilon^2),$$

$$\gamma_{2k}^{(m)}=u_0\delta_m\frac{\varepsilon_1}{\varepsilon}\varepsilon_2^{2k+m}, \quad \varepsilon_2=la_1^{-1}, \quad \varepsilon_1=(1-\varepsilon_2^2)^{1/2}. \tag{2.9}$$

3. Estimates made in Ref. 3 indicate that the system (2.9) is quasiregular for $\varepsilon<1$ and regular for $\varepsilon\leqslant 0.75$. These estimates also indicate that the solution of system (2.9) belongs to the number space of sequences of converging squares l_2 and the following inequalities are found

$$|a_{2k}^{(m)}|\leqslant KD_k^{(m)}, \quad |b_{2k}^{(m)}|\leqslant\gamma_{2k}^{(m)}+KD_k^{(m)}\varepsilon,$$

$$D_k^{(m)}=\varepsilon^{m+1}(m+k)!\,\Gamma^{-1}(m+k+5/2),$$

$$K=\text{const}>0. \tag{3.1}$$

The inequalities (3.1) show that the series (2.8) and the series $\Sigma_{m=0}^{\infty}\sigma_m(\rho)\cos m\varphi$ converge absolutely and uniformly when $a_1\leqslant\rho\leqslant a$. Returning systematically from the variables x_3,y_3,z_3 to x,y,z, we write the expression for the density of the charge distribution at an eccentric annular disk

$$\sigma(\rho,\varphi)=\frac{\varepsilon^*}{2\pi\rho^{*3}}\sum_{m=0}^{\infty}\sigma_m(\rho_3)\cos m\varphi_3,$$

$$\rho^*=\sqrt{(x+k_2)^2+y^2},\quad \rho_3^2=x_3^2+y_3^2,$$

$$x_3=\frac{x+k_2}{(x+k_2)^2+y^2}+l,\quad y_3=\frac{y}{(x+k_2)^2+y^2},$$

$$\cos\varphi_3=\frac{x_3}{\rho_3},\quad \sin\varphi_3=\frac{y_3}{\rho_3}.$$

Finally we give the solution of infinite systems as the first terms of an expansion in powers of the small parameter ε

$$a_{2k}^{(m)}=\frac{\varepsilon^{m+1}(m+k)!}{2\sqrt{\pi}\Gamma(m+k+5/2)}\times\left[\gamma_0^{(m)}+\varepsilon\frac{m+k+1}{(m+k+5/2)}\left(\frac{1}{2}\gamma_0^{(m)}+\gamma_2^{(m)}\right)+O(\varepsilon^4)\right],$$

$$b_{2k}^{(m)}=\gamma_{2k}^{(m)}+\varepsilon^{2m+3}\times\left[\frac{(m+k)!m!}{4\pi\Gamma(m+k+5/2)\Gamma(m+5/2)}\gamma_0^{(m)}+O(\varepsilon^2)\right].$$

Summing the series, we obtain

$$f_m^{(1)}(\rho)=\frac{\gamma_m\varepsilon^{m+1}}{\pi^{3/2}a\sqrt{a^2-\rho^2}}x_0^m\times\left[\psi_m(x_0^2)+\varepsilon^2\left(\frac{1}{2}+\varepsilon_2^2\right)\psi_{m+1}(x_0^2)+O(\varepsilon^4)\right],$$

$$f_m^{(2)}(\rho)=\frac{2\gamma_m}{\pi a_1\sqrt{\rho^2-a_1^2}}y_0^{m+2}\left[\frac{\rho^2}{\rho^2-l^2}+\varepsilon^{2m+3}\times\frac{m!}{4\pi\Gamma(m+5/2)}\psi_m(y_0^2)+O(\varepsilon^{2m+7})\right],$$

$$x_0=\rho/a,\quad y_0=a_1/\rho,\quad \gamma_m=u_0\delta_m\varepsilon_2^m\sqrt{1-\varepsilon_2^2},$$

$$\psi_m(x)=\frac{m!}{\Gamma(m+5/2)}F(1,m+1;m+5/2;x). \tag{3.2}$$

Note that the proposed method can be used to solve the problem of the electric charge distribution at an annular concentric disk in the presence of concentrated charges distributed on the plane $z=0$ outside the disk.

[1] G. A. Grinberg and V. N. Kuritsyn, Zh. Tekh. Fiz. **31**, 1017 (1961) [Sov. Phys. Tech. Phys. **6**, 743 (1961)].
[2] A. N. Tikhonov and A. A. Samarskiĭ, *Equations of Mathematical Physics* (Pergamon Press, Oxford, 1964; Nauka, Moscow, 1972, 735 pp.).
[3] V. S. Protsenko, Prikl. Mekh. **4**(9), 83 (1968).
[4] J. C. Cooke, Proc. Edinburgh Math. Soc. **13**, 303 (1963)

Translated by R. M. Durham