Energy balance of a pulsed emitting sodium-vapor/xenon discharge

F. G. Baksht and V. F. Lapshin

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

1. A sodium-vapor discharge is an efficient source of visible light.[1] In recent years there have been active attempts to create mercury-free low-power high-pressure sodium lamps on its basis.[2–4] Most of this research, however, focused on steady-state discharges. At the same time, a pulsed burning regime makes it possible to achieve considerable improvement in the output characteristics of the lamp.[1–3] The study of this arc burning regime is plagued by great difficulties, which are caused by the highly inhomogeneous state of the plasma and by the great variety and nonstationary character of the processes occurring in it. It was shown in Ref. 5 that a pulsed discharge is characterized by a complex pattern of gas-dynamic flows of components, whose relative concentrations vary across the radius. This paper presents the results of calculations of the energy balance in the column of a pulsed discharge of low power $W = 5 - 60$ W/cm. It is shown that the fraction of energy emitted in the visible range reaches 50–55% of the power W released in a unit length of the discharge column, which is significantly greater than in a steady-state high-pressure discharge.[1]

2. A pulsed discharge in a long ($L \gg R$) cylindrical tube with an internal radius $R = 1.5$ mm is considered in the calculations. It is assumed that the tube is filled with xenon under a pressure of 20 Torr at 293 K. The quantity of sodium is determined from the saturating pressure P_{sat} in the coldest part of the tube with a temperature T_{cold}. It is assumed that $T_{cold} = T_w - 300$ K, where T_w is the temperature of the inner surface of the tube wall in the working zone of the torch. A steady-state lamp operating regime, in which a current pulse of assigned shape $I(t)$ is passed through the plasma of each successive discharge is considered below. The pulse repetition rate $\nu = 800$ Hz, and the pulse duration $t_{pul} = 0.1/\nu$.

The system of gas-dynamic equations for the individual plasma components, the boundary conditions on it, and the procedure for calculating the electrical and thermal conductivity of a plasma were described in Ref. 5. The calculation of the radiation losses from a unit volume of the plasma took into account the emissions in the lines corresponding to transitions to the ground state from the $3P$, $4P$, and $5P$ levels, transitions to the $3P$ state from the $4S$, $5S$, $6S$, $3D$, $4D$, and $5D$ levels, and to the $4S$ state from the $5P$ and $6P$ levels, as well as the emission from the recombination and electron-ion bremsstrahlung continua. The emission in all the lines was calculated with consideration of reabsorption in the effective lifetime approximation for a homogeneous plasma.[6] When the absorption coefficient was calculated, the profiles of all the lines were assumed to be Lorentzian. Two broadening mechanisms were taken into account for the $3P \to 3S$ line: by atoms upon resonant excitation transfer[7] and by Stark

electrons.[8] Only the broadening by electrons was taken into account for all the remaining lines. The broadening of resonance lines due to the formation of NaXe* excimer molecules and Na$_2$ molecules was disregarded in view of their small concentration (for further information on this subject see Refs. 9 and 10). The recombination continuum for electron capture in state k was calculated using the photoionization cross section σ_k^{ph} of the respective level:

$$W_{rec}^{(k)} = \sqrt{\frac{2}{\pi m k_B T}} \frac{1}{mc^2} n_i n_e \int_0^\infty e^{-x} (\varepsilon + E_{ion}^{(K)})^3 g_k \sigma_k^{ph} dx.$$

Here m is the electron mass, $n_e = n_i$ is the plasma concentration, g_k and $E_{ion}^{(k)}$ are the statistical weight and the ionization energy of an Na atom in state k, and $\varepsilon = x k_B T$ is the energy of the recombining electron. The photoionization cross sections of the nS, nP ($n \geq 3$), and $3D$ states were borrowed from Refs. 11–14, and the values of the cross sections in the quasiclassical approximation for hydrogenic atoms[15] were employed for the nD, nF ($n \geq 4$), and nG ($n \geq 5$) states.[15] In calculating the radiation losses it was taken into account that the discrete spectrum contains only excited states of Na atoms having an effective principal quantum number $n_l < n_{max}$, where n_{max} was determined using the Inglis-Teller formula $n_{max} = 0.5(a_0^3 n_e)^{-2/15}$, where a_0 is the Bohr radius.

3. Figures 1 and 2 present the results of calculations of the principal parameters of the discharge plasma for a power $W = 34$ W/cm. The shape of the current pulse was selected such that rapid heating of the plasma of each successive discharge to high temperatures would occur within a time $t/t_{pul} \sim 0.1$ and the temperature of as large as possible a part of the plasma would subsequently be identical across the radius of the tube and would not vary during the pulse (Figs. 1 and 2a). At the end of the pulse, the hot emitting plasma occupies most of the tube and is separated from its walls by considerably denser and cooler gas (Fig. 2b). The corresponding pulse shape includes a portion $0 \leq t/t_{pul} \leq 0.1$ with a rapid rise in the current from $I(0) = 0.1$ A in each successive discharge to $I(0.1) = 5$ A and a portion $0.1 \leq t/t_{pul} \leq 1.0$ with a subsequent smoother increase in the current to $I_{max} = 20$ A. We note that the strength of the longitudinal electric field E has a sharp maximum ($E_{max} = 129$ V/cm) at the very beginning of the pulse, which is characteristic of a pulsed regime. The total pressure P varies relatively smoothly during the pulse ($P_{max} = 863$ Torr). Figure 3a presents calculations of the fractions of the power supplied to the plasma that are contributed to line emission in the visible (η_{vis}) and nonvisible (η_{nonvis}) spectra, to the recombination continuum (η_{rec}), and to thermal radiation of the tube walls (η_w). The quantity

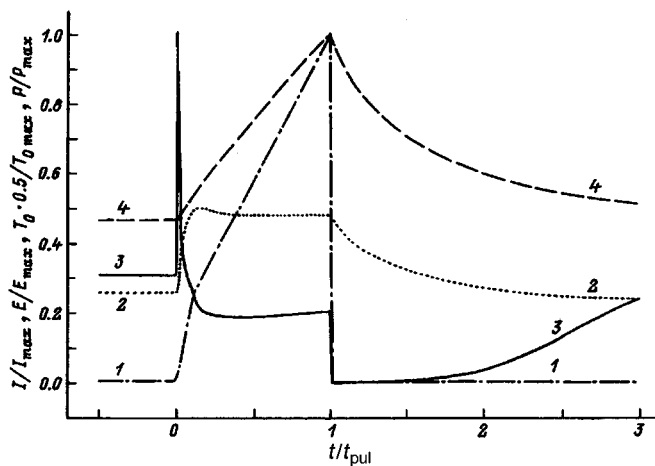FIG. 1. Time dependence of the principal plasma parameters. $1 - I/I_{max}$, $2 - T_0 \cdot 0.5/T_{0max}$, $3 - E/E_{max}$, $4 - P/P_{max}$; T_0 — temperature on the discharge axis, $T_{0max}=5900$ K.

η_{vis} includes emission corresponding to the $5D \rightarrow 3P$ (498 nm), $6S \rightarrow 3P$ (515 nm), $4D \rightarrow 3P$ (569 nm), $3P \rightarrow 3S$ (589 nm), and $5S \rightarrow 3P$ (616 nm) transitions. Since the part of the recombination continuum corresponding to the capture of electrons in $3D$ and $4P$ states lies in the visible part of the spectrum, the energy efficiency of the lamp in the visible part
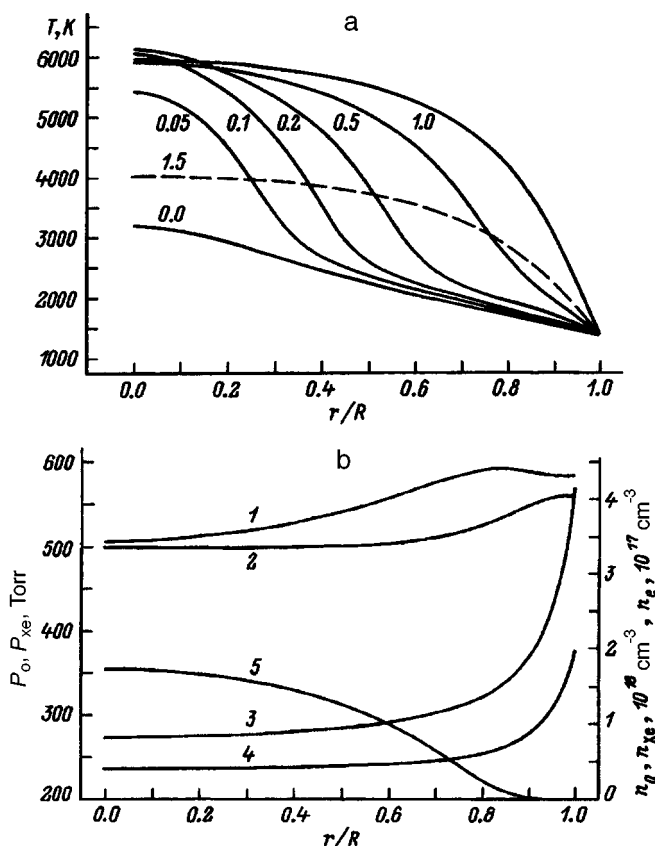

FIG. 2. Radial distributions of the plasma parameters: a — temperature at various moments in time from the beginning of a pulse (the time in units of t/t_{pul} is indicated on the curves); b — the concentrations n_e and $n_0=n_i+n_{Na}$, the pressure $P_0=P_i+P_{Na}$, and $P_{Xe}\cdot 2$ at the time $t=t_{pul}$: $1 - P_0$, $2 - P_{Xe}$, $3 - n_0$, $4 - n_{Xe}$, $5 - n_e$.
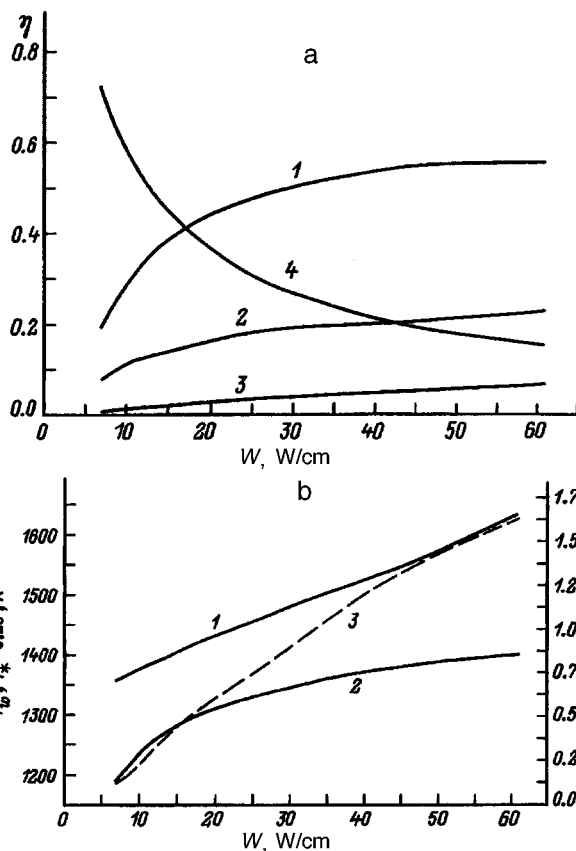

FIG. 3. Dependence of the energy balance (a) and the discharge parameters (b) on the power W. a: $1 - \eta_{vis}$, $2 - \eta_{nonvis}$, $3 - \eta_{rec}$, $4 - \eta_w$; b: $1 - T_* \cdot 0.25$, $2 - T_w$, $3 - C_{Na}$.

of the spectrum will be somewhat greater than η_{vis}. Figure 3b presents the dependence of the temperature T_* of the hot plasma on the discharge axis at the time of completion of a pulse, the temperature of the tube walls T_w, and the amount of sodium C_{Na} per unit length of the tube, i.e.,

$$C_{Na}=2\pi P_{sat}\int_0^R rT^{-1}dr,$$

on the power W. Here $T(r)$ is the temperature profile in the respective discharge before the pulse. It is seen that the temperatures of the plasma and the tube wall increase with increasing W. Within the model adopted this leads to a dependence of C_{Na} on W.

Thus, when the shape and on-off time ratio of the pulses are optimized, a pulsed emitting sodium-xenon discharge has a markedly greater energy efficiency in the visible region of the spectrum in comparison to a steady-state discharge.

[1] G. N. Rokhlin, *Discharge Light Sources* [in Russian], Énergoatomizdat, Moscow (1991).

[2] Yu. V. Aĭzenberg, Svetotekhnika, No. 5–6, 1 (1993).

[3] V. O. Gornov, Svetotekhnika, No. 1, 19 (1994); (1), 24 (1995); (9), 24 (1995).

[4] S. V. Alyshev, A. M. Kokinov, L. I. Kirdyashkina, and V. V. Merkushkin, Svetotekhnika, No. 10, 6 (1995).

[5] F. G. Baksht and V. F. Lapshin, Zh. Tekh. Fiz. **66** (11), 170 (1996) [Tech. Phys. **41**, 1177 (1996)].

[6] L. M. Biberman, V. S. Vorobev, and I. T. Yakubov, *Kinetics of Nonequilibrium Low-Temperature Plasmas*, Consultants Bureau, New York (1987)

[7] I. I. Sobelman, L. A. Vainshtein, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral Lines*, Springer-Verlag, Berlin (1981).

[8] H. R. Griem, *Spectral Line Broadening by Plasmas*, Academic Press, New York (1974) [Russian trans., Mir, Moscow (1978)].

[9] G. York, R. Scheps, and A. Gallagher, J. Chem. Phys. **63**, 1052 (1975).

[10] A. Gallagher, in *Excimer Lasers*, C. K. Rhodes (ed.), Springer-Verlag, New York (1979) [Russian trans., Mir, Moscow (1981)].

[11] M. Aymar, E. Luc-Koenig, and F. Combet Farnoux, J. Phys. B: At. Mol. Phys. **9**, 1279 (1976).

[12] M. Aymar, J. Phys. B: At. Mol. Phys. **11**, 1413 (1978).

[13] A. Z. Msezane and S. Manson, Phys. Rev. A **30**, 1795 (1984).

[14] J. M. Preses, C. E. Burkhardt, R. L. Corey *et al.*, Phys. Rev. A **32**, 1264 (1985).

[15] I. I. Sobelman, *Atomic Spectra and Radiative Transitions*, Springer-Verlag, Berlin (1979).

Translated by P. Shelnitz

Magnetic trap for confining a hot plasma

K. B. Abramova and A. A. Semenov

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

The magnetic field of a two-helix Tornado trap has all the properties needed to efficiently confine a hot plasma. However, its practical utilization has been restricted because of the ponderomotive interaction between the turns of the helices, which disrupts the structure of the magnetic field. A modification of a Tornado trap, which permits significant reduction of the ponderomotive interaction is considered, and arguments in favor of using a magnetic field to maintain the properties of the field in an unmodified trap are presented. © *1997 American Institute of Physics.* [S1063-7842(97)00509-6]

The advances in the field of controlled fusion in Tokamak facilities[1,2] have made it possible to undertake the construction of a fusion reactor. However, there are physical factors that render the use of Tokamaks as reactors complicated and costly. The plasma is confined in a high magnetic field and is in a fundamentally nonequilibrium state because of synchrotron radiation. In addition, the magnetized plasma is diamagnetic and, therefore, tends to leave the magnetic field with the resultant appearance of convective instabilities. A radical method for overcoming these difficulties is to use the magnetic field not as a medium in which the plasma resides, but as a barrier surrounding the plasma on all sides.

At the current level of knowledge it seems that the confining magnetic fields must have the following properties to provide effective thermal insulation to the plasma:[3–10] 1) they must be closed, i.e., the magnetic field lines must not leave the working volume or intersect structural elements of the system, so as to eliminate the departure of charged particles from the plasma body along the field lines; 2) the working volume of the trap must include a region where the absolute value of the magnetic field strength is less than the absolute value of the magnetic field strength on its boundary, so as to eliminate large-scale, highly dangerous convective instabilities and reduce the synchrotron radiation; 3) the structure of the magnetic field in the working volume of the trap must have a layer of toroidal magnetic surfaces adjacent to the boundary, so as to ensure the stability of the magnetic field toward perturbations; 4) the poloidal component of the magnetic field must exceed the toroidal component, so as to eliminate any toroidal drift of plasma particles.

All of these requirements are satisfied by the magnetic field of a Tornado trap,[11–16] which consists of two geometrically similar, concentric spherical helices, which are wound with a constant pitch and are arranged so that if a radius drawn from the center of the system intersects the conductor of the inner helix, it also intersects the conductor of the outer helix. The helices are joined by crosspieces at the poles. Oppositely directed currents flow along the helices. The ratio between them is $I_{out}/I_{in} = \sqrt{R_{in}/R_{out}}$, where R_{in} and R_{out} are the radii of the inner and outer helices, respectively. The current circuits are closed by semi-infinite leads lying on the axis running through the poles of the helices (Fig. 1). As experiments have shown, such a trap effectively confines a

plasma.[14] However, to confine a hot plasma and carry out a thermonuclear reaction, besides the properties already enumerated, the magnetic field of the trap must have a sufficiently high strength and persist for a long time.

Nevertheless, because of the ponderomotive interaction between the current-carrying elements of the trap, the unimmobilized turns of the inner helix of a Tornado are set into motion, the structure of the magnetic field is destroyed, and the helix itself rapidly becomes useless. As was shown in Ref. 17, loads in the direction of the z axis, which tend to squeeze the inner helix, are of greatest significance. Supports cannot be introduced for the inner helix, which is located in the space occupied by the plasma, since this would lead to destruction of the plasma. A very simple way to increase the magnetic field strength is to increase the strength of the helix by using high-strength materials and increasing the thickness
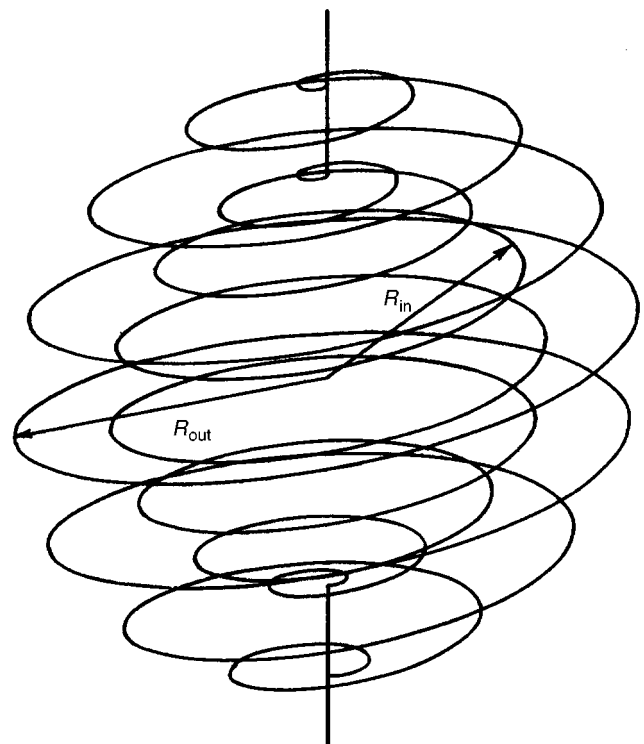

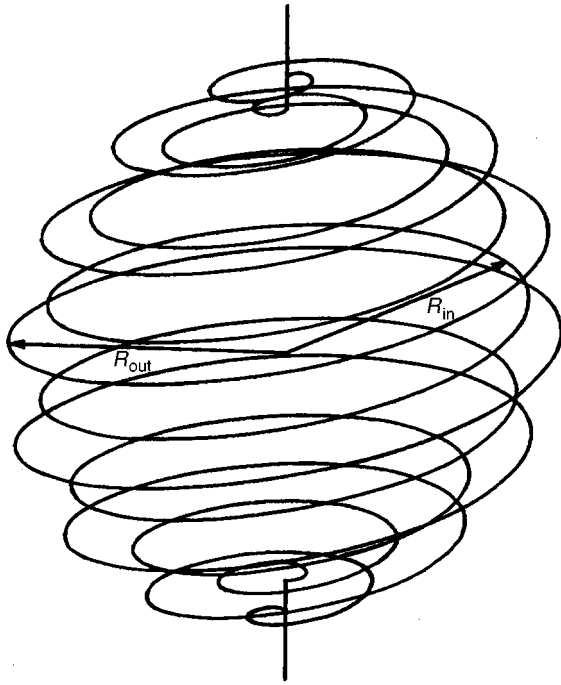FIG. 1. Schematic representation of a Tornado trap.

FIG. 2. Schematic representation of the modified Tornado trap.

FIG. 3. Schematic configuration of the current-carrying rings.

of the conductor. This route was followed in designing the Tornado-X device. The helices in this trap were fabricated from a steel rod with a diameter of 0.05 m. According to the calculations in Ref. 17, a field with a strength up to 1.5 T can be created in the barrier in Tornado-X, but only in a pulsed regime (the pulse duration τ is determined by the oscillation period T of the inner helix of the trap from the condition $\tau/T \ll 1$). The stationary magnetic field in the barrier of the Tornado-X trap cannot exceed 0.5–0.9 T. Therefore, the only possibility for further increasing the magnetic field strength is to modify the design of the trap. Modifying the design would be useful, if it would be possible to find a geometric configuration for the conductors of the inner helix in which the ponderomotive forces have a minimum (and the merits of the unmodified trap are preserved) that is sufficient for it to be possible to significantly increase the strength of the confining field.

For this purpose, in 1983 Varshavskiĭ et al.[18] proposed positioning the turns of the outer helix on extensions of radii passing exactly between the turns of the inner helix, i.e., turning the helices to an angle $\Delta\varphi = \beta = \pi$ relative to one another (Fig. 2). The outer helix can be rigidly immobilized, since it located outside the working volume of the trap. It is difficult to analytically treat the problem of the ponderomotive interaction between spherical helices, but stability criteria can be obtained from an examination of simple models, and then experimental verification can be made on a real trap.

The possibility of stabilizing the position of the unimmobilized turns can also be demonstrated in the case of two infinite series of rings with a current that rest freely on coaxial cylinders of different diameter (Fig. 3). Let the distance between the rings in each of the series be equal to a, and let the series be shifted relative to one another by $0.5a$. The
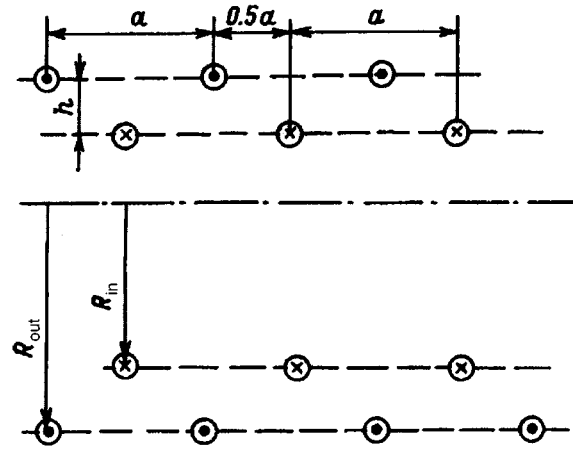
currents in the rings in each series are equal in magnitude and oppositely directed. In this case the forces exerted on each of the rings with a current by the remaining rings are clearly equal to zero, i.e., they are in equilibrium positions. Now, let us assume that the rings lying on the outer cylinder are rigidly immobilized and that the radii of both series coincide. If one of the unimmobilized rings is shifted parallel to the z axis by a distance less than $0.5a$, a force which returns it to the equilibrium position will act on that ring. As the calculations showed, a stabilization effect is observed, if the difference between the radii of the rings $h \leqslant 0.5a$.[18]

The possibility of the existence of a stable position for the inner helix was tested experimentally on a 12-turn model of the Tornado trap by A. S. Varshavskiĭ, A. V. Voronin, and V. M. Kuznetsov. The diameter of the inner helix was 400 mm, and the diameter of the outer helix was 440 mm, i.e., $h = 0.35a$. The helices were fabricated from a 14 mm rod and turned to an angle equal to π relative to one another about the axis passing through their poles. The ratio between the currents in the inner and outer helices was selected to correspond to the existence of a spherical separatrix.[13] When a current (0.5−20 kA) was supplied, some displacement of the turns of the inner helix occurred in the trap, but it was insignificant compared with the distance between the two helices, and they were fixed in a position of stable equilibrium. No experiments were performed for $h > 0.5a$, but the ponderomotive forces acting on the inner helix of an unmodified and a modified five-turn trap with a ratio between the radii of the helices $R_{in}/R_{out} = 0.7$, i.e., with $h = 0.68a$, were calculated in Ref. 17. It is shown that the ponderomotive forces acting on the turns of the inner helix of the modified trap in the z direction are smaller than the corresponding forces acting on the turns of the unmodified trap, although they are still comparable. Therefore, the criterion obtained for solenoids can also be used for spherical helices. Thus, these findings allow us to state that the inner helix of a trap can be fixed in a position of stable equilibrium by properly selecting the ratio between the radii of the helices and turning the helices to an angle $\Delta\varphi = \pi$ relative to one another about the axis passing through the poles. Therefore, the currents in the helices can

be increased, and the magnetic field strength in the trap can thereby be increased.

It was shown in Ref. 17 that a magnetic field in the barrier of the order of 1.5 T can be obtained in a stationary regime for helices turned to an angle equal to π relative to one another provided they are fabricated from high-strength materials. However, as we have already indicated above, the calculations in Ref. 17 were performed for $h = 0.68a$, i.e., for conditions that do not correspond to the stability criterion $h < 0.5a$; when it is satisfied, the magnetic induction which can be achieved in the system can be considerably greater than 1.5 T.

The turning of the helices violates one of the conditions for the existence of a spherical separatrix,[13] and questions regarding the closed state, structure, and stability of the magnetic field arise. Experimental investigations of the magnetic field in a so-called trap with opposing turns (one of the helices is a left-hand winding, and the other is a right-hand winding) of Tornado II[11] and the results of experiments on the confinement of a cold plasma in it[19,20] allow us to hope that the structure of the magnetic field is stable even toward such gross perturbations.

A quantitative analysis of the magnetic field can also be performed. Let us consider an electromagnetic system consisting of two similar concentric spherical helices wound with the same constant pitch. The helices are turned to an angle β relative to one another about the axis passing through their points, and oppositely directed currents related by the ratio $I_{out}/I_{in} = \sqrt{R_{in}/R_{out}}$ flow in them.[13] Here I_{in} and I_{out} are the values of the currents in the inner and outer helices, respectively, and R_{in} and R_{out} are the radii of the helices. The current circuits are closed by two rectilinear crosspieces that connect the ends of the helices and two semi-infinite conductors lying on the same axis as the crosspieces.

The equations of the helices in the case under consideration have the form

$$\theta_0 = \alpha\varphi_0, \quad R = R_{in}, \quad -N\pi \leq \varphi_0 \leq N\pi,$$

$$\theta_0 = \alpha(\varphi_0 - \beta), \quad R = R_{out}, \quad -N\pi \pm \beta \leq \varphi_0 \leq N\pi + \beta, \tag{1}$$

where N is the number of turns in the helix, and $\alpha = 1/2N$.

Generalizing the results in Ref. 13, we can easily obtain expressions for the components of the magnetic induction. As a result, it is found that the symmetric part of the magnetic induction, i.e, the part which does not depend on the angle φ, does not depend on the angle of rotation of the helices relative to one another.

A similar treatment was performed for a trap with opposing turns. When the outer helix is turned relative to the inner helix, their equations have the form

$$\theta_0 = \alpha\varphi_0, \quad R = R_{in},$$

$$-N\pi \leq \varphi_0 \leq N\pi, \quad -\pi/2 \leq \theta_0 \leq \pi/2,$$

$$\theta_0 = -\alpha(\varphi_0 - \beta), \quad R = R_{out},$$

$$-N\pi \pm \beta \leq \varphi_0 \leq N\pi + \beta, \quad -\pi/2 \leq_0 \leq t\pi/2. \tag{2}$$

In this case, too, it was found that the symmetric part of the components of the magnetic induction does not depend on the angle of rotation of the helices and is equal in magnitude to the analogous part of the magnetic induction for the case described above.

When the ratio between the currents $I_{out}/I_{in} = \sqrt{R_{in}/R_{out}}$, this has the following consequences.

1. The magnetic flux within the annular contour, whose plane is perpendicular to the z axis and which lies at $R_c = \sqrt{R_{in}, R_{out}}$, is equal to zero in all the cases.

2. The expression for the magnetic induction on the axis of the system within the volume of the outer helix is identical for all cases and has the following form

$$B_z = \frac{\mu_0 N I_{in}}{\pi z}$$

$$\times \begin{cases} \frac{R_{in}}{z}\left[K\left(\frac{z}{R_{in}}\right) - E\left(\frac{z}{R_{in}}\right)\right] - \frac{\sqrt{R_{in}R_{out}}}{z} \\ \times \left[K\left(\frac{z}{R_{out}}\right) - E\left(\frac{z}{R_{out}}\right)\right], \quad R \leq R_{in}, \\ \left[K\left(\frac{R_{in}}{z}\right) - E\left(\frac{R_{in}}{z}\right)\right] - \frac{\sqrt{R_{in}R_{out}}}{z} \\ \times \left[K\left(\frac{z}{R_{out}}\right) - E\left(\frac{z}{R_{out}}\right)\right], \quad R_{in} < R < R_{out}, \end{cases} \tag{3}$$

where $K(x)$ and $E(x)$ are compete elliptic integrals.

In particular, the field at the center of the system is

$$B_z(0,0,0) = \frac{\mu_0 N I_{in}}{4 R_{in}}\left[1 - \left(\frac{R_{in}}{R_{out}}\right)^{3/2}\right]. \tag{4}$$

3. The magnetic flux through a surface which is perpendicular to the z axis and is bounded by an annular contour lying on a sphere of radius R_{in} is equal in magnitude and opposite in sign to the magnetic flux piercing the annular surface bounded by the same contour and the contour lying on a sphere of radius $R_c = \sqrt{R_{in}R_{out}}$. This is the magnetic flux of the poloidal field, and it equals

$$\Phi_{pol} = \frac{\mu_0 I_{in} R_{in}}{2}\sum_{n=1}^{\infty}\left\{1 - \sqrt{\frac{R_{in}}{R_{out}}}\left(\frac{R_{in}}{R_{out}}\right)^n\right\}$$

$$\times A_n P_n^1(\cos\theta)\sin\theta. \tag{5}$$

Here

$$A_n = \frac{1}{n(n+1)}\int_{-N\pi}^{N\pi} P_n^1(\sin\alpha\varphi_0)\cos\alpha\varphi_0 d\varphi_0$$

$$= \frac{2N}{n(n+1)}\int_{-1}^{1} P_n^1(x)dx$$

$$= \begin{cases} 0 & n = 2k, \\ -\frac{\pi N}{2}\left[\frac{(2k-1)!!}{2^k(k+1)!}\right]^2 & n = 2k+1. \end{cases}$$

Thus,

$$\Phi_{\text{pol}} = -\frac{\pi\mu_0 N I_{\text{in}} R_{\text{in}}}{4} \sum_{k=1}^{\infty} \left\{ 1 - \sqrt{\frac{R_{\text{in}}}{R_{\text{out}}}} \left(\frac{R_{\text{in}}}{R_{\text{out}}}\right)^{2k+1} \right\}$$

$$\times \left[\frac{(2k-1)!!}{2^k(k+1)!}\right] P_{2k+1}^1(\cos\theta)\sin\theta. \tag{6}$$

Here $(-1)!! \equiv 1$. Expression (6) allows us to obtain the mean value of the poloidal component of the magnetic induction within the inner helix and between the inner helix and $R_c = \sqrt{R_{\text{in}}R_{\text{out}}}$. The ratio of the field in the barrier to the field within the trap can be evaluated.

4. The toroidal component of the field within the inner helix averaged over φ is equal to zero, and the field between the outer and inner helices is equal to the magnitude of the magnetic induction of an infinite straight conductor

$$B_\varphi = \frac{\mu_0 I_{\text{in}}}{2R_{\text{in}}\sin\theta}. \tag{7}$$

All of the foregoing material allows one to hope that the closed state of the magnetic field in the working volume of the trap can be preserved. If the distance between the helices $h < 0.5a$ and they are turned to an angle equal to π relative to one another, it will be possible to achieve magnetic fields considerably greater than the fields in the Tornado traps that have been and are still being investigated investigated, because of the decrease in the ponderomotive interaction in such a modified Tornado trap.

At the present time we do not see a way to perform an exact analytical investigation of the closed state of the structure of the magnetic field in the proposed modified Tornado trap; therefore, this investigation is being carried out by numerical methods.

[1] M. Keilhacker, in *Abstracts of Invited and Contributed Papers, 22nd EPS Conference on Controlled Fusion and Plasma Physics*, Bournemouth, England (1995), p. 1.

[2] B. J. D. Tubbing, A. Chankin, S. Clement *et al.*, in *Abstracts of the 22nd EPS Conference on Controlled Fusion and Plasma Physics*, Bournemouth, England, (1995), Vol. 19C, Part 3, pp. 453–456.

[3] L. A. Artsimovich, *Controlled Thermonuclear Reactions*, A. Kolb and R. S. Pease (Eds.), Gordon and Breach, New York (1964).

[4] L. S. Solov'ev and V. D. Shafranov, in *Reviews of Plasma Physics*, Vol. 5, edited by M. A. Leontovich (Consultants Bureau, New York, 1970) [Vopr. Teor. Plazmy, No. 5, 3 (1967)].

[5] A. I. Morozov, Fiz. Plazmy **18**, 305 (1992) [Sov. J. Plasma Phys. **18**, 159 (1992)].

[6] G. V. Skornyakov, Zh. Tekh. Fiz. **32**, 261 (1962) [Sov. Phys. Tech. Phys. **7**, 187 (1962)].

[7] G. V. Skornyakov, Zh. Tekh. Fiz. **32**, 777 (1962) [Sov. Phys. Tech. Phys. **7**, 571 (1963)].

[8] G. V. Skornyakov, Zh. Tekh. Fiz. **32**, 1494 (1962) [Sov. Phys. Tech. Phys. **7**, 1104 (1963)].

[9] G. V. Skornyakov, Zh. Tekh. Fiz. **34**, 1126 (1964) [Sov. Phys. Tech. Phys. **9**, 875 (1964)].

[10] G. V. Skornyakov, J. Nucl. Energy, Part C **8**, 561 (1966).

[11] K. B. Abramova, G. A. Galechyan, and B. P. Peregud, Zh. Tekh. Fiz. **36**, 1426 (1966) [Sov. Phys. Tech. Phys. **11**, 1064 (1967)].

[12] A. N. Kozyrev and B. P. Peregud, Zh. Tekh. Fiz. **40**, 950 (1970) [Sov. Phys. Tech. Phys. **15**, 734 (1970)].

[13] B. P. Peregud and A. A. Semenov, Zh. Tekh. Fiz. **41**, 2297 (1971) [Sov. Phys. Tech. Phys. **16**, 1824 (1972)].

[14] B. P. Peregood and B. Lehnert, Nucl. Instrum. Methods Phys. Res. **180**, 357 (1981).

[15] K. B. Abramova, G. A. Galetchyan, M. L. Lev *et al.*, in *Abstracts of the 22nd EPS Conference on Controlled Fusion and Physics*, Bournemouth, England (1995), Vol. 19C, Part 2, pp. 164–168.

[16] K. B. Abramova, A. V. Voronin, M. L. Lev *et al.*, Zh. Tekh. Fiz. **67**, 30 (1997) [Tech. Phys. **42**, 25 (1997)].

[17] A. V. Voronin, V. M. Kuznetsov, B. P. Peregud *et al.*, Zh. Tekh. Fiz. **57**, 469 (1987) [Sov. Phys. Tech. Phys. **32**, 283 (1987)].

[18] A. S. Varshavskiĭ, A. V. Voronin, V. M. Kuznetsov *et al.*, USSR Inventor's Certificate (Patent) No. 1,145,900 (15 November 1984).

[19] G. A. Galechyan and B. P. Peregud, Zh. Tekh. Fiz. **39**, 1696 (1969) [Sov. Phys. Tech. Phys. **14**, 1272 (1970)].

[20] G. M. Vorobiev *et al.*, in *3rd European Conference on Controlled Fusion and Plasma Physics. Confinement of Plasma in Tornado Traps*, Utrecht (1969).

Translated by P. Shelnitz

Nonlinear dynamics of the critical state in hard superconductors and composites based on them

V. R. Romanovskiĭ

Kurchatov Institute Russian Science Center, 123182, Moscow, Russia
(Submitted December 18, 1995)
Zh. Tekh. Fiz. **67**, 29–33 (September 1997)

The nonlinear dynamics of the magnetic flux within a superconducting plate in response to the continuous rise in temperature over the course of the entire process of applying the magnetic field is investigated within the critical-state model. The results of numerical simulations based on a method developed to solve the system of Fourier and Maxwell equations with an unknown internal magnetic flux penetration boundary are compared with the corresponding analytical expressions of the isothermal theory. It is shown that the difference between the isothermal and nonisothermal models increases as the heat transfer coefficient decreases and as the rate of increase in the magnetic field strength and the transverse dimensions of the superconductor increase. The errors appearing in the isothermal approximation are very significant in the case of a thermally insulated, massive conductor. Consequently, the calculated values of the thermal losses occurring during the time period preceding the flux jump in the isothermal approximation can be significantly lower than the corresponding nonisothermal values. © *1997 American Institute of Physics.* [S1063-7842(97)00609-0]

INTRODUCTION

Hard superconductors are known to pass into a critical state in response to any external influence that gives rise to an electric field in them.[1] The study of the dynamics of the critical state underlies the solution of several fundamental problems in the physics of superconductors and composite structures based on them in the form of a regular composite of a large number of filaments of a hard superconductor in a normally conducting matrix. A theory of thermomagnetic instability,[2–4] which permits determination of the conditions for stability of the superconducting state with respect to infinitesimal perturbations, was developed using this conception. The results of numerous investigations of the energy losses in superconductors and of the conditions for the appearance of instabilities of various nature within a linear approximation that presumes an insignificant increase in the temperature of the superconductor were presented in Refs. 5–8. The linear models greatly simplify the original system of equations used to describe the dynamics of the critical state. However, no attention was focused on the influence of several factors (for example, the features of the heat transfer between the conductor and the coolant, the character of the variation of the external magnetic field, etc.) on the shaping of the temperature field in the superconductor before the appearance of instability. At the same time, it was shown in Ref. 9 that the variation of the background temperature of the superconductor plays a significant role under conditions in which the critical state is stable. Assigning its initial temperature *a priori* will distort the results of an analysis of the boundary of the stable states.

It is also noteworthy that the few existing studies of the problem of the stability of the critical state in the nonisothermal approximation[10–14] not only do not permit the formulation of general laws for determining the permissible increase in the temperature of the superconductor, but also contain some erroneous results because of inadequate estimation of the role of the temperature factor in the appearance of instability. For example, according to Ref. 13, the dependence of the critical current on the rate at which the current is introduced is nonmonotonic. The solution of this problem from the standpoint of the existence of a correlation between the permissible increase in the temperature of a composite and the rate at which the current is introduced into it[15] does not support this conclusion: as the rate of introduction increases, the critical current decreases monotonically, approaching its adiabatic limit. The determination of the conditions for adiabatic stability in Ref. 14 yielded a criterion, according to which consideration of the influence of the background temperature of the superconductor leads to a decrease in the size of the region of stable states in comparison with the known isothermal criterion. However, the analysis of the adiabatic conditions for stability of the critical state performed under the assumption of the existence of a finite permissible increase in the temperature of the superconductor in Ref. 9 reveals the opposite tendency, i.e., the isothermal stability criterion imposes excessively strong restrictions on the parameters of the superconductor in comparison with the corresponding nonisothermal stability condition.

Thus, a correct investigation of the dynamics of the critical state should be performed on the basis of an analysis of the continuous variation of the temperature and electromagnetic fields within the superconductor. This paper presents the results of numerical simulations of the nonisothermal diffusion of magnetic flux into a superconducting plate in the case of an increase in the external uniform magnetic field at a constant rate.

FORMULATION OF THE PROBLEM AND METHOD FOR ITS SOLUTION

Let us consider the simplest problem of the penetration of magnetic flux into a cooled plane-parallel superconducting

plate in a uniform external magnetic field, which is parallel to its surface and increases at a constant rate. We shall determine the distribution of the temperature in a transverse section of a plate initially cooled to the coolant temperature T_0 from the solution of a heat-conduction equation of the form

$$c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\lambda \frac{\partial T}{\partial x}\right) + \begin{cases} 0, & 0 \leqslant x \leqslant x_p, \\ EJ, & x_p < x < a, \end{cases} \tag{1}$$

in which the size of the heat-evolving region and the density of the Joule losses depend on the magnetic flux penetration depth $x_p(t)$, which can be described by its continuity equation

$$\mu_0 \int_{x_p}^{a} J(x,t)dx = \frac{dB}{dt} t, \tag{2}$$

as well as on the corresponding distribution of the shielding current and the electric and magnetic fields, which satisfy the system of equations ($x_p < x < a$, $t > 0$)

$$\frac{\partial B}{\partial x} = \mu_0 J, \quad \frac{\partial E}{\partial x} = \frac{\partial B}{\partial t}, \quad J = \eta J_C + \frac{1}{\rho}E. \tag{3}$$

Here c is the volumetric specific heat of the plate, λ is its thermal conductivity, a is the half width of the plate, ρ is its resistivity, η is the percentage of the plate filled by the superconductor, and J_C is the critical current density of the superconductor.

To simplify the analysis of the results obtained, it is assumed that J_C depends only on the temperature:

$$J_C = J_{C0} \frac{T_{CB} - T}{T_{CB} - T_0} \quad (J_{C0} \text{ and } T_{CB} \text{ are constants}).$$

The initial and boundary conditions imposed on system of Eqs. (1)–(3) has the form

$$T(x,0) = T_0, \quad B(x,0) = 0,$$

$$\lambda \frac{\partial T}{\partial x}(a,t) + h(T(a,t) - T_0) = 0, \quad B(a,t) = \frac{dB}{dt} t, \tag{4}$$

$$\frac{\partial T}{\partial x}(0,t) = 0, \quad B(x_p,t) = 0.$$

The problem consisting of Eqs. (1)–(4) describes a dissipative magnetic flux diffusion process, as a result of which the shielding currents penetrate into the plate from its surface and exist only in the region $x_p < x < a$ in accordance with the critical-state model. The system of equations written down permits investigation of the dynamics of the critical state both in a hard superconductor and in a superconducting composite under the assumption that the superconductor is distributed uniformly over the cross section of the plate within the one-dimensional model. In the former case we should set $\eta = 1$, and in the latter case the theoretical model should be supplemented by the so-called mixture relations[16]

$$c = \eta c_s + (1-\eta)c_m, \quad \frac{1}{\rho} = \frac{\eta}{\rho_s} + \frac{1-\eta}{\rho_m},$$

$$\lambda = \lambda_m \frac{(1-\eta)\lambda_m + (1+\eta)\lambda_s}{(1+\eta)\lambda_m + (1-\eta)\lambda_s},$$

in which the subscripts s and m refer to the superconductor and the matrix, respectively.

The boundary conditions take into account the convective heat transfer between the plate and the coolant with an assigned heat-transfer coefficient h and an increase in the external magnetic field at a constant rate dB/dt. The presence of the special condition (2), which contains the unknown moving magnetic-field penetration boundary, renders the problem defined by Eqs. (1)–(4) significantly nonlinear even when both the original different equations and the boundary conditions are linear. Moreover, the implicit form in which the temporal variation law $x_p(t)$ is written complicates the use of known methods for solving equations of the parabolic type for multiphase regions with an unknown phase boundary. Therefore, a numerical method which takes into account the specific features of the problem under consideration was developed to solve it. It was based on a through-calculation algorithm for the difference analog of system of Eqs. (1)–(4), which approximates the values of the temperature, electric field strength, and magnetic induction sought on a spatiotemporal net with an implicit template.[17] In addition, x_p was determined by an iterative procedure based on the existence of a root for the nonlinear equation (2), in which x_p is the quantity sought. Within this method the calculations performed are confined to isolating and then refining the root of Eq. (2). In the first stage the value of x_p for each new temporal step is assumed to be known (the corresponding value is taken either from the preceding temporal step or from the preceding iteration). After the necessary calculations have been performed for each successive iteration number $s = 1, 2, \ldots$, the sign of the expression

$$r^{(s)} = \mu_0 \int_{x_p^{(s)}}^{a} J(x,t)dx - \frac{dB}{dt} t$$

is found.

It is not difficult to see from the simple physical meaning of Eq. (2) that when $x_p^{(s)}$ is greater than the true value of x_p, the sign of $r^{(s)}$ is negative. Conversely, if $x_p^{(s)} < x_p$, then $r^{(s)} > 0$. Therefore, the stage of isolating the root is ended when the signs of $r^{(s)}$ for two successive iterations differ. After this, it is not difficult to refine the root with an assigned accuracy in the range

$$x_p^{(s)}|r^{(s)} < x_p < x_p^{(s-1)}|r^{(s-1)}.$$

Since an implicit difference scheme was used to determine the net functions, both the temporal and spatial mesh widths are determined only by accuracy considerations. Therefore, a large temporal mesh width can be selected in the initial stage of application of the magnetic field for the purpose of reducing the computation time. After instability appears, it must be diminished, since this process has an avalanche character and proceeds very rapidly (the computational process is actually carried out in two stages: at first, the calculations are performed with a large temporal mesh
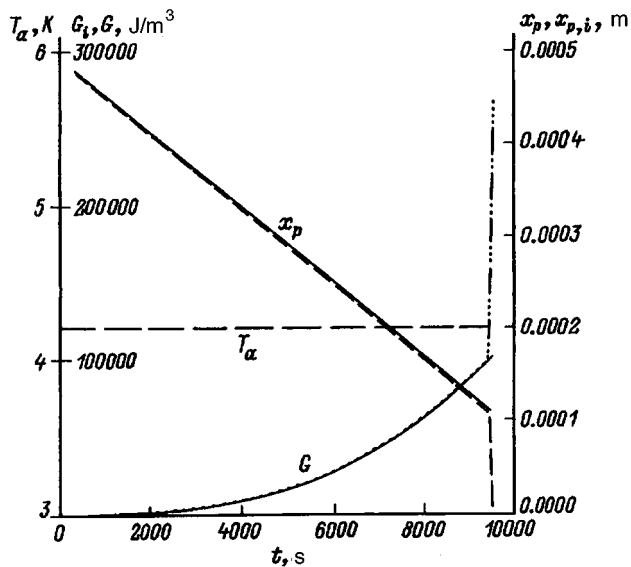
FIG. 1. Time dependence of the temperature, the penetration depth, and the thermal losses for slow application of the external magnetic field to a cooled superconducting composite.
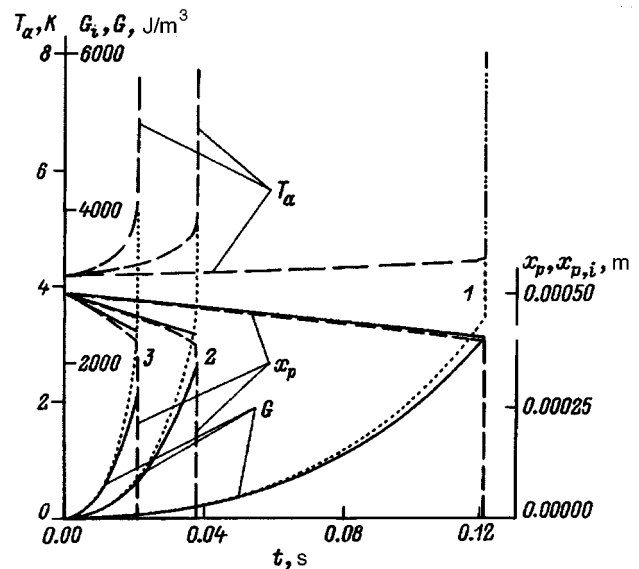

FIG. 2. Time dependence of the temperature, the penetration depth, and the thermal losses for various rates of application of the external magnetic field to a cooled superconducting composite ($h=100$ W/m$^2\cdot$K). dB/dt, T/s: 1 — 2, 2 — 6, 3 — 10.

width, and then after instability appears, they are repeated with a smaller mesh width using the results already obtained).

The algorithm described is implemented without significant difficulties in the analysis of the dynamics of the critical state in superconductors of cylindrical shape in a varying longitudinal or transverse magnetic field, as well as in the determination of the boundaries of their metastable states when a current is introduced.

RESULTS OF NUMERICAL EXPERIMENTS

To verify the proposed solution method, the results of numerical simulations were compared with the known analytical expressions following from the isothermal model.[5–8] In particular, for a plane-parallel plate it is not difficult to write formulas which describe the temporal variation of the penetration boundary of the shielding currents and the bulk density of the thermal losses occurring during the diffusion of magnetic flux. In terms of the present work these functions are described by the expressions

$$x_{p,i}(t)=a-\frac{dB/dt\ t}{\mu_0\eta J_{C0}}, \quad G_i(t)=\frac{(dB/dt)^3 t^3}{6\mu_0^2\eta J_{C0}a}.$$

Figure 1 presents the results of the corresponding analytical (solid lines) and numerical (dashed lines) calculations of the surface temperature of a cooled plate, the penetration depth, and the thermal losses

$$G=\frac{1}{a}\int_0^t\int_{x_p}^a EJdxdt.$$

The starting values of the parameters used to perform the calculations were $a=5\times10^{-4}$ m, $c=10^3$ J/m^3 K, $\lambda=200$ W/m K, $h=10$ W/m^2 K, $dB/dt=10^{-3}$ T/s, $\eta=0.5$ $\rho_m=2\times10^{-10}$ $\Omega\cdot$m, $\rho_s=5\times10^{-7}$ $\Omega\cdot$m, $J_{C0}=4\times10^9$ A/m^2, $T_{CB}=9$ K, and $T_0=4.2$ K. They describe the aver-

aged thermo- and electrophysical parameters of a niobium-titanium superconductor in a copper matrix cooled by liquid helium.

As follows from Fig. 1, when there is cooling and the rate of variation of the external magnetic field is very small, the temperature of the plate increases only slightly up to the appearance of instability, i.e., a flux jump. Therefore, in this stage of the diffusion of magnetic flux, the finite-difference approximation used and the analytical calculations coincide with a high degree of accuracy. However, the calculations of the dynamics of the critical state after the development of instability within the isothermal and nonisothermal models exhibit significant disparity. It is not difficult to see that this difference is based on the mutual influence of the corresponding variations of the thermal and electromagnetic fields occurring within the plate on one another. Consequently, as is clearly shown in Fig. 1, the development of instability is characterized by practically instantaneous filling of the entire transverse section of the composite by shielding currents. As a result, its temperature increases sharply with a resultant subsequent decrease in the shielding current density. Therefore, the final result of a flux jump can be a loss of superconducting properties by the sample.

The difference between the isothermal and nonisothermal models can also be manifested over the course of the entire magnetic flux diffusion process. Curves describing the variation of the composite temperature, the magnetic flux penetration depth, and the thermal losses with time as a function of the heat transfer coefficient, the rate of increase in the magnetic field, and the thickness of the plate are plotted in Figs. 2 and 3. It is seen that the temperature of the composite rises more rapidly when h decreases and dB/dt and a increase. As a result, the corresponding plots of $x_p(t)$ and $G(t)$ obtained in the nonisothermal approximation (the dashed curves) deviate to a greater degree from their isothermal val-
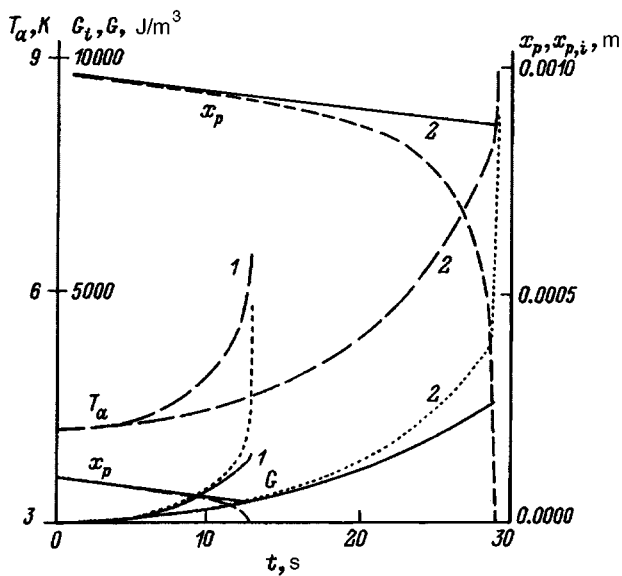
FIG. 3. Time dependence of the temperature, the penetration depth, and the thermal losses in an uncooled superconducting composite for $dB/dt=0.01$ T/s, $h=0$, and $a=10^{-4}$ (*1*) and $a=10^{-3}$ (*2*).
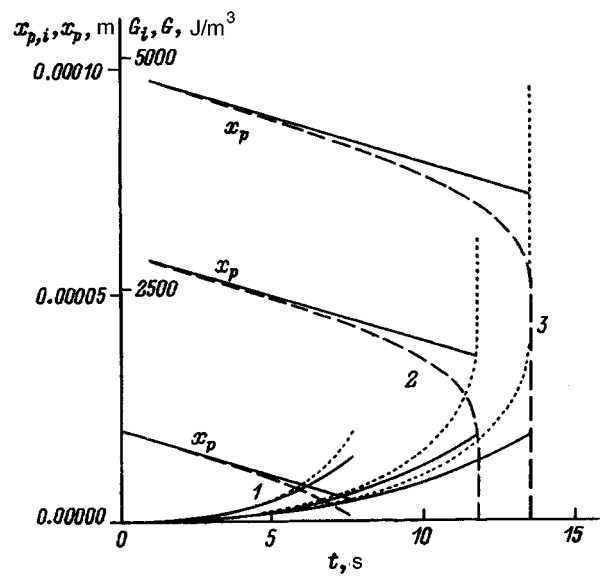

FIG. 4. Time dependence of the penetration depth and the thermal losses in an uncooled superconductor for $dB/dt=0.01$ T/s and $h=0$. a, m: *1* — 2×10^{-5}, *2* — 6×10^{-5}, *3* — 10^{-4} m.

ues (the solid curves). Since the error introduced by assuming an isothermal state increases with time, its value can become very significant before the appearance of instability. For example, the disparity between the calculated values of the thermal losses for high rates of increase of the magnetic field can reach 100% even in the case of a cooled composite. Since the influence of the nonisothermal state on the processes taking place is manifested to a greater degree in uncooled current-carrying elements, similar errors can occur in massive superconducting composites even when the rate of variation of the external magnetic field is relatively small (Fig. 3).

The calculation results presented in Figs. 1, 2, and 3 were obtained for a composite with a matrix of high thermal conductivity. As we know, its low electrical conductivity has a damping effect on the conditions for the appearance of instability. Therefore, when the plate has a poorly conducting matrix, the nonisothermal diffusion of magnetic flux can have a more nonlinear character before and especially after the appearance of instability. To illustrate this point, Fig. 4 presents dashed curves which describe the dynamics of the magnetic flux penetration depth and the thermal losses in the most unfavorable case from the standpoint of the variation of the thermal state of the plate, i.e., in a thermally insulated, hard superconductor without a stabilizing matrix ($\eta=1$). The starting parameters were $a=5\times10^{-4}$ m, $c=30\,T^3$ J/m$^3\cdot$K, $\lambda=0.0075\,T^{1.8}$ W/m$\cdot$K, $h=0$ W/m^2 $\cdot$K, $dB/dt=0.01$ T/s, $\rho_s=5\times10^{-7}\Omega\cdot$m, $J_{C0}=4\times10^9$A/m^2, $T_{CB}=9$ K, and $T_0=4.2$ K. Here the solid lines show the respective dependences following from the isothermal model. It is not difficult to see that the nonlinear dynamics of the critical state are characterized by more intense penetration of magnetic flux into the superconductor than in the linear approximation. Therefore, its nonisothermal state, on the one hand, promotes uniform distribution of the shielding currents over the cross section of

the superconductor and, on the other hand, leads to an increase in the level of thermal losses. Taking into account the corresponding decrease in the critical current density, we should expect that the thermal history of the superconductor can have a significant influence on the conditions for the appearance and development of instability under adiabatic conditions. Figure 5 shows the results of the calculation of the surface temperature of a hard superconductor ($a=10^{-4}$ m) for various functional relationships between the specific heat and the temperature. The solid lines describe the temperature increase in cases where the shielding currents do not completely fill the cross section of the plate $[0<x_p(t)<a]$, and the dashed curves describe the increase after complete penetration of the magnetic flux into the superconductor $[x_p(t)=0]$. The results presented graphically demonstrate that the transition from one theoretical dependence of the specific heat on the temperature to another not only alters the
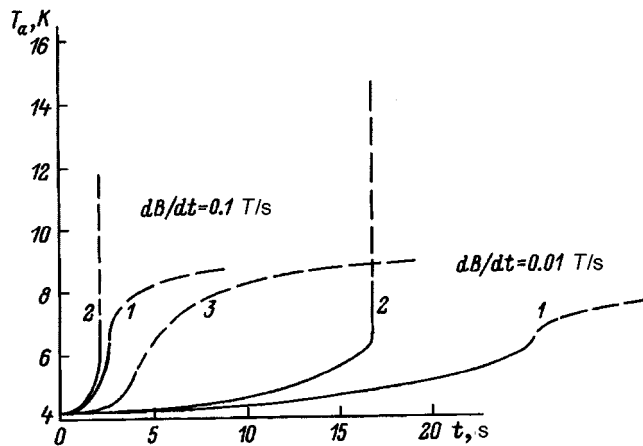

FIG. 5. Increase in the surface temperature of a thermally insulated superconductor for various dependence of the specific heat on the temperature. *1* — $c(T)=30T^3$ J/(m^{3}K), *2* — $c=c(T)|_{T=4.2\,K}$, *3* — $c=c(T)|_{T=9\,K}$.

form of the $T_a(t)$ curves, but also influences the final temperature, to which the superconductor is heated during the entire magnetic flux diffusion process. The results of the calculation of the stability boundary of the critical state of a hard superconductor under adiabatic conditions are modified accordingly.

Thus, the distinguishing feature of the nonisothermal thermal state of hard superconductors and composites based on them is the more intense penetration of magnetic flux into the sample. Therefore, consideration of the temperature factor in an analysis of the dynamics of the critical state leads to a more uniform distribution of the shielding currents over a cross section of the superconductor. At the same time, however, the errors of the isothermal approximation increase when the thermal losses are calculated. These laws are most pronounced in the absence of cooling, at high rates of variation of the external magnetic field, and when the transverse dimensions of the conductor are increased. This must be taken into account in developing large-scale superconducting magnetic systems, primarily with poorly cooled, massive current-carrying elements immersed in rapidly varying magnetic fields.

[1] C. P. Bean, Rev. Mod. Phys. **3**, 31 (1964).

[2] S. L. Wipf, Phys. Rev. **161**, 404 (1967).

[3] P. S. Swartz and C. P. Bean, J. Appl. Phys. **39**, 4991 (1968).

[4] M. G. Kremlev, JETP Lett. **17**, 223 (1973).

[5] H. Brechna, *Superconducting Magnet Systems*, Springer-Verlag, Berlin (1973) [Russian translation, Mir, Moscow, 1976].

[6] V. A. Al'tov, V. B. Zenkevich, M. G. Kremlev, and V. V. Sychev, *Stabilization of Superconducting Magnetic Systems*, Plenum Press, New York (1977).

[7] M. N. Wilson, *Superconducting Magnets*, Oxford University Press, London (1983) [Russian translation, Mir, Moscow, 1985].

[8] A. V. Gurevich, R. G. Mints, and A. L. Rakhmanov, *The Physics of Composite Superconductors*, CRC Press, Boca Raton, FL (1995).

[9] V. R. Romanovskiĭ, JETP Lett. **59**, 709 (1994).

[10] N. Morton and M. I. Darby, Cryogenics **13**, 232 (1973).

[11] J. F. Bussière and M. A. R. LeBlanc, J. Appl. Phys. **46**, 406 (1975).

[12] E. A. Gijsbertse, L. J. M. van de Klundert, and W. J. van Weezep, Cryogenics **21**, 419 (1981).

[13] E. Yu. Klimenko and N. N. Martovetsky, IEEE Trans. Magn. **MAG- 24**, 1167 (1988).

[14] E. Yu. Klimenko, N. N. Martovetskiĭ, and S. I. Novikov, Sverkhprovodimost' (KIAE) **2** (11), 152 (1989) [Supercond. Phys. Chem. Technol. **2**, 171 (1989)].

[15] V. E. Keilin and V. R. Romanovskii, Cryogenics **33**, 986 (1993).

[16] *Fibrous Composite Materials*, J. W. Weeton (ed.) [Russian translation, Metallurgiya, Moscow, 1978].

[17] A. A. Samarskiĭ, *Theory of Difference Schemes* [in Russian], Nauka, Moscow (1983).

Translated by P. Shelnitz

Characteristics of the relaxation to steady-state deformation in solids

S. G. Psakh'e, A. Yu. Smolin, E. V. Shil'ko, S. Yu. Korostelev, A. I. Dmitriev, and S. V. Alekseev

Institute of Physics of Strength and Materials Science, Siberian Branch, Russian Academy of Sciences, 634055 Tomsk, Russia
(Submitted March 11, 1996)
Zh. Tekh. Fiz. **67**, 34–37 (September 1997)

The characteristics of the transient process of relaxation to steady-state deformation in solids are investigated theoretically. Various loading regimes are modeled by the method of mobile cellular automata. It is shown that the stressed state in a material is highly inhomogeneous in the relaxation stage; this property, in turn, can produce stable structures in the velocity field of the material particles and influence the evolution of deformation in later stages. © *1997 American Institute of Physics.* [S1063-7842(97)00709-5]

The behavior of materials under mechanical loading is usually investigated, both experimentally and theoretically, in the stages of steady-state deformation and at prefracture strain levels.[1,2] On the other hand, scarcely any attention has been given to the specific attributes of the initial stage of deformation. The only way this problem can be studied experimentally at the present time is on the basis of global characteristics of the response of materials, such characteristics as the total strain of the sample, acoustic-emission spectra, etc. Even in this case the instrumentation must have high resolution (in the range 10^{-6}–10^{-9} s). The spatial distribution of strains can be measured very accurately on the basis of the special television-optical system described in Refs. 3 and 4. This device is capable of measuring displacement vectors on the surface of a deformable material with a resolution of 1500 points/mm^2. In combination with sufficiently high-performance electronics, such instruments can be used to investigate the characteristics of the transient process leading to steady-state deformation. Methods based on continuum mechanics are currently used for theoretical investigations of the response of materials at the mesoscopic level. At the same time, discrete approaches are beginning to make greater inroads into modeling.[5–12] Apart from their other advantages, these approaches are far less demanding in computer resources, an asset that is especially useful in the solution of problems requiring high spatial and temporal resolution. For this reason we have developed a mobile (kinetic) cellular automata method[8–12] as an elaboration of the particle method[5] and element dynamics.[8,9]

FORMALISM

The material modeled in the given method is represented by an ensemble of discrete elements (cellular automata), which interact with each other by certain rules, relations, and laws. The automata, in turn, comprise elements of heterogeneous media, in particular, individual grains of a polycrystalline material, individual particles of a powder mixture, etc. Of course, the dimensions of the automata depend on the conditions of the specific problem.

Since the individual automata are mobile, this approach can be used to model various processes encountered in a real material, including permeation effects, mass transfer, frag-

mentation effects, and the formation of defects, cracks, and voids. Various mechanical loading regimes (compression, tension, shear deformation, etc.) can be simulated by imposing additional conditions on the boundary of the modeled sample.

The modeled system is characterized by the following quantities in the method of mobile cellular automata: the radius vectors of the centers of the automata $\{\mathbf{R}^i\}$, the translational velocities of the automata $\{\mathbf{V}^i\}$, their angles of rotation $\{\boldsymbol{\theta}^i\}$, and their angular velocities $\{\boldsymbol{\omega}^i\}$. In addition, each cellular automaton is characterized by a dimensional parameter d^i, a mass m^i, and an inertia tensor $\hat{J}^i$. Interaction between the automata is analyzed in the pair approximation and depends on the central forces, the forces of viscous and dry friction, and the forces of resistance to shear deformation. The interacting pairs of automata are divided into two types: bound, when chemical bonds are present (elements of one particle), and independent, when such chemical bonds are not present (elements of different particles).

In the simplest case a cellular automaton contains one type of material — an initial component of a mixture or a reaction product, and in the general case (composite cellular automaton) it contains a set of several different kinds of materials, whose characteristics govern its state. In describing the properties of a composite cellular automaton, we use an analog of the virtual crystal model, which stipulates that all the specific characteristics of the automaton and the parameters of interaction with its neighbors are determined by averaging the constituents of its composition over the number of atoms.

A cellular automaton can change its state both as a result of internal transformations (phase transitions) and in the course of chemical reaction with its neighbors.

The evolution of an ensemble of cellular automata is determined by solving numerically the system of equations of motion

$$m^i \frac{d^2 \mathbf{R}^i}{dt^2} = \sum_j \mathbf{F}^{ij},$$

$$\hat{J}^i \frac{d^2 \boldsymbol{\theta}^i}{dt^2} = \sum_j \mathbf{K}^{ij},$$
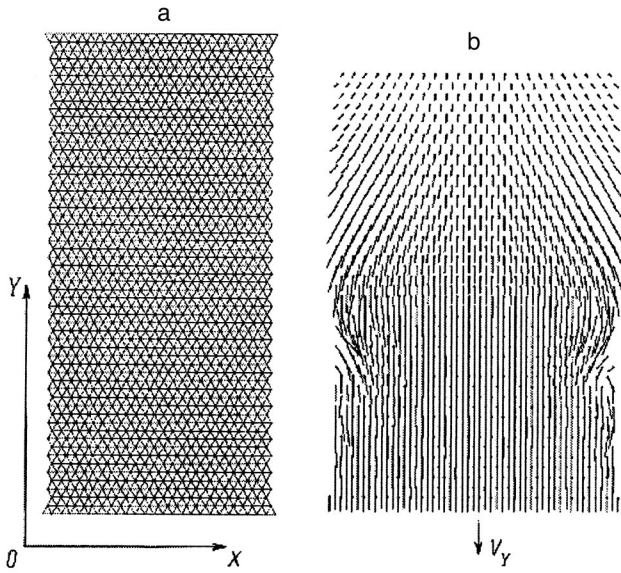
1063-7842/97/091016-03$10.00

FIG. 1. a) Structure of the bonds between automata in the sample model (dimensions: 0.1 mm along the x axis; 0.2 mm along the y axis; b) velocity field at $t = 0.26$ μs for a loading rate $V_y = 10$ cm/s.

FIG. 2. a) Velocity field for the loading scheme compression + shear at $T = 0.022$ μs; b) schematic distribution of the directions of elastic displacements (the arrows indicate the directions of the displacements in different regions of the sample).

where $\mathbf{F}^{ij} = (\mathbf{p}^{ij} + \mathbf{f}^{ij})$, $\mathbf{K}^{ij} = q^i(\mathbf{n}^{ij} \times \mathbf{F}^{ij})$, $\mathbf{p}^{ij}$ describes central interaction, and $\mathbf{f}^{ij}$ is the tangential component of the interaction force.

The unit vector $\mathbf{n}^{ij}$ is defined as $\mathbf{n}^{ij} = (\mathbf{R}^j - \mathbf{R}^j)/r^{ij}$, where r^{ij} is the distance between centers of the automata, and q^i is the distance from the center of the ith automaton to its point of contact with the jth automaton.

The modeled object (a material in our case) can consist of a single species of automata (sample of pure material) or of different species (inclusions, surfacing, etc.). Special algorithms are used to take account of the changes of the surface of a material subjected to fracture.

RESULTS OF CALCULATIONS AND DISCUSSION

The object modeled in the present study to investigate the response of a deformable material at the mesoscopic level is a planar sample, for which Fig. 1a shows a typical structure of the bonds between automata. The scale of one automaton is varied from 4 μm to 400 μm. The parameters used for interaction between automata correspond to pure Ni or Al and are taken from Ref. 13.

We consider the behavior of the material under the influence of a simple loading scheme: uniaxial tension and compression at a constant rate V_y. The load is applied to the lower part of the sample. The calculations are carried out for three loading rates: $V_y = -10$ cm/s, -1 cm/s, and -0.1 cm/s.

It has been shown[11] that in the tension of Ni samples steady-state deformation is established at a time $t = 5$ μs, the deformation relaxation front having begun to move at the time $t = 0.05$ μs. As should be expected, the front is plane for a homogeneous sample. It becomes evident from the subsequent evolution of the velocity field that a zone of compressive stresses of length $y \approx 30$ μm along the y axis must exist behind the front over the surface of the material (Fig. 1b). This compression pulse propagates directly behind the
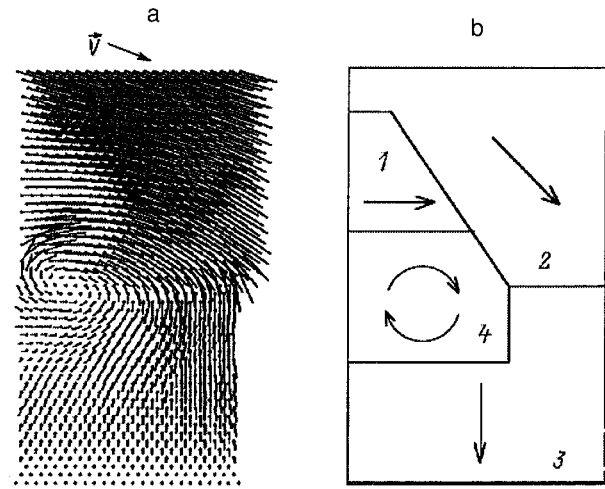
front until steady-state deformation is attained, and it is stable under interaction with surface defects. It is important to note that the internal structure of the pulse changes in interaction with a surface defect. The experimental study of the observed effect should therefore be based on an investigation of the corresponding acoustic-emission spectra.

Further investigation of the effect discovered in Ref. 11 has shown that the existence of such a transverse compression pulse is explained by the tendency of the material to preserve its initial volume. This pulse first appears in the surface layer of the material, where freedom of motion exists in the direction perpendicular to the applied load. The onset of the compression pulse is not observed when periodic boundary conditions are present on the lateral surfaces of the modeled sample. Once it appears, the compression pulse propagates along the surface with the velocity of transverse sound and into the depth of the material with the velocity of longitudinal sound. Calculations have shown that compression pulses propagating into the depth of the material from two opposite boundaries interact and cancel one another. If the surface is hardened, i.e., if the elastic moduli are increased, specifically doubled and quadrupled for surface automata, the velocity of propagation of the compression pulse along the surface of the material increases by 10% and 15%, respectively. An analogous expansion pulse is observed when the samples are compressed.

In reality, excluding specially designed experiments, the applied load is not oriented along the axis of the sample, owing to peculiarities of the loading scheme and the inhomogeneous internal structure of the material, and shear stresses appear as a result. Of major interest in practice, therefore, is the investigation of the characteristics of deformation of a material in the presence of a shear component. Accordingly, the following two loading regimes for an Al sample are modeled in the present study: tension + shear ($V_x = 12.5$ m/s, $V_y = 3.75$ m/s) and compression + shear ($V_x = 12.5$ m/s, $V_y = -3.75$ m/s). The structure of the sample
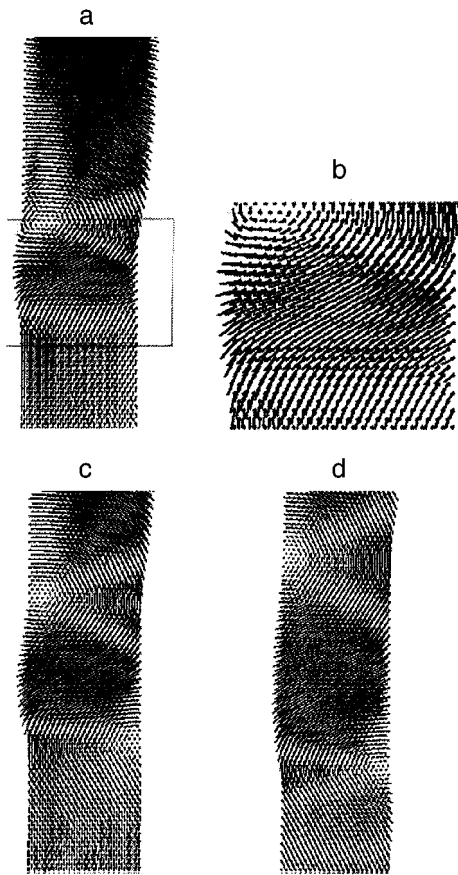
FIG. 3. Velocity field for a long sample at various times. a) $t = 0.065$ μs; b) fragment outlined by the rectangle in part a of the figure; c) $t = 0.095$ μs; d) $t = 0.165$ μs.

ing from the load surface and the other reflected from the lateral surface.

Modeling of a solid under tension shows that the strain also exhibits a vortex character due to the free boundary of the sample, but now the vortex forms at the opposite boundary of the sample from the case of compression.

Our investigation has shown that for a sufficiently long sample such a vortex initiates the formation of a new, similar vortex, but at the opposite boundary and in the opposite direction. The nucleation of the new vortex can be divided into four stages. In the first stage a fairly broad strip of homogeneous displacements with a shear component is distinguishable downstream of the propagating first vortex (Figs. 3a and 3b). In the second stage the core of a new vortex appears below the indicated strip near the lateral surface, opposite in direction to the shear component (Fig. 3c). In the third stage the offspring grows and moves away from the parent vortex (Fig. 3d). In the fourth stage the dimensions and propagation velocities of the offspring and parent vortices equalize. Afterward, if the dimensions of the sample allow, the new vortex generates another offspring by the same mechanism. The distance between the vortices depends on the dimensions and elastic characteristics of the sample. In a sufficiently long sample, therefore, the relaxation to steady-state deformation can be implemented by the propagation of a characteristic wavelike displacement field.

In summary, according to the final results, the stressed state of the material is strongly inhomogeneous early in the deformation relaxation stage; this property, in turn, can cause the velocity field of the material particles to acquire stable structures that propagate near the surface of the material at the transverse sound velocity.

also corresponds to Fig. 1a. The characteristic scale of each automaton is 3 μm.

Preliminary calculations based on the mobile cellular automata model show that the presence of the free surface and a shear component of the load in such a complex loading scheme results in the formation of a stable vortex of elastic strains, which moves in the direction of propagation of the perturbation front. The velocity of the vortex depends on the transverse sound velocity. It has been shown[12] that the strain distribution in the transient load relaxation stage is highly inhomogeneous. Figure 2 shows the velocity field (a) and the schematic distribution of the directions of elastic displacements (b) at a time when vortex formation has already taken place. It is important to note that such an inhomogeneous pattern in the velocity field, as in uniaxial loading, is due to the influence of the free surface of the sample. The properties of the surface affect the shape of the vortex as well as the characteristics of its formation and propagation. For example, in the case of a hardened surface the vortex is more localized and begins to evolve at a greater depth from the load surface. The vortex nucleation depth is also influenced by the width of the sample: The greater the width, the greater is the depth at which the vortex forms. This relationship is most likely attributable to the fact that the vortex forms as the result of the interaction of two load pulses, one propagat-

[1] *Plastic Deformation of Structural Materials*, edited by A. F. Pimenov [in Russian], Nauka, Moscow (1988).

[2] J. F. Bell, *Experimental Foundations of Solid Mechanics* (Vol. 1 of *Mechanics of Solids*, edited by C. Truesdell, Springer-Verlag, New York, 1984) [Russian translations, Nauka, Moscow, 1984; Part I, ''Small Deformations''] .

[3] V. I. Syryamkin, V. E. Panin, E. E. Deryugin et al., in *Physical Mesomechanics and Computer Construction of Materials* [in Russian], Nauka, Novosibirsk (1995), Vol. 1, pp. 176–194.

[4] V. E. Panin, Izv. Vyssh. Uchebn. Zaved. Fiz. **38**, 6 (1995).

[5] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New York, 1981) [Russian translations, Mir, Moscow, 1987].

[6] S. Wolfram, *Theory and Application of Cellular Automata* (World Scientific, Singapore, 1986).

[7] G. Philippou, H. Kim, and R. Rajagopalan, Comput. Mater. Sci. **4**, 181 (1995).

[8] V. E. Panin, V. A. Klimenov, S. G. Psakh'e et al., *New Materials and Technologies: Construction of New Materials and Hardening Technologies* [in Russian], Nauka, Novosibirsk (1993).

[9] S. G. Psakh'e, S. I. Negreskul, K. P. Zol'nikov et al., in *Physical Mesomechanics and Computer Construction of Materials* [in Russian], Nauka, Novosibirsk (1995), Vol. 2, pp. 77–105.

[10] S. G. Psakh'e, Ya. Khori, S. Yu. Korostelev et al., Izv. Vyssh. Uchebn. Zaved. Fiz. **38**, 58 (1995).

[11] S. G. Psakh'e, A. Yu. Smolin, S. Yu. Korostelev et al., Pis'ma Zh. Tekh. Fiz. **21**(20), 72 (1995) [Tech. Phys. Lett. **21**, 849 (1995)].

[12] S. G. Psakh'e, E. V. Shil'ko, A. I. Dmitriev et al., Pis'ma Zh. Tekh. Fiz. **22**(2), 90 (1996) [Tech. Phys. Lett. **22**, 687 (1996)].

[13] *Physical Quantities: Handbook* edited by I. S. Grigor'ev and E. Z. Meĭlikhov, [in Russian], Énergoatomizdat, Moscow (1991).

Translated by James S. Wood

Pinning of planar vortices and magnetic field penetration in a three-dimensional Josephson medium

M. A. Zelikman

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

The behavior of planar (laminar) vortices in a three-dimensional, ordered Josephson medium as a function of the parameter *I*, which is proportional to the critical junction current and the cell size, is investigated with allowance for pinning due to the cellular structure of the medium. The minimum possible distances between two isolated vortices are calculated. A system of vortices formed in a sample in a monotonically increasing external magnetic field is analyzed. The minimum distance from the outermost vortex to the nearest neighbor is proportional to $I^{-1.1}$. For $I \leq 1.3$ each vortex contains a single flux quantum Φ_0, and the distance between them does not decrease in closer proximity to the boundary but remains approximately constant, implying that the magnetic field does not depend on the coordinate in the region penetrated by vortices. These facts contradict the generally accepted Bean model. The sample magnetization curve has a form typical of type II superconductors. Allowance for pinning raises the critical field H_c and induces a sudden jump in the curve at $H = H_c$. © *1997 American Institute of Physics.* [S1063-7842(97)00809-X]

INTRODUCTION

Studies in recent years have demonstrated the important role of vortices in processes occurring in bulk high-temperature (high-T_c) superconductors exposed to moderate magnetic fields. Vortices do not exist in weak fields; the field is expelled from the sample. Conditions become energetically favorable for the onset of vortices at a certain field strength. Without pinning, a uniform vortex lattice is established in the sample, corresponding to uniform field penetration, which becomes denser as the field increases. If pinning is present, the vortices cannot fill up the entire sample all at once; they originate initially near the surface and gradually advance into the depth of the sample as the field increases. The nature of this process is governed by the structure of the vortices, their pinning, and the dependence of the latter on the magnetic field.[1–7]

The penetration of the magnetic field into the sample is usually calculated on the basis of the Bean model,[1] according to which all vortices in the region penetrated by the magnetic field exist in a critical state, i.e., the force exerted on each vortex by all other vortices is equal to the maximum force with which it is pinned to structural defects. A calculation of the critical state leads to the relation[1]

$$\left| \frac{B}{4\pi} \frac{\partial B}{\partial x} \right| = \alpha_m, \tag{1}$$

where α_m is the maximum pinning force.

To calculate $B(x)$ in the critical state, it is necessary to know the dependence of α_m on the magnetic field B (i.e., on the density of vortices). Bean originally postulated a linear dependence $\alpha_m(B)$, which gives a linear dependence of B on the coordinate in the sample interior. Kim *et al.*[2] have experimentally justified the hypothesis that α_m is independent of B, imparting a parabolic profile to $B(x)$. Other forms of the dependence $\alpha_m(B)$ obtained from empirical analyses of

experimental data, for example, $\alpha_m \sim B/(B + B_0)^\beta$ (Refs. 3 and 4), etc., are considered at the present time.

The form of the dependence $\alpha_m(B)$ depends on the physical nature of the pinning. Various mechanisms exist for the pinning of vortices to all possible lattice defects capable of acting as pinning centers. For example, the interaction of a continuous vortex with discretely arrayed pinning centers has been investigated[5,6] for various relations between the vortex dimensions and the distance between their centers. In Refs. 5 and 6 the Josephson medium essentially comprises a superconductor crosscut by a lattice of one-dimensional weak links. The pinning centers are not tied directly to this lattice, but are created by other factors: impurities, discontinuities, etc. Other authors[7–10] have analyzed a Josephson medium having a different kind of structure: a cubic lattice, in which each link contains a single point Josephson junction. In this type of medium a vortex is not described by a continuously distributed phase difference, but by discrete values of the latter at the individual junctions. Pinning is present in this case; it is associated with the cellular structure of the medium and depends on the finite energy required to move the vortex center to an adjacent cell.

The present author[8,9] has analyzed in detail some possible self-sustaining current structures in such a medium and has obtained a system of equations for the quantization of a fluxoid in loops for the cases of screening currents as well as planar and linear vortex structures. The pinning energy of a solitary planar vortex in a three-dimensional sample due to the cellular structure of the Josephson medium has been calculated.[10] However, as mentioned above, the dependence of the pinning force on the vortex density must be known in order to calculate the magnetic field penetrating the medium. In this regard one needs to be aware that the nonlinearity of the system of equations[9] rules out the possibility of analyzing vortex interaction by the superposition principle. In other words, the interaction of a vortex with its neighbors causes it
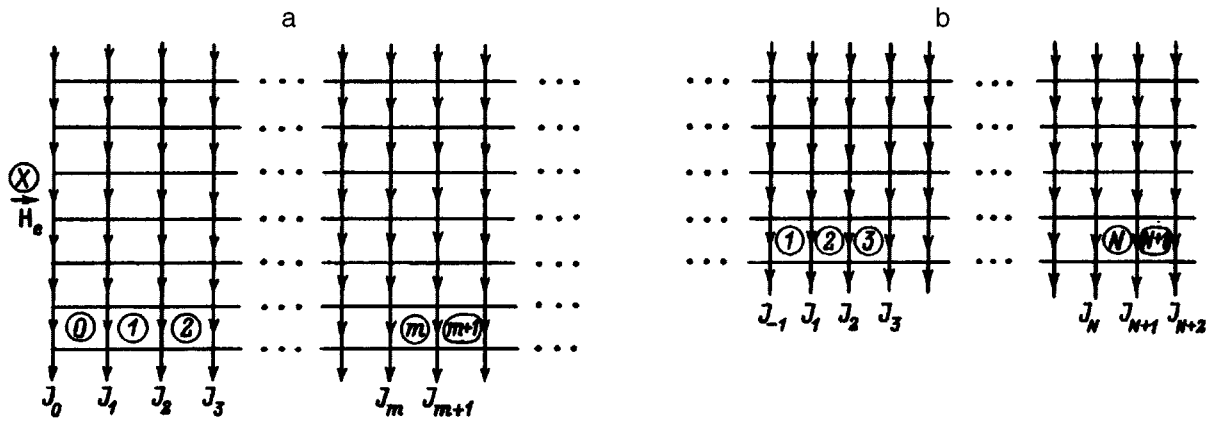
FIG. 1. Distribution of currents in the plane perpendicular to the external field H_e. a) Screening currents near the boundary; b) two interacting vortices far from the boundary.

to change shape, rendering inapplicable the results obtained for a solitary vortex. We illustrate this fact in the following example. Consider a chain of massive spheres connected by springs, which rests on a single sharp peak. The system has two equilibrium stages: 1) a stable state with the peak passing through the midpoint of the interval between adjacent spheres; 2) an unstable state with one of the spheres on the tip of the peak. The ''pinning'' energy in this case can be defined as the difference in the energy of these two states. This is the analytical scheme used in Ref. 10. Let us now assume that the profile of the structure is changed, for example, by elevating the region to the right of the peak. At some time the entire chain begins to move to the left. The energy of the structure is then equal to the pinning energy calculated above. Indeed this is what happens when a vortex interacts with its neighbors.

Consequently, to find the possible current configurations in such a Josephson medium and, on the basis thereof, the profile of the penetrating magnetic field, the entire configuration as a whole must be calculated directly without any reliance on the pinning forces calculated for an isolated vortex. Parodi and Vaccarone[7] have analyzed numerically the possible current distributions and corresponding magnetic field profiles near the boundary of a Josephson sample of the type in question. They have confirmed the validity of the critical-state concept, i.e., the Bean model. However, the discussion in Ref. 7 covers only the case of high critical current in the junctions, when it is impossible to separate individual vortices in such a medium (this problem is discussed in detail below).

Here we investigate the interaction and pinning of planar vortices in a three-dimensional Josephson medium of the type discussed in Refs. 7–10 for a low critical current in the junctions. We calculate the possible distances between vortices for the case of two solitary vortices and also for a system of vortices formed in the medium when it is placed in an external magnetic field. On the basis of these results we analyze the profile of the magnetic field penetrating the sample. We show that Bean's concept, stipulating that all vortices reside at the surface of motion, is invalid in this range of currents I. The results are also applicable in the one-dimensional case of long Josephson junctions.

EQUATIONS FOR THE QUANTIZATION OF A FLUXOID IN CELLS

As in Refs. 8 and 9, we work with the simplified model of a cubic lattice having a lattice constant h and consisting of superconducting wires, where each link of the lattice contains one Josephson junction, and all the junctions have the same critical current J_c. The current distribution has a planar structure, i.e., the current is identically distributed in all parallel planes perpendicular to the external magnetic field or the vortex axis and separated by a distance h.

Let a sample in the form of a thick plate of infinite extent in two dimensions be placed in an external magnetic field H_e parallel to the plane of the plate. In weak fields H_e we have the Meissner effect: Screening currents appear in the sample, running along its surface and closing at infinity. When the magnetic field H_e exceeds a certain threshold, vortices begin to appear in the sample. In the ensuing discussion we consider planar (laminar) vortex structures, even though conditions are more favorable for the onset of linear vortices, which are also the kind encountered in practice. The case of planar vortices, on the other hand, is analytical and therefore offers insight into all mathematical and physical nuances; at the same time, the results can be extended qualitatively to the case of linear vortices. In the article we devote special attention to the validity of the Bean model.

Figure 1a shows the cross section of the sample in a plane perpendicular to the field H_e. In the case of the screening currents (Meissner case) the currents decay into the depth of the sample and are equal to zero in its interior. For a solitary planar vortex far from the surface the currents decay and tend to zero with increasing distance from the middle row.[8] In general, there are also surface screening currents and a system of vortices which is periodic without pinning or thins out with increasing depth into the sample in the presence of pinning.

The condition for the quantization of a fluxoid in the mth cell (the numbers of the cells are circled in Fig. 1a) has been derived previously[8] and is represented by Eq. (8) in Ref. 8:

$$x_e - I \sum_{i=0}^{m} \sin \varphi_i - (bI \sin \varphi_m + \varphi_m)$$

$$+ (bI \sin \varphi_{m+1} + \varphi_{m+1}) = 2\pi K_m, \qquad (2)$$

where $x_e \equiv 2\pi\Phi_e/\Phi_0$ is the flux, normalized to the flux quantum Φ_0, of the external magnetic field $\Phi_e = \mu_0 H_e h^2$ through the cell, φ_i is the phase difference at the ith junction,

$$I \equiv 2\pi\mu_0 h J_c/\Phi_0, \quad b \equiv -\frac{1}{2\pi} \ln\left(2 \sinh \frac{\pi\delta}{h}\right), \qquad (3)$$

δ is the radius of the wire, and K_m is the number of quanta Φ_0 in the mth cell.

For simplicity we shall assume below that $b=0$. Eq. (2) then assumes the form

$$\varphi_{m+1} - \varphi_m = I \sum_{i=0}^{m} \sin \varphi_i - x_e + 2\pi K_m. \qquad (4)$$

The number $K_m=1$ in cells of the central column of each vortex, and $K_m=0$ in all other columns.

Subtracting from (4) the analogous equation for the $(m-1)$st cell, we obtain

$$\varphi_{m+1} - 2\varphi_m + \varphi_{m-1} = I \sin \varphi_m (\pm 2\pi), \qquad (5)$$

where the term $+2\pi$ or -2π can appear on the right side of the equation if the value of n corresponds to the center of one of the vortices.

It might appear that the distribution of φ_m for a given x_e (i.e., for a given external magnetic field) could be found by solving the system of Eq. (5) with the boundary condition obtained form (4) at $m=0$:

$$\varphi_1 - \varphi_0 = I \sin \varphi_0 - x_e. \qquad (6)$$

However, this is not the case. The existence of pinning means that a given x_e can correspond to an uncountable set of distinct combinations of φ_m. In other words, the profile of the magnetic field in the sample with pinning depends on the prior history. We are interested in the case of a monotonically increasing external field. The vortices gradually penetrate the interior of the sample, progressing from the surface. The corresponding distribution of the vortices in the sample can be determined from the condition that the vortex farthest from the boundary resides at the surface of motion.

As the parameter I increases, the vortex decreases in size,[8,9] and the pinning force increases (as is readily understood in light of the fact that $I \sim J_c$). Consequently, as I increases, the distance between vortices decreases until, for large I, the centers of vortices can be situated in adjacent cells, the concept of individual vortices becomes meaningless.

INTERACTION AND PINNING OF TWO ISOLATED VORTICES

We consider two planar vortices situated far from the boundary, their centers separated by N cells (Fig. 1b), in the rows numbered 1 and $(N+1)$ (the numbers are circles in Fig. 1b). For $I \geq 1$ (Ref. 10) the values of φ_m outside the central cells are small, the system (5) is linearized, and its solution has the form

$$\varphi_m = \varphi_{-1}\gamma^{-m-1} \quad (m \leq -1), \qquad (7a)$$

$$\varphi_m = C_1\gamma^{m-1} + C_2\gamma^{N-m} \quad (1 \leq m \leq N), \qquad (7b)$$

$$\varphi_m = \varphi_{N+1}\gamma^{m-N-1} \quad (m \geq N+1), \qquad (7c)$$

where

$$\gamma = 1 + \frac{I}{2} - \sqrt{I + \frac{I^2}{4}}$$

is the solution of the equation $\gamma^2 - (2+I)\gamma + 1 = 0$.

We assume that the distribution of the currents and the phases is symmetric about the midpoint of the distance between vortices. We then have $C_1 = -C_2 = \varphi_1/(1-\varphi^{N-1})$, and

$$\varphi_2 = \varphi_1 k, \qquad (8)$$

where the coefficient $k = \gamma(1-\gamma^{N-3})/(1-\gamma^{N-1})$.

Substituting Eq. (8) into the boundary conditions on the central cell of the vortex:

$$I \sin \varphi_1 = \varphi_{-1} - 2\varphi_1 + \varphi_2 + 2\pi, \qquad (9a)$$

$$I \sin \varphi_{-1} = \varphi_1 - 2\varphi_{-1} + \varphi_{-2} - 2\pi, \qquad (9b)$$

we obtain the system of equations for φ_{-1} and φ_1:

$$\varphi_{-1} = I \sin \varphi_1 + (2-k)\varphi_1 - 2\pi, \qquad (10a)$$

$$\varphi_1 = I \sin \varphi_{-1} + (2-\gamma)\varphi_{-1} + 2\pi. \qquad (10b)$$

In Eq. (10) only the coefficient k depends on the spacing of the vortices: $k(\infty) = \gamma$, $k(2) = -1$, $k(3) = 0$, etc. Figure 2 shows graphs of the functions (10) for $I = 1, 2.85, 4$, from which it is evident that for every value of I the system (10) has solutions in a definite interval of N, i.e., vortices can exist at different distances from each other, ranging from a certain minimum to infinity. This is the result of pinning, because without it the vortices would blow up at infinity, i.e., solutions would not exist for finite N. Having found N, we can determine the minimum distance for which the pinning forces are able to cancel out the mutual repulsion of vortices. It is evident from Fig. 2 that $N_{\min} = 2$ for $I = 2$, and that for $I = 4$ the centers of vortices can even be located in neighboring cells.

We now find the minimum value of I for which the vortex centers can be situated in neighboring cells. In this case $N = 1$, $k = \infty$, and it follows from (10a) that φ_1 must be equal to zero (as is readily perceived from the symmetry of the pattern). The minimum value of I satisfying Eqs. (10) for $\varphi_1 = 0$ is represented by the curve tangent to the φ_{-1} axis in Fig. 2. An exact numerical calculation gives $I_{\min} = 2.9$.

EQUILIBRIUM OF THE OUTERMOST VORTEX

Here we determine the values of I for which the concept of vortices is well defined. In the preceding section we have shown that the vortex centers cannot be situated in neighboring cells for $I < 2.9$. However, this conclusion applies only to the case of two solitary vortices, because in this case they are, in effect, located at the most distant possible sites in the corresponding cells. But if other vortices are placed to the right of the right vortex in the original pair, allowing it to retreat farther to the right in the same cell, the left vortex can
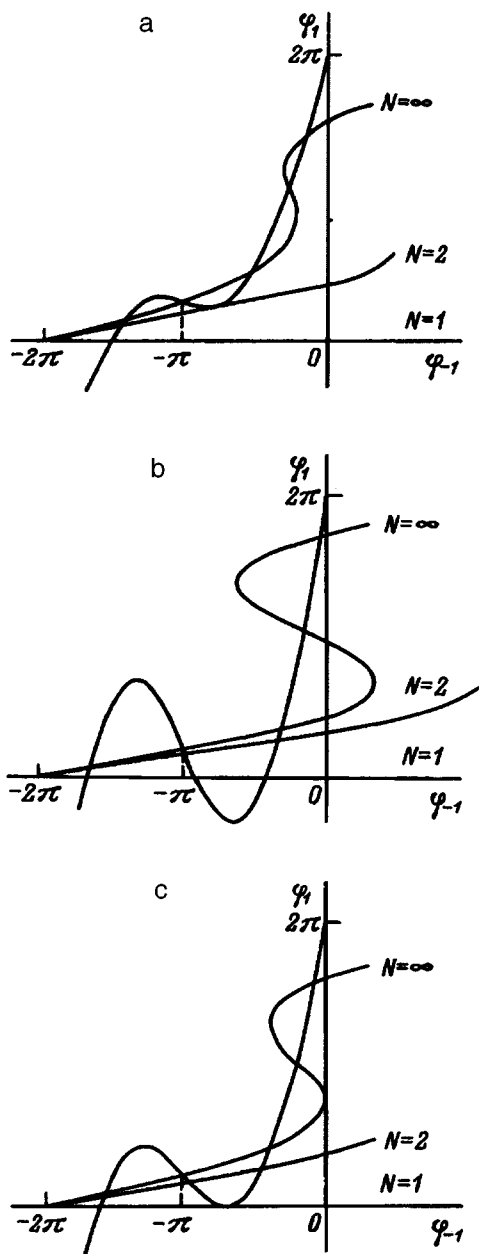
FIG. 2. Graphical solution of the system (10) for various values of the parameter I. The curves emanating from the point $\varphi_1 = 2\pi$ correspond to Eq. (10b), and those from the point $\varphi_{-1} = -2\pi$ correspond to Eq. (10a). a) $I=2$; b) $I=4$; c) $I=2.85$.

then remain at a distance shorter than that deduced from the assumption of two isolated vortices, and the actual value of $N_{\min}$ becomes lower. In other words, the value of I corresponding to a given $N_{\min}$ is smaller than the value obtained for two vortices; in particular, carrying out a numerical calculation by the procedure described below, for $N_{\min} = 1$ we obtain $I_{\min} = 2.7$.

In this section we consider values of I for which the outermost vortex is well defined, i.e., $I < 2.7$.

The system of Eq. (5) can be regarded as a recursion law for determining the next value of φ_{m+1} (or φ_{m-1} from the known φ_m and φ_{m-1} (or φ_{m+1}) in transition to the left (or to the right). The problem of including $\pm 2\pi$ terms in (5) can be eliminated by requiring that $|\varphi_m| \leq \pi$. In this case, if the

new value $\varphi_{m+1}(\varphi_{m-1}) < \pi$, then 2π must be added to it, but if $\varphi_{m+1}(\varphi_{m-1}) > \pi$, then the same quantity must be subtracted. This generalized recursion law can be used to calculate the distribution of φ_m throughout the entire region once the values of φ have been specified at two adjacent points. Various vortex distributions in the sample and the corresponding magnetic fields can be found by varying the specified values of φ.

The situation in question, where the outermost vortex is situated at the surface of motion, corresponds to a pattern in which the neighbor to the right of a certain vortex is separated from it by the maximum possible distance (at infinity in the ideal situation) and the neighbor to the left is situated at the minimum possible distance.

The postulated recursion law has been implemented numerically on a computer. The two initial values chosen for φ are the values φ_- and φ_+ at the right and left boundaries of the central cell of the selected initial vortex. Its right and left neighbors must have the same orientation as itself, i.e., repulsion in both directions takes place. The following specific computational algorithm is used: For a fixed value of φ_+ the value of φ_- is decreased monotonically until the instant at which the opposite to the initial orientation is first acquired to the right of the initial vortex. This value is successively refined, the right and left neighbors retreating monotonically from the initial vortex. The distance from the right neighbor gradually exceeds the distance from the left neighbor and, beginning at a certain time, the distance from the left neighbor no longer changes, whereas the right neighbor moves still farther away. This value of d is recorded. The same procedure is then repeated for a different value of φ_+. The minimum distance $d_{\min}$ is chosen from the total set of successive calculated distances d; the corresponding distribution of φ_m is then the solution of the stated problem.

To find the starting values φ_+ and φ_-, we investigate the behavior of one vortex as another approaches it. It has been shown[10] that for values of $I \leq 1$ the shape and energy of a solitary vortex are satisfactorily described by expressions obtained in the approximation $I \ll 1$, when the discrete set φ_m goes over to a continuous function $\varphi(z)$, where z is the coordinate normalized to the cell size h. Of course, this approach cannot be used to find the pinning forces, because they are a product of the discreteness of the medium, but it suffices for a qualitative analysis of the changes in shape of the vortex as the other one converges toward it, as long as the distance between them remains large in comparison with the size of the vortex itself. The system[5] now goes over to the differential equation

$$\frac{d^2\varphi}{dz^2} = I \sin \varphi. \tag{11}$$

The solution of Eq. (11) for a solitary vortex is shown in Fig. 3 (solid curves). As another vortex of like orientation approaches from the left, the given vortex begins to move to the right (dashed curves in Fig. 3). From condition[5] on the central cell of the vortex we deduce the relation

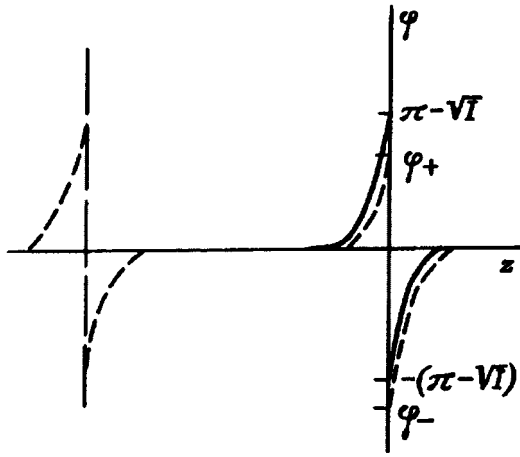$$\varphi_+ - \varphi_- \approx 2\pi - \varphi_0' \sqrt{I}, \tag{12}$$

FIG. 3. Solution of Eq. (11) for a solitary vortex (solid curve) and its displacement as a neighboring vortex approaches from the left (dashed curves).

FIG. 4. Graphs of the distributions of φ_m. a) $\varphi_- = -2.96$; b) -2.95; c) -2.9524; d) -2.9523; e) $-2.952\,358\,812$; f) $-2.952\,358\,811$.

where φ_0' is the value of the derivative φ_z' at $z=0$.

For a solitary vortex we have $\varphi_0'=2$. Since the second vortex is far from the first, we can assume that φ_0' differs very little from 2. Condition (12) therefore assumes the form

$$\varphi_+ - \varphi_- = 2\pi - 2\sqrt{I}. \tag{13}$$

The solitary vortex has three equilibrium states[10]: 1) a stable state with $\varphi_+^s = -\varphi_-^s = \pi - \sqrt{I}$ (solid curve in Fig. 3); 2) a right unstable state with $\varphi_-^u = -\pi$ and $\varphi_+^u = \pi - 2\sqrt{I}$; 3) a left unstable state with $\varphi_-^u = \pi$ and $\varphi_+^u = -\pi + 2\sqrt{I}$.

Consequently, the values $\varphi_+^s = \pi - \sqrt{I}$ and $\varphi_+^u = \pi - 2\sqrt{I}$ corresponding to an infinite distance from the neighboring vortex. The value of φ_+ lies between these two extremes for finite distances between the vortices. It is also necessary to choose from this interval the values of φ_+ for numerical calculations. The corresponding starting values of φ_- are chosen as follows on the basis of (13):

$$\varphi_- = \varphi_+ + 2\sqrt{I} - 2\pi. \tag{14}$$

Figure 4 shows graphs of the distribution of φ_m to clarify the above-described algorithm. Calculations have been carried out for $I=0.3$ and $\varphi_+=2.24$; the initial value $\varphi_-=-2.95$. For $\varphi_-=-2.95$ the closest vortices to the left and to the right of the central vortex have the same orientation as the latter. For $\varphi_-=-2.96$ (a) the left and right vortices become oppositely oriented, as is evident from the corresponding kinks of the curve at the ends (the calculations are terminated when such kinks occur). For $\varphi_-=-2.95$ (b) the left and right vortices acquire the necessary orientation. The unknown value of φ_- therefore lies in the interval $(-2.96, -2.95)$. For $\varphi_-=-2.9524$ (c) the right vortex has the opposite orientation, and for $\varphi_-=-2.9523$ (d) it now has the required orientation. The situation is analogous for graphs e and f. It is evident from Fig. 4 that as the value of φ_- is refined, the vortices move away from the central vortex to the left and to the right (graphs b and d), then the left vortex stops, while the right vortex continues to move away (graphs d and f). In graph f the right vortex is so much farther from the central vortex than the left one that it can be
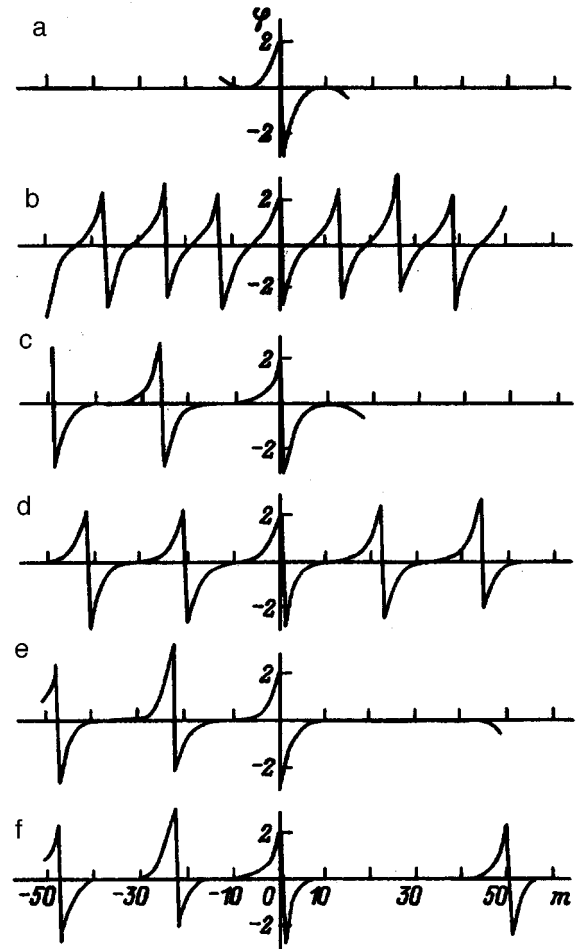
assumed to no longer influence the central one; this assumption is corroborated by the invariance of the pattern to the left of center as the orientation of the right vortex changes, i.e., as attraction is replaced by repulsion (graphs e and f). As a result, for $\varphi_+=2.24$ the distance between the left and central vortices is equal to 23 cells.

It is evident from the results of Fig. 4 that the accuracy of the calculations in solving the stated problem must be very high, in some cases extending to the 15th decimal place.

Figure 5 shows the dependence of the distance d (measured in cells) on the value of φ_+ for $I=0.2$ (the pattern is similar for other values of I). The minimum possible distance $d_{\min}$ is established in a broad interval of φ_+, so that $d_{\min}$ can be found by choosing, for example, $\varphi_+ = \pi - 1.5\sqrt{I}$, i.e., its value at the midpoint of the interval.

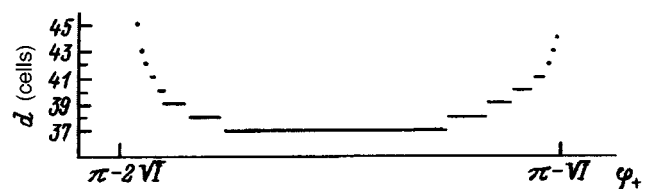

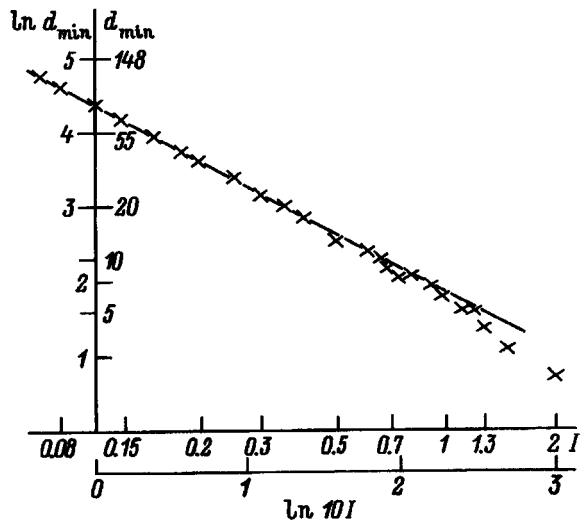FIG. 5. Distance d between the outermost vortex and its neighbor vs φ_+.

FIG. 6. Minimum vortex spacing $d_{\min}$ (in units of cells) versus I. $\times$) Calculated values.

FIG. 7. a) Results of calculations of the distances Δ from the $(n-1)$st to the nth vortex in the critical state for various values of I; b) results of averaging Δ over five successive vortices. The values of I are indicated alongside the curves.

We note that in calculations accurate to the 15th decimal place the right vortex is situated 40–50 cells from the central vortex for values of I in the interval 0.5–1.3, 70 cells for $I = 0.3$, 90 cells for $I = 0.2$, and 120 cells for $I = 0.1$. We can therefore assume that no vortices exist to the right of the central one.

Figure 6 shows the dependence of $d_{\min}$ on I in logarithmic scale. From the linearity of this dependence we deduce the power-law relation

$$d_{\min} = 6.1 \cdot I^{-1.1}, \tag{15}$$

which is obeyed up to $I \approx 1$.

"CRITICAL" STATE OF A SYSTEM OF VORTICES UPON A MONOTONIC INCREASE OF THE EXTERNAL FIELD

The Meissner effect occurs in weak external fields, the field is expelled from the sample as a result of surface currents, and vortices do not exist. At a certain field strength conditions become energetically favorable for vortex generation, and the first vortex forms near the boundary. Pinning keeps it from advancing into the interior of the sample. With a further increase in the field the next vortex is formed near the boundary, driving the first vortex to the minimum possible distance $d_{\min}$. Next comes a third vortex, causing the first two to advance farther inward to corresponding distances, etc. According to the Bean model, in the critical state all vortices are situated at the surface of motion, implying a reduction in the distance between vortices (i.e., an increase in the mean field B) in closer proximity to the boundary. Finally, a situation arises where the forces of repulsion of the vortices, even those located in neighboring cells, is not sufficient to move a vortex, being overpowered by repulsion from vortices on the other side. When the external field is increased, another flux quantum Φ_0 penetrates the vortex closest to the boundary, causing the repulsive forces to increase until they are sufficient to move the preceding vortex. With a further increase in the field and advancement of the entire vortex lattice deeper into the interior, the new vortices
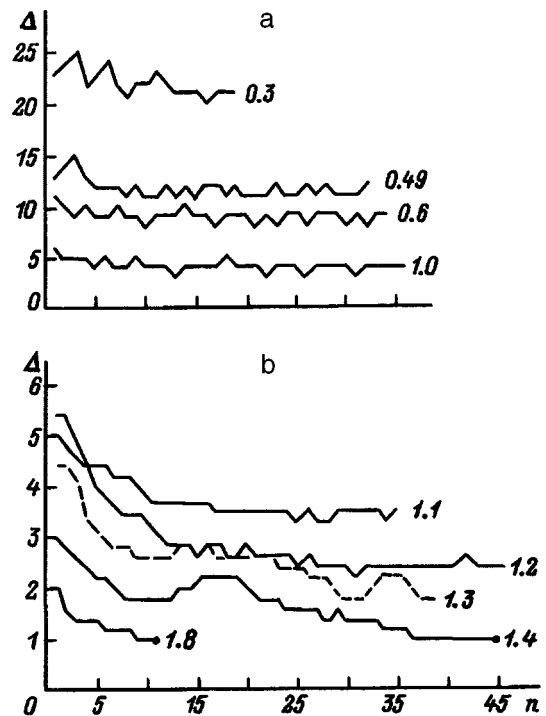
near the boundary contain an ever-increasing number of quanta Φ_0; as a result, the effective magnetic field increases as the boundary of the sample is approached. This is how the situation looks within the context of the Bean model. In this section, however, we show that the above-described model is invalid for $I < 1.3$.

We consider the critical state of the vortex lattice, when the farthest vortex from the boundary is situated at the surface of transition to the next cell. It is readily perceived that this is the vortex from which the total vortex lattice must begin to move inward.

In the preceding section we have determined the interval of φ_+ corresponding to the shortest distance $d_{\min}$ to the next vortex. We minimize the distance to the next vortex by varying φ_+. Repeating the same procedure, each time narrowing the interval of φ_+, we can find the required critical state of the vortex lattice as the external field H_e is monotonically increased. Calculations with accuracy to the 15th decimal place enable us to find the positions of 20–50 vortices (depending on I).

Figure 7a gives the results of a computer calculation of the distances Δ (in cells) from the $(n-1)$st to the nth vortex (the outermost vortex is numbered 0 and is followed by number 1, etc.) for the most tightly bunched configuration against the boundary and various values of $I \leq 1$. For $I > 1$ the distances Δ are equal to several cells ($\Delta = 2–3$ for $I = 1.2$), and fluctuations of Δ in the first couple of cells obscure the functional relation in Fig. 7a. Figure 7b shows the results of averaging Δ over five adjacent vortices:

$$\bar{\Delta}(n)=0.2\sum_{i=0}^{4}\Delta(n+i). \tag{16}$$

The following conclusions can be drawn from the graphs in Fig. 7.

1) For $I \geqslant 1.4$ the average distance between vortices $\bar{\Delta}$ decreases monotonically to unity as n increases. The vortex centers are then situated in neighboring cells. The number of flux quanta Φ_0 in the vortex begins to increase upon closer approach to the boundary of the sample. The dots in Fig. 7b indicate the positions of the center of the vortex in which the number of flux quanta Φ_0 is two. It follows, therefore, that the proposed approach based on the concept of individual, interacting vortices is invalid for $I \geqslant 1.4$. This case requires special analysis.

2) For $I \leqslant 1.3$ it follows from Figs. 7b and 7a that the average distance $\bar{\Delta}$ does not decrease to unity. As it changes for small n, i.e., near the boundary of the structure, $\bar{\Delta}$ reaches saturation, remaining approximately constant and fluctuating slightly about the average. In this interval of I vortices containing more than one quantum Φ_0 do not form as the external field is monotonically increased. Since the average distance between vortices $\bar{\Delta}$ does not depend on the distance from the boundary, the magnetic induction inside the sample, in the region penetrated by vortices, is also independent of the coordinate. This fact, which contradicts the Bean model, is true because not all the vortices reside immediately at the surface of motion as postulated in the Bean model. When the outermost vortex moves into the next cell, its neighbor remains fixed in the same place, and in order for it to move, the entire vortex lattice must shift slightly inward; the same is true of the next vortex, etc. This situation can be clarified in the following example: Let a chain of massive spheres lie in a system of periodically spaced indentations and be interconnected by slightly compressed springs. The chain can be moved by the application of a force along the surface. When the end sphere shifts into the next indentation, its neighbor stays in place and shifts afterward; the next one after that moves even later, etc. The chain does not advance as a unit whole, but in successive advances of the individual spheres. Consequently, for the chain to move it is necessary to overcome the pinning force of one or more spheres, not all the spheres in the chain. It is readily perceived that appreciable differences in the degree of compression can occur only in a few springs closest to the end. With increasing distance from the end the situation tends to average out, and the compression of the springs does not increase monotonically toward the boundary of the sample as it would if the spheres all moved simultaneously.

The magnetic flux through the area $h \times h\Delta$ corresponds to a single vortex and is equal to Φ_0; the average magnetic induction can be found by dividing this flux by the area of the vortex:

$$B=\Phi_0/h^2\Delta. \tag{17}$$

We now determine the external field H_e for which vortices begin to form in the sample. The Gibbs thermodynamic potential of unit volume is
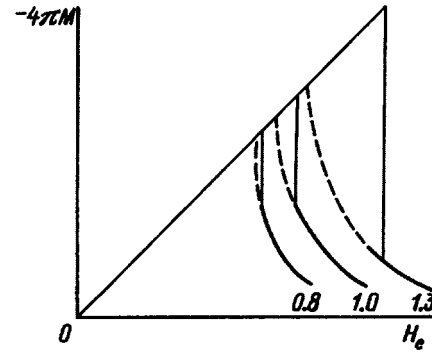

FIG. 8. Sample magnetization curves based on data in Table I.

$$G=N_bE_0/h-BH_e/2=0.5B(H_c-H_e), \tag{18}$$

where N_b is the number of planar vortices per meter, i.e., $B=\Phi_0/S=N_b\Phi_0/h$, E_0 is the average energy of one vortex, and

$$H_c=2E_0/\Phi_0. \tag{19}$$

For $H_e<H_c$ the potential G increases as B increases, i.e., G has a minimum at $B=0$, and we have the complete Meissner effect. If $H_e>H_c$, then G decreases as B increases, i.e., vortex generation is favorable. In this case the vortices fill up the entire cross section of the sample all at once. The density of the vortices depends on the external field and the interaction between them. It can be calculated with the energies of interaction between vortices included in the Gibbs potential (see, e.g., Ref. 11). The magnetization curve of the sample can be plotted as a result of the calculations. It has the typical form for type II superconductors (Fig. 8). Allowance for pinning raises H_c and creates an abrupt drop in the curve, since the period of the vortex lattice in the sample cannot exceed the period corresponding to the critical state (Fig. 7). At $H_e=H_c$ the field in the sample jumps abruptly from zero to $H_0/\bar{\Delta}$, where $\bar{\Delta}$ is the average period of the lattice, and $H_0\equiv\Phi_0/\mu_0h^2$ is the external field at which the flux through each $h\times h$ cell is Φ_0. For $I<0.5$ we can assume that $\bar{\Delta}$ corresponds approximately to $d_{\min}$ from (15), and the energy per meter of height and length of the vortex is approximately equal to[10]

$$E_0=8\sqrt{I}\varepsilon_0=2\Phi_0^2\sqrt{I}/\pi^2\mu_0h^2. \tag{20}$$

The quantity $-4\pi M=H_e-B$ decreases abruptly from $4\sqrt{I}H_0/\pi^2$ to $(4\sqrt{I}/\pi^2-I^{1.1}/6.1)H_0$. For $I\leqslant 1$ the distance between vortices can exceed the vortex size, so that the vortex interaction energy is small in comparison with the self-energy of the isolated vortex. Consequently, H_c scarcely changes from the nonpinning case, and the jump of $4\pi M$ is small, i.e., pinning has almost no influence on the form of the magnetization curve. For $I\gtrsim 1$ the magnetization curve exhibits significant changes.

Table I gives the average distances between vortices $\bar{\Delta}$, the numerically calculated average energies per vortex E_0, and the solitary-vortex energies $E_{0\,\text{sol}}$ (Ref. 10), along with

TABLE I.

I	$\bar{\Delta}$	E_0	$E_{0\,\mathrm{sol}}$	H_c	$\Delta(4\pi M)$
0.8	6	$7.16\varepsilon_0$	$7.09\varepsilon_0$	$0.36H_0$	$0.17H_0$
1.0	4	$8.60\varepsilon_0$	$7.80\varepsilon_0$	$0.44H_0$	$0.25H_0$
1.3	2	$12.1\varepsilon_0$	$8.90\varepsilon_0$	$0.61H_0$	$0.50H_0$

the critical fields $H_c=2E_0/\Phi_0=E_0H_0/2\pi^2\varepsilon_0$ and the jump $\Delta(4\pi M)=H_0/\bar{\Delta}$ with pinning taken into account for $I=0.8$, 1, 1.3.

Figure 8 shows magnetization curves plotted on the basis of data in Table I. The dashed curves correspond to zero pinning for the chosen values of I (shown alongside the curves).

CONCLUSION

We have investigated the behavior of planar (laminar) vortices in a three-dimensional, ordered Josephson medium comprising a cubic lattice, each link of which contains one Josephson junction. The cellular character of the medium leads to pinning, which depends on the energy required to move the vortex center into the next cell.

1. We have shown that the distance between two isolated vortices, at rest and identically oriented, can vary from infinity to a minimum value. This fact is a consequence of pinning, because without it interaction between the vortices would scatter them to infinity. As the parameter I increases (in proportion to the critical current of the junctions and the cell size), the degree of pinning increases, and the repulsion between vortices diminishes, causing the minimum distance between vortices to decrease. For $I>2.9$ the vortex centers can even be situated in adjacent cells.

2. To plot the magnetization curve of a sample, we have investigated a system of vortices formed in the sample in a monotonically increasing external magnetic field. The minimum distance from the vortex farthest from the boundary to

its nearest neighbor exhibits a power-law dependence on I: $d_{\min}\approx 6.1\cdot I^{-1.1}$.

3. For $I\geqslant 1.4$ the average distance between vortices $\bar{\Delta}$ decreases to one cell as the boundary is approached, and then the vortex centers are located in neighboring cells. Upon closer approach to the boundary the number of flux quanta Φ_0 in the vortices begins to increase. The whole concept of individual interacting vortices breaks down for these values of I. This case requires special consideration.

4. For $I\leqslant 1.3$ the average distance $\bar{\Delta}$ does not decrease toward the boundary, but remains approximately constant, so that the magnetic field is independent of the coordinate in the region penetrated by vortices.

5. For $I\leqslant 1.3$ the magnetization curve of the sample has the form typical of type II superconductors. Allowance for pinning raises the critical field H_c and creates an abrupt change in the curve, owing to the fact that the period of the vortex lattice cannot exceed the period corresponding to the calculated critical state.

We can conclude from these facts that the Bean model is invalid in the case of interacting vortices containing at most one flux quantum Φ_0 for pinning associated with the cellular structure of the medium.

[1] C. P. Bean, Rev. Mod. Phys. **36**, 31 (1969).
[2] Y. B. Kim and P. W. Anderson, Rev. Mod. Phys. **36**, 39 (1964).
[3] K. H. Muller, J. C. Macfarlane, and R. Driver, Rev. Mod. Phys. **158**, 69 (1989).
[4] Q. H. Lam, Y. Kim, and C. D. Jeffries, Phys. Rev. B **42**, 4848 (1990).
[5] Y. S. Kivshar and B. A. Malomed, Rev. Mod. Phys. **61**, 763 (1989).
[6] V. V. Bryksin and S. N. Dorogovtsev, Zh. Éksp. Teor. Fiz. **102**, 1025 (1992) [Sov. Phys. JETP **75**, 558 (1992)].
[7] F. Parodi and R. Vaccarone, Physica C **173**, 56 (1991).
[8] M. A. Zelikman, Sverkhprovodimost' (KIAE) **5**, 60 (1992).
[9] M. A. Zelikman, Sverkhprovodimost' (KIAE) **5**, 1819 (1992).
[10] M. A. Zelikman, Sverkhprovodimost' (KIAE) **7**, 946 (1994).
[11] I. O. Kulik and I. K. Yanson, *Josephson Effect in Superconducting Tunnel Structures* [in Russian], Nauka, Moscow (1970), p. 270.

Translated by James S. Wood

Self-excited oscillatory regimes in the growth of thin films from a multicomponent vapor: dynamics and control

P. Yu. Guzenko, S. A. Kukushkin, A. V. Osipov, and A. L. Fradkov

*Institute of Problems in Mechanical Engineering, Russian Academy of Sciences,
199178 St. Petersburg, Russia*

A model system describing the nucleation of films from a multicomponent vapor with allowance
for chemical reactions between different components in the initial phase is investigated in
detail. It is shown that the condensation of thin films can proceed by different avenues, depending
on the values of external parameters such as the temperature or the precipitation rate of
particles of the component that limits the chemical reaction. In particular, low precipitation rates
are characterized by a stable condensation regime, in which any deviations from equilibrium
die out. At medium precipitation rates the phase transition takes place in a self-excited oscillatory
regime corresponding to a stable limit cycle. Finally, at high precipitation rates the stable
limit cycle breaks up, and the new phase usually condenses in a sawtooth (accretion) regime. A
procedure is developed for controlling the given oscillatory processing by judicious time
variation of the external parameters. The investigated system is found to have a special kind of
memory in that for external parameters with identical values but different histories the
films condense differently; even a slight difference in the past behavior of the external parameters
can lead to different precipitation regimes. It is concluded that these memory effects are in
fact responsible for the poor reproducibility encountered in some cases of experiments on film
growth utilizing chemical reactions. © *1997 American Institute of Physics.*
[S1063-7842(97)00909-4]

INTRODUCTION

Research on the nucleation and growth of new phase on
the surfaces of solids has attracted considerable interest for
some time now.[1,2] This is because, first, thin films are widely
used in microelectronics, optics, etc., and are the basis of
many technologies; second, the processes in question are
typical of many first-order phase transitions. A characteristic
feature of such transitions is the presence of various nonlin-
ear relationships, which generate a host of nonlinear
phenomena[3–6]: self-organization; the generation of solitons,
kinks, and shock waves; the self-similarity of many param-
eters; the growth of instabilities; self-excited oscillations in
multicomponent systems. In particular, the onset of self-
oscillation is attributable to the nonlinear interaction of
chemical reaction in the primary phase with phase transition
on the part of the reaction product. Indeed, on the one hand,
chemical reaction forces material into the primary phase,
thereby accelerating phase transition; on the other hand, the
new phase consumes the reaction product, which is a catalyst
and therefore slows down the chemical reaction. This situa-
tion is characteristic of many thin-film growth techniques
utilizing chemical reactions, in particular, metal-organic
chemical vapor deposition (MOCVD). The present article is
devoted to an investigation of the properties of such self-
excited oscillations and the development of a procedure for
controlling them.

MODEL SYSTEM AND ITS BASIC PROPERTIES FOR CONSTANT EXTERNAL PARAMETERS

We consider a chemical reaction of the type $A + B \leftrightharpoons C$.
We assume that the concentrations of substances A and B is
sufficiently low for their mixture to condense and (the more
so) for A and B to condense separately, whereas the reaction
product C forms with a concentration higher than the equi-
librium value C_e, so that the reaction product undergoes a
first-order phase transition.[4] If C does not form solid solu-
tions with A and B, a film of substance C with stoichiometric
composition grows. We assume for definiteness that sub-
stance B is so abundant on the substrate as to limit the
chemical reaction only of substance A. Let A and C denote
the concentrations of the corresponding substances; also, φ is
the chemical reaction rate, $\Psi(C-C_e)$ is the rate of forma-
tion of islands of the new phase, N is the number density of
such islands, and $\Psi(N,C)$ is the rate of decay of the reaction
product C into new-phase islands. Then in the simplest
model the kinetics of chemical reaction and phase transition
is described by the system of equations

$$dA/d\tau = J_0 - \varphi(A,C,N),$$
$$dC/d\tau = \varphi(A,C) - \Phi(N,C),$$
$$dN/d\tau = \Psi(C-C_e). \tag{1}$$

Here τ is the time, and J_0 is the translational velocity of
substance A onto the substrate. In the most commonly en-
countered diffusion growth regime all the new-phase islands
consume the same number of C molecules, i.e., $\Phi = \gamma NC$,
where γ is a proportionality factor. The dependence of Ψ on
$C - C_e$ is very complex;[1] however, considering that singu-
larities are still analyzed in the linear approximation and that
$\Psi(0) = 0$, we confine our discussion to a linear dependence
$\Psi = \beta_0(C - C_e)$, where β_0 is the corresponding proportion-
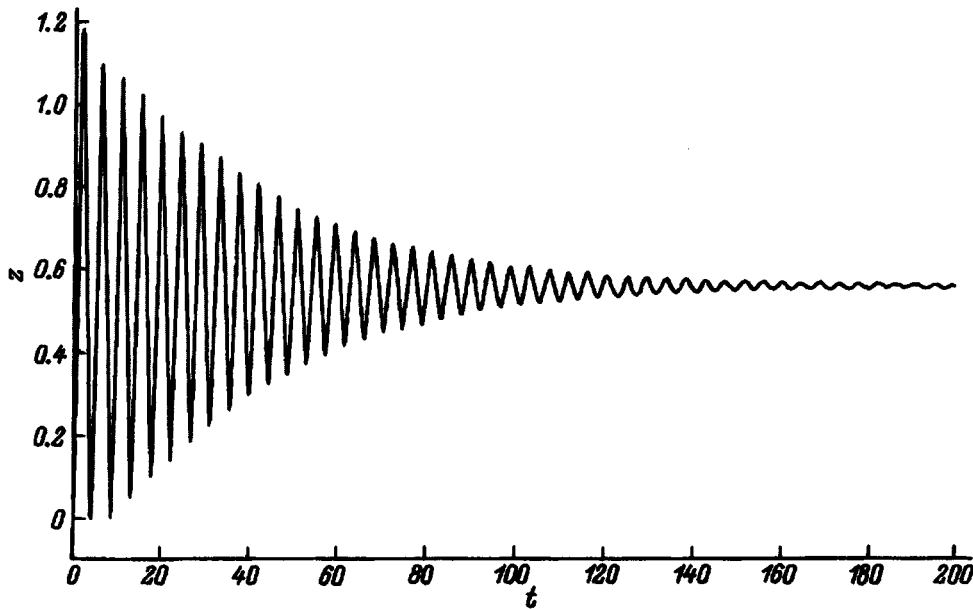ality factor. Following Ref. 4, for the reaction rate φ we

FIG. 1. Density of new-phase islands z versus time t in the equilibrium film growth regime ($J=0.7$).

choose an elementary function with positive feedback, $\varphi = k_0 A C^2$, where k_0 is the reaction constant. We introduce the new dimensionless variables and the dimensionless constants $x = A k_0^{2/3} \beta_0^{-1/3} \gamma^{-1/3}$, $y = C k_0^{2/3} \beta_0^{-1/3} \gamma^{-1/3}$, $z = N k_0^{1/3} \beta_0^{-2/3} \gamma^{-1/3}$, $t = \tau k_0^{-1/3} \beta_0^{2/3} \gamma^{2/3}$ and the dimensionless constants $J = k_0 J_0 / \beta_0 \gamma$, $y_0 = C_e k_0^{2/3} \beta_0^{-1/3} \gamma^{-1/3}$; the system (1) then assumes the form

$$\dot{x} = J - xy^2, \quad \dot{y} = xy^2 - yz, \quad \dot{z} = y - y_0 \quad (z \geqslant 0). \quad (2)$$

An analysis of the singularities of this system in the linear approximation shows that the point $J = y_0^3 - 1$ is a bifurcation point, which leads to the formation of a stable limit cycle. For definiteness we set y_0 equal to 5/4, whereupon bifurcation takes place at a point $J_1 \approx (5/4)^3 - 1 \approx 0.95$. The value of this quantity found by computer simulation is $J_1 \approx 0.888$ (to

an accuracy of 0.0002), i.e., for constant fluxes J smaller than J_1 the system tends to equilibrium (Fig. 1), and for $J_1 < J < J_2 \approx 1.049$ the system undergoes undamped oscillations corresponding to a stable limit cycle (Fig. 2). Finally, for $J > J_2$ this cycle breaks up, and the film grows in an unstable accretion regime (Fig. 3). The problem of controlling the oscillations in the system is extremely crucial in this situation, because the structure and properties of the films will depend on the amplitude and period of the oscillations.[1] One question, in particular, is how the external flux J must be varied with time so that the maximum island density z_{max} will approach a specified value z_* at large times. In the present study, for this purpose, we have developed a general algorithm to solve the stated problem; in general it is valid for a broad class of functions φ, Ψ, and Φ.
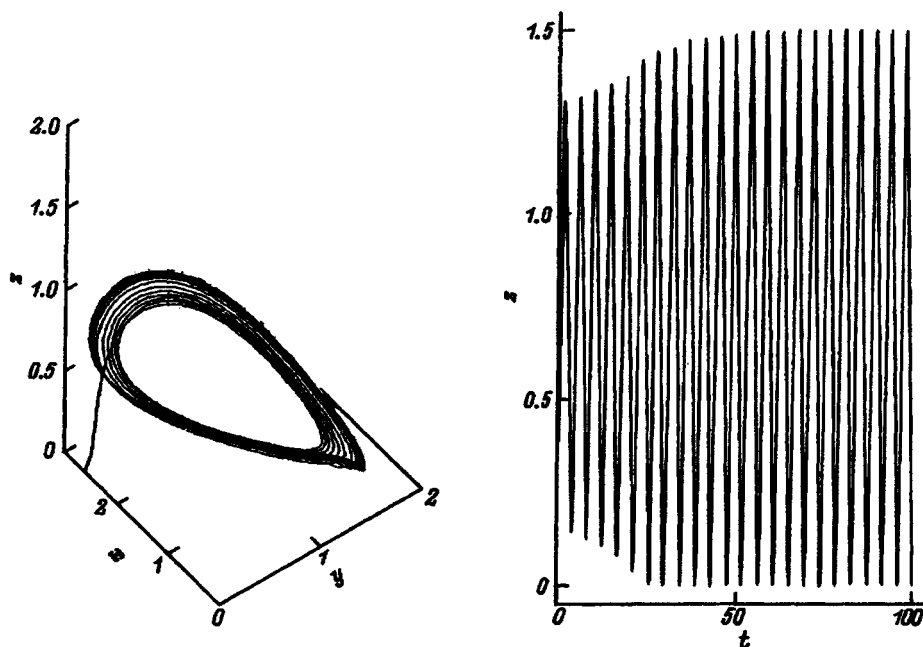

FIG. 2. Density of new-phase islands z vs time t in the stable oscillatory growth regime and the corresponding phase portrait ($J=0.96$).
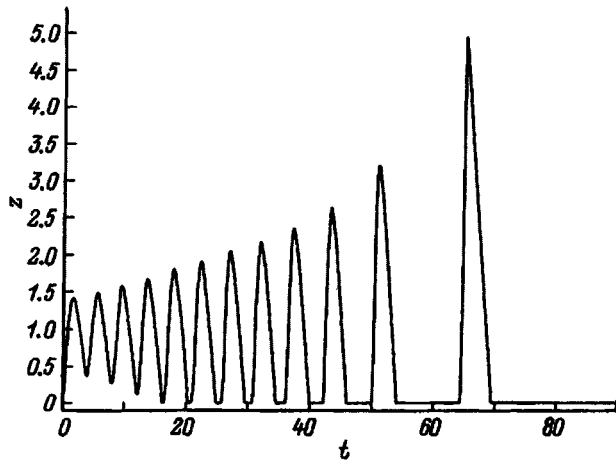
FIG. 3. Density of new-phase islands z versus t in the unstable accretional growth regime ($J = 1.2$).

METHOD OF CONTROLLING OSCILLATORY REGIMES IN THE GROWTH OF THIN FILMS

The control problem for the system (2) is treated as the problem of maintaining the local maxima of $z(t)$ at a prescribed level by varying the function $J(t)$. Consequently, the control function is $J(t)$, the measured (sensed) variable is $z(t)$, and the control objective is expressed in the form

$$|z_k - z_*| \leq \Delta, \qquad (3)$$

where $z_k = z(t_k)$, and t_k is the time at which the kth local maximum of $z(t)$ is attained.

We propose to solve the problem on the basis of adaptive control algorithms that do not require the values of the right sides of the model (2) (Ref. 7). The value of the control function $J(t)$ is changed at the times t_k with allowance for the measurement of z_k, where the law governing the variation of $J_k = J(t_k)$ also changes during the operation as improved estimates are obtained from calculations of the parameters of the model of the controlled system by the adaptation algorithm. A special feature of the algorithm is the transition from the continuous nonlinear model of the controlled system (2) to a linear discrete model obtained by linearizing the Poincaré transform at the points of successive local maxima of $z(t)$ and transforming to the difference equation

$$z_{k+1} + a_1 z_k + a_2 z_{k-1} + a_3 z_{k-2}$$
$$= b_0 J_k + b_1 J_{k-1} + b_2 J_{k-2} + f_k \qquad (4)$$

in the measured variables z_k and the control functions J_k. In Eq. (5) a_1, a_2, a_3, b_0, b_1, and b_2 are unknown coefficients, and f_k is a bounded perturbation (error of the model). It can be shown that the error f_k has the upper bound

$$|f_k| \leq C_1 |z_k - z_*| + C_2 |J_k - J_*|, \qquad (5)$$

where $C_1 > 0$, $C_2 > 0$, and J_* is the value of the control function under the condition $z_k = z_*$.

Consequently, the accuracy of the model increases as the solution is approached. The upper bound (5) is valid not only for the model (2), but also for a more general class of Eq. (1)
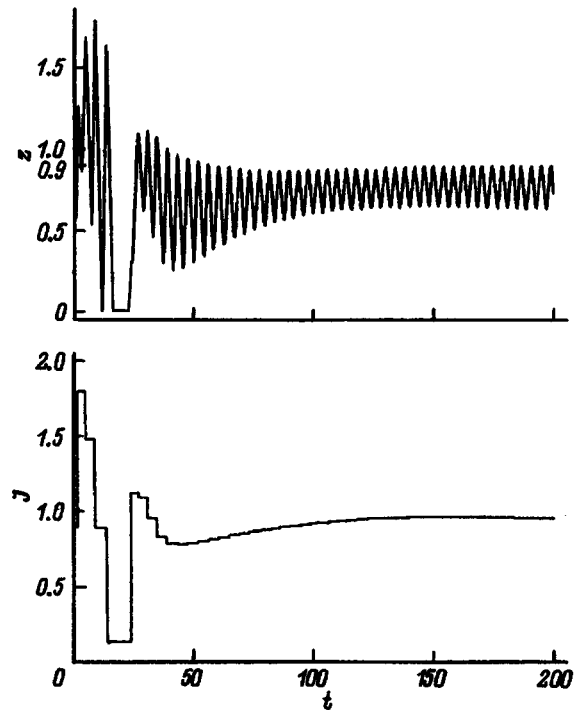
with arbitrary, piecewise-smooth right sides satisfying the Lipschitz condition in a certain neighborhood of a periodic solution.

The adaptive control algorithm includes a main loop algorithm, which computes a new value of the control function J_k, and an adaptation algorithm, which adjust the estimates
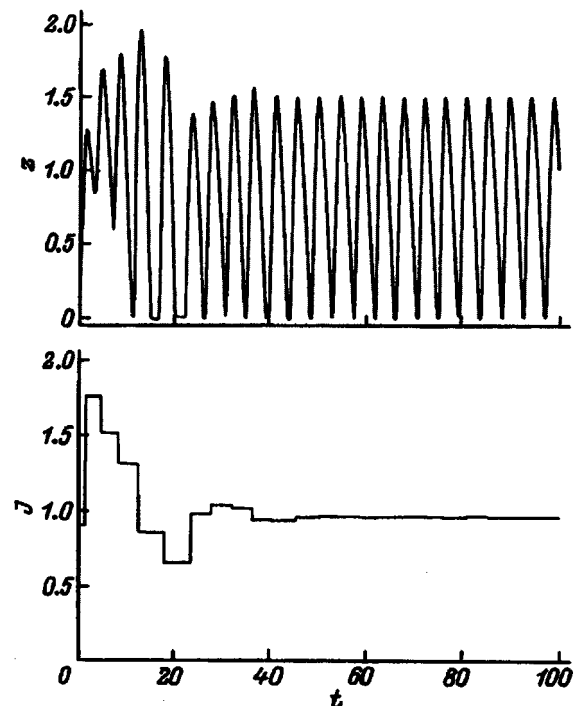

FIG. 4. External flux J and density of new-phase islands z versus time t in the control problem $z_* = 0.9$.

FIG. 5. External flux J and density of new-phase islands z versus time t in the control problem $z_* = 1.5$.

$\hat{a}_{1k}$, $\hat{a}_{2k}$, $\hat{a}_{3k}$, $\hat{b}_{0k}$, $\hat{b}_{1k}$, and $\hat{b}_{2k}$ of the model parameters (4). The main-loop algorithm, written in the form

$$J_k = [z_* + \hat{a}_{1k}z_k + \hat{a}_{2k}z_{k-1} + \hat{a}_{3k}z_{k-2}$$
$$- \hat{b}_{1k}J_k - \hat{b}_{2k}J_{k-1}]/\hat{b}_{0k}, \tag{6}$$

is chosen to ensure that the objective (3) will be achieved in one step if the estimates coincide with the true parameters of the controlled-system model. The adaptation algorithm used to refine the parameter estimates is chosen by the method of recursive objective inequalities[8] and has the form

$$\hat{a}_{i,k+1} = \hat{a}_{i,k} - \alpha \vartheta_k z_{k-i+1}, \quad i = 1, 2, 3,$$

$$\hat{b}_{i,k+1} = \hat{b}_{i,k} - \alpha \vartheta_k J_{k-i}, \quad i = 0, 1, 2,$$

$$\vartheta_k = \begin{cases} z_{k+1} - \hat{z}_{k+1}, & |z_{k+1} - \hat{z}_{k+1}| > \Delta, \\ 0, & |z_{k+1} - \hat{z}_{k+1}| \leq \Delta, \end{cases} \tag{7}$$

where $\hat{z}_{k+1} = -\hat{a}_{1k}z_k - \hat{a}_{2k}z_{k-1} - \hat{a}_{3k}z_{k-2} + \hat{b}_{0k}J_k + \hat{b}_{1k}J_{k-1} + \hat{b}_{2k}J_{k-2}$ is the value of the local maximum in the next step, predicted from the current model, and $\alpha > 0$ is the gain.

The results of Ref. 7 can be used to show that when the model error satisfies the constraint

$$|f_k| \leq \Delta_\varphi < \Delta \tag{8}$$

and if α is sufficiently small, the control objective (3) is attained in a finite number of steps, i.e., inequality (3) is satisfied for $k > k_*$ for a certain k_*. This result proves the hypothetical workability of the proposed adaptive control method.

MAIN RESULTS AND DISCUSSION

The accuracy and rate of convergence of the algorithm (6) and (7) has been analyzed by computer simulation. Simulation is implemented by means of the ADAM software package[9] operating in a MATLAB environment.

Graphs of $J(t)$ and $z(t)$ obtained for the control objective (3) with $z_* = 0.9$ are shown in Fig. 4. The following initial conditions and parameters have been chosen: $J(0) = 0.9$, $x(0) = 0$, $y(0) = 2.6$, $z(0) = 0$, $y_0 = 5/4$. It is evident that $J(t) \to 0.96$ at large times, where the amplitude of the oscillations of the density of new-phase islands decreases, dropping to half the oscillation amplitude for a constant $J = 0.96$. Figure 5 shows the same graphs, but for the objective $z_* = 1.5$. As should be expected, the amplitude of the oscillations z increases in this case, where $J(t) \to 0.96$ as $t \to \infty$. The proposed control procedure is therefore efficient. Clearly, the behavior of the system at large times depends not only on the asymptotic values of the external parameters of the problem, but also on how they change at the initial times; this dependence is indicative of memory effects in systems undergoing first-order phase transition with chemical reactions. These effects are probably responsible for the poor reproducibility of many pertinent experiments, specifically in regard to the growth of high-T_c superconducting films by the MOCVD method.

[1] S. A. Kukushkin and Al V. Osipov, Prog. Surf. Sci. **51**, 1 (1996).

[2] V. I. Trofimov and V. A. Osadchenko, *Growth and Morphology of Thin Films* [in Russian], Énergoatomizdat, Moscow (1993).

[3] J. J. Kozak, Adv. Chem. Phys. **40**, 229 (1979).

[4] S. A. Kukushkin and A. V. Osipov, J. Phys. Chem. Solids **56**, 831 (1995).

[5] S. A. Kukushkin and A. V. Osipov, Fiz. Tverd. Tela (St. Petersburg) **37**, 2127 (1995) [Phys. Solid State **37**, 1157 (1995)].

[6] S. A. Kukushkin and A. V. Osipov, Phys. Rev. B **53**, 856 (1996).

[7] A. L. Fradkov, *Adaptive Control in Complex Systems* [in Russian], Nauka, Moscow (1990).

[8] V. N. Fomin, A. L. Fradkov, and V. A. Yakubovich, *Adaptive Control of Dynamic Systems* (Nauka, Moscow, 1981).

[9] D. L. Egorenov, A. L. Fradkov, and V. Yu. Kharlamov, *Fundamentals of Mathematical Modeling: Construction and Analysis of Models with Examples in the MATLAB Language* [in Russian], BGTU, St. Petersburg (1994).

Translated by James S. Wood

Formation of defects in gallium phosphide grown in the presence of oxygen

A. V. Skazochkin, Yu. K. Krutogolov, V. I. Maĭor, Yu. I. Kunakin, and A. A. Matyash

Scientific-Research Institute of Electronic Engineering Materials, 248650 Kaluga, Russia

G. G. Bondarenko

Moscow State Institute of Electronics and Mathematics, Moscow, Russia

The influence of oxygen introduced in the gaseous phase on the formation of defects in GaP epitaxial layers is investigated by deep-level transient spectroscopy. The extremal dependences of the concentrations of charge carriers and electron traps with energy $E_c - 0.24$ eV on the oxygen flux are discussed. © *1997 American Institute of Physics.* [S1063-7842(97)01009-X]

INTRODUCTION

Together with silicon and carbon, oxygen is one of the principal uncontrollable impurities present in gallium phosphide.[1] Its occurrence in epitaxial layers is not always desirable. For example, the detection of oxygen in GaP green light-emitting diodes alters the purity of the emitted light as a result of the incursion of a red peak in the spectrum due to recombination at Zn–O complexes. The sources of the contaminants in gas-phase epitaxy are usually quartz, surface oxide films on the primary components of the compound, oxygen dissolved in the metals, etc. The oxygen content in GaP can attain 2×10^{18} cm^{-3}, even though the maximum concentration of electrically active oxygen does not exceed $(2-3) \times 10^{17}$ cm^{-3} (Ref. 1). The main mass of the oxygen is assumed to exist in the form of electrically inactive predeposits of the gallium oxide Ga$_2$O$_3$ (Ref. 2).

Oxygen present in the phosphorus sublattice forms a deep donor level with ionization energy $E_c - 0.89$ eV (Ref. 3). Oxygen can also create complexes involving intrinsic lattice defects and other impurities, for example, Si (Ref. 1), Cd, Mg, and C (Ref. 4), and it can influence the density of point defects themselves in growing layers. The formation of deep nonradiative recombination centers in this case has not been investigated. Previously proposed[5–7] models of deep centers in GaP are conducive to the application of deep-level transient spectroscopy (DLTS) for analyzing the attendant processes of defect formation in gallium phosphide structures in this case.

SAMPLES AND MEASUREMENT PROCEDURE

Epitaxial layers of GaP were grown in the system PH$_3$–HCl–Ga–H$_2$–HCl on single-crystal GaP:Te substrates with an electron density $n = 1 \times 10^{17}$ cm^{-3} and orientation (100). A GaP:Te buffer layer of thickness 10μm was grown on the substrate with the aid of an additional GaP:Te source. The sample was doped with oxygen from the gaseous phase. The total thickness of the layers was 25μm. DLTS and $C-V$ measurements were performed on diodes with a $p-n$ junction of diameter 400 μm; the p layer was formed by diffusing Zn to a depth of 5μm. The hole density on the surface of the p layer was $(6-9) \times 10^{18}$ cm^{-3}. Mesa diodes were formed by standard photolithography with etching of the

structure to the depth of the $p-n$ junction. An Ohmic contact was created on the p layer by the galvanic deposition of Au–Zn with a subsequent ''burning in'' at $T = 500$ °C in a hydrogen atmosphere. After the reverse side had been polished, an Ohmic contact was formed on it by spark sputtering of tin.

The variation of the oxygen concentration in the layers was determined from photoluminescence spectra according to the intensity variation of the peak of the red band (~ 1.77 eV), which is due to the presence of Zn–O pairs. Excitation was provided by a HeCd layer emitting at a power of 10 mW and wavelength of 441.6 nm.

The DLTS spectra were measured on the spectrometer described in Ref. 8 with a sensitivity $\Delta C/C_0 = 10^{-6}$. The DLTS data were recorded, stored, and processed with the aid of a computer using programs described in detail in Refs. 6 and 9. The spectra were measured by means of a cryostat for measurements in the temperature range 80–400 K, within ± 0.2-K error limits of determination of the temperature, and by means of a heating unit for measurements in the range 300–573 K, within ± 1 K error limits.

EXPERIMENTAL RESULTS AND DISCUSSION

Figure 1 shows the dependence of the free-carrier density $n = N_d - N_a$, measured by the $C-V$ method, on the oxygen flux. The main background impurity in GaP epitaxial layers is silicon[7,10] owing to the reaction of hydrogen with the quartz equipment.[1] Amphoteric Si in gas-phase GaP, predominantly in the place of gallium, is a donor with ionization energy $E_c - 0.082$ eV, inducing n-type conductivity in nominally undoped layers.[7] Silicon atoms can also exist in the phosphorus sublattice, where it functions as an acceptor with $E_v + 0.202$ eV. The degree of self-compensation of silicon in GaP has not been investigated, but if it is assumed to behave as in GaAs, we can expect Si$_P$/Si$_{Ga}$ = 0.1–0.3 (Ref. 11).

The observed (see Fig. 1) decrease of the free-electron density during the initial increase of the oxygen flux is associated with a drop in the concentration of background Si in the growing epitaxial layers.[1] A further increase in the oxygen flux causes $N_d - N_a$ to increase somewhat (sample 263-7), probably on account of an increase in the concentration of
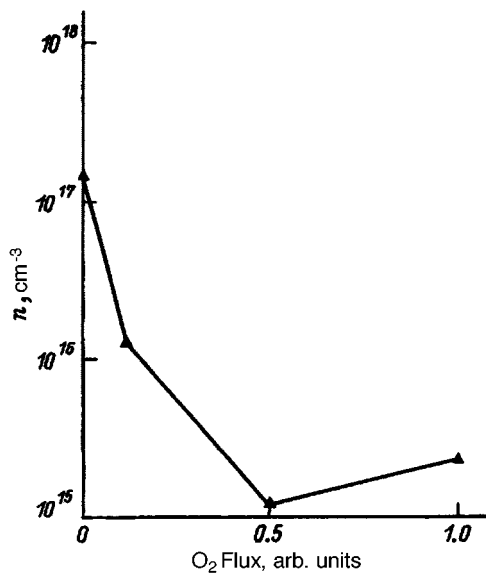
FIG. 1. Free-carrier density versus oxygen flux.

FIG. 2. Deep-level relaxation spectrum of electron traps in sample 263-4, recorded on a diode with a p–n junction. Reverse bias voltage $U_0 = -5$ V, carrier pulse duration 0.5 ms, sampling times $t_1 = 1$ ms, $t_2 = 5$ s.

FIG. 3. Deep-level relaxation spectrum of hole traps in sample 263-4, recorded on a diode with a p–n junction. Reverse bias voltage $U_0 = -5$ V, carrier pulse voltage $U_1 = +3$ V, carrier pulse duration 0.5 ms, sampling times $t_1 = 1$ ms, $t_2 = 5$ s.

intrinsic lattice defects and donor states related to oxygen, which produce free electrons.

The deep-level spectra of the majority and minority carriers, measured for sample 263-4, are shown in Figs. 2 and 3. The depth of the levels in the band gap and the capture cross sections of the majority and minority carriers are determined from the dependence $\ln(tT^2)$ for these centers on the reciprocal temperature $1/T$, where t is the reciprocal of the carrier emission rate from a deep level. Each Arrhenius curve spans at least 20 points. Data on the concentrations, thermal emission activation energies E_a, and the corresponding deep-center capture cross sections σ^∞ are given in Table I, along with data from other papers. To preserve the level indexing convention in Refs. 6 and 7, the numbering of the electron traps begins from the center $E_c - 0.24$ eV.

The electron $T6$ trap appears when GaP is doped with tellurium.[12,13] Tellurium is present in our structures as a background impurity left in the reactor after the growth of the buffer layer. The slight increase in the $T6$ concentration
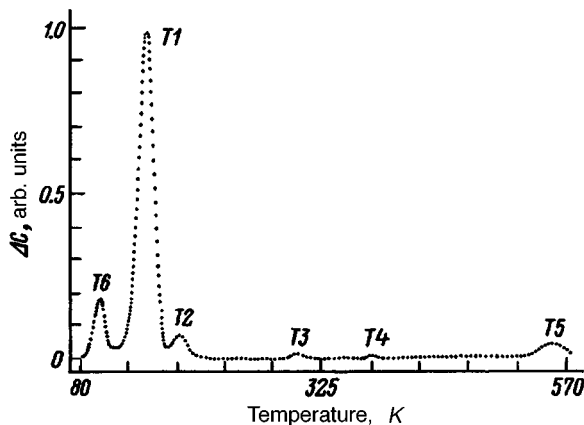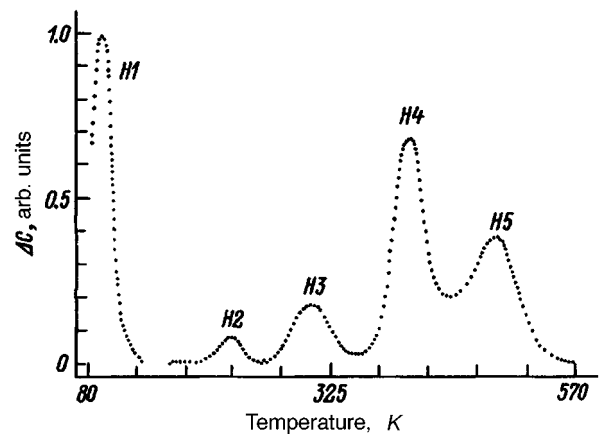
as the GaP layers are doped with oxygen is probably attributable to the improved infusion of tellurium into the phosphorus sublattice due to an increase in the concentration of intrinsic defects, V_P in particular.

It has been established experimentally[6,14] that the $T1$ center is associated with phosphorus vacancies V_P. It has been shown[15] that the $T1$ trap can also contain a background Si atom, i.e., the $T1$ center comprises a Si_{Ga}–V_P complex. On the basis of this model of the $T1$ defect we attempt to explain the variations in the concentration of this center as the oxygen flux increases in the growth of GaP layers. In sample 263-3 the concentration of background Si_{Ga} is much higher than the calculated concentration of V_P ($[V_P] \sim 10^{16}$ cm^{-3}) (see, e.g., Refs. 7 and 15). In this case the variation of the concentration of the $T1$ complex depends on the variation of the concentration of the more dilute component of the complex, i.e., V_P. When oxygen is supplied, the Si_{Ga} decreases, but the concentration of structural defects, including V_P, increases, raising the $T1$ concentration in sample 263-4. With a further increase in the oxygen flux the Si_{Ga} concentration continues to drop sharply, and $[V_P]$ increases, leading to the relation $[Si] \ll [V_P]$ in samples 263-6 and 263-7. In this case the variation of the concentration of the $T1$ defect depends on the variation of the Si concentration and decreases accordingly (Fig. 4).

The deep center $T2$ has been registered in several papers.[5,6,14] Experimental data that might provide a basis for a model of this trap are lacking. The parameters of the $T3$ trap are most likely the same as for the $E5$ trap,[16] which was first detected in the electron-beam irradiation of GaP crystals. In the opinion of Krispin and Maege,[16] $E5$ traps are attributable to phosphorus vacancies. The deep center $T4$ has been detected in samples grown at elevated temperatures or after irradiation with electrons.[5] Kol'tsov et al.[5] regard this deep center as similar to the center $EL2$ in GaAs and therefore endowed with a $P_{Ga}V_{Ga}V_P$ structure. The increase in the concentration of $T3$ and $T4$ defects in sample 263-7 in comparison with sample 263-4 indicates an increase in the concentration of intrinsic defects as the oxygen flux increases.

Kol'tsov et al.[5] and Masse et al.[17] hypothesize that the

TABLE I. Parameters and concentrations of deep centers in GaP layers grown in various oxygen fluxes.

Sample No.	263-3	263-4	263-6	263-7	E_a, eV	σ^∞, cm^2	E_a, eV	References
O$_2$ flux, arb. units	0	0.11	0.5	1.0				
Defect	Concentration ($\times 10^{13}$ cm^{-3})							
T6	6.2	7.7	8.4	9.0	0.16±0.02	$(1-8)\times10^{-16}$	0.16	12
T1	5.9	45.2	8.9	7.2	0.24±0.01	$(1-4)\times10^{-15}$	0.24	14
T2	*	2.6	3.3	3.4	0.28±0.02	$(2-10)\times10^{-16}$	0.28	20
T3	*	0.3	1.8	2.1	0.63±0.02	$(8-30)\times10^{-16}$	0.65	16
T4	*	0.3	1.9	2.9	0.79±0.02	$(3-15)\times10^{-15}$	0.79	5
T5	*	2.1	90.4	200	0.98±0.02	$(2-10)\times10^{-17}$	0.97	5
H1	24.1	21.0	42.8	47.4	0.15±0.01	$(4-20)\times10^{-17}$	0.16	19
H2	*	1.7	1.9	1.9	0.40±0.02	$(1-5)\times10^{-16}$	0.40	19
H3	*	3.6	4.2	5.2	0.64±0.02	$(3-15)\times10^{-15}$	0.64	20
H4	4.4	14.2	35.0	80.1	0.75±0.01	$(8-30)\times10^{-15}$	0.75	21
H5	*	8.7	9.3	13.6	0.93±0.01	$(5-20)\times10^{-16}$	0.95	22
H6	12.3	*	*	*	1.21±0.02	$(1-10)\times10^{-17}$	1.22	7

Note: The prefix T denotes electron traps, the prefix H denotes hole traps, and E_a is evaluated relative to the edges of the corresponding bands; *) concentration below the level of detection.

$T5$ defect is an intricate complex containing gallium and phosphorus vacancies in addition to an oxygen atom. The concentration of the center $T5$ in sample 263-7 is higher than $0.1\cdot(N_d-N_a)$, and to determine it we used the capacitance method proposed in Ref. 18. This method can be used to measure the concentration of defects N_T when N_T is commensurate with N_d-N_a.

The hole traps $H1$ and $H2$ have been detected in Ref. 19, and $H3$ in Ref. 20. The nature of these defects is unknown.

The deep center $H4$ controls nonradiative recombination in high-quality GaP grown by vapor-phase and liquid-phase epitaxy.[21,22] A model for this defect in the form $V_PV_{Ga}Ga_PV_P$. has been proposed.[7,15] The $H4$ concentration increases as the oxygen flux increases, indicating an attendant increase in the concentration of intrinsic defects.

The $H5$ trap was detected in Refs. 7 and 21. Hamilton *et al.*[21] have identified it with nickel. The deep center $H6$ was recorded in Ref. 7. As in the present study, the $H6$ defect was observed in a sample having a high Si content. It may be that Si$_P$ is a constituent of $H6$.

Consequently, as the oxygen flux increases, we observe an increase in the concentration of deep centers associated with intrinsic defects, implying an increase in the concentration of the latter. The introduction of oxygen into the reactor most likely influences the defect concentration in an indirect way. It is a well-known fact that, apart from inserting an oxygen atom into a substitutional site, the introduction of O$_2$ into a continuous-flow system using a hydrogen gas carrier results in the formation of H$_2$O. On the one hand, H$_2$O molecules can inhibit the influx of growth-stimulating components, thereby increasing the concentration of intrinsic defects in the corresponding deep centers. On the other hand, oxygen interacting with the primary growth-targeted and reactor materials, Si in particular, can cause various inclusions to enter the growing GaP layer, a process that also ultimately increases the concentration of deep centers.
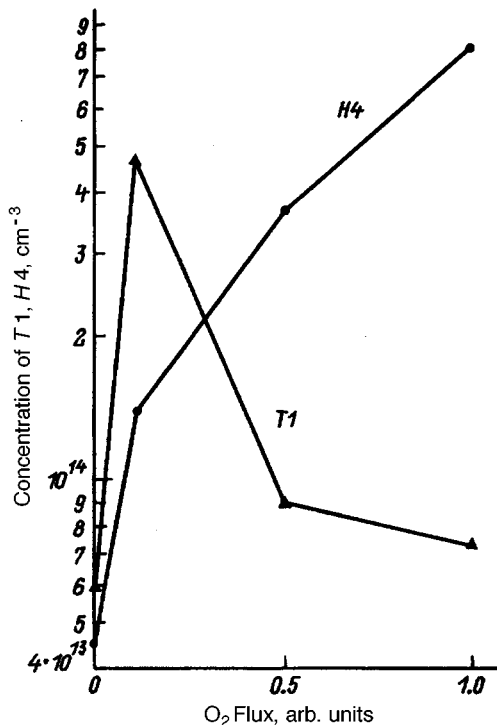

FIG. 4. Concentration of $T1$ and $H4$ traps versus oxygen flux.

[1] V. F. Masterov and B. E. Samorukov, Fiz. Tekh. Poluprovodn. **12**, 625 (1978) [Sov. Phys. Semicond. **12**, 363 (1978)].

[2] M. Kovalchik, A. S. Jordan, and M. H. Read, J. Electrochem. Soc. **119**, 756 (1972).

[3] P. J. Dean, Physica (Utrecht) **117**, 140 (1983).

[4] A. A. Bergh and P. J. Dean, Proc. IEEE **60**, 156 (1972) [Russian translation, Mir, Moscow, 1973].

[5] G. I. Kol'tsov, S. Yu. Yurchuk, V. D. Aleshin, and Yu. I. Kunakin, Fiz. Tekh. Poluprovodn. **24**, 782 (1990) [Sov. Phys. Semicond. **24**, 494 (1990)].

[6] A. V. Skazochkin, Yu. K. Krutogolov, and Yu. I. Kunakin, Semicond. Sci. Technol. **10**, 634 (1995).

[7] A. V. Skazochkin, Yu. K. Krutogolov, and G. G. Bondarenko, J. Adv. Mater., No. 6, 380 (1995).

[8] U. A. Shakirov, O. R. Zhuravlev, A. Sh. Remeev *et al.*, Pribory Tekh. Éksp., No. 5, 250 (1987).

[9] Z. Su and J. W. Farmer, J. Appl. Phys. **68**, 4068 (1990).

[10] G. B. Stringfellow, H. T. Hall, and R. A. Burmeister, J. Appl. Phys. **46**, 3006 (1975).

[11] V. I. Fistul', *Amphoteric Impurities in Semiconductors* [in Russian], Metallurgiya, Moscow (1992).

[12] E. Calleja, E. Muñoz, B. Jimenez *et al.*, J. Appl. Phys. **57**, 5295 (1985).

[13] A. V. Skazochkin, Yu. K. Krutogolov, G. G. Bondarenko *et al.*, in *Third Russian–Chinese Symposium* (1995), p. 110.

[14] P. Kaminski, W. Strupinski, and K. Roszkiewicz, J. Cryst. Growth **108**, 699 (1991).

[15] A. V. Skazochkin, Yu. K. Krutogolov, and G. G. Bondarenko, in *Third Russian–Chinese Symposium* (1995), p. 111.

[16] P. Krispin and J. Maege, Phys. Status Solidi A **84**, 573 (1984).

[17] G. Masse, M. F. Lawrence, and J. M. Lacroix, J. Phys. Chem. Solids **49**, 1349 (1988).

[18] X. Boddaert, D. Deresmes, D. Stievenard *et al.*, J. Appl. Phys. **65**, 5228 (1989).

[19] M. Matyas Jr., Phys. Status Solidi A **97**, 297 (1986).

[20] B. Tell and F. P. J. Kuijpers, J. Appl. Phys. **49**, 5938 (1978).

[21] B. Hamilton, A. R. Peaker, and D. R. Wight, J. Appl. Phys. **50**, 6373 (1979).

[22] A. R. Peaker and B. Hamilton, in *Deep Centers in Semiconductors*, edited by S. T. Pantelides (New York, 1986), Chap. 5.

Translated by James S. Wood

Electrical conduction mechanisms of γ-irradiated amorphous germanium telluride films

I. S. Dutsyak

Lvov State University, 290005 Lvov, Ukraine

Structurally disordered films made from alloys of the system Ge–Te have lately found applications as memory elements.[1,2] The short-range order structure and physical properties of *a*-GeTe films have been well studied to date.[3–8] Their short-range order is described by the model of a tetrahedral environment of atoms with coordination 4(Ge) and 2(Te). Films of *a*-GeTe are semiconductors of the lone-pair type, the majority carriers are holes, the width of the optical gap is 0.7–0.8 eV, and the thermal activation energy of conduction is 0.3–0.4 eV. The structure and physical properties of *a*-GeTe films are sensitive to additives of Bi_2Te_3, Bi, Y, Gd, Tb and also to γ irradiation.[9–10]

Here we report new experimental results on the dark conductivity and thermoelectric power (Seebeck coefficient) of *a*-GeTe films exposed to γ rays at doses of 10^2–10^6 Gy, and we discuss carrier transport mechanisms.

Amorphous GeTe films of thickness 0.3–$1\,\mu$m were prepared by electron-beam deposition in a vacuum of 10^{-6} torr and by rf magnetron sputtering in an argon plasma. Coplanar Al electrodes were first sprayed onto glass or quartz substrates, and then the *a*-GeTe layers were deposited between them at $T_r = 293$ K. These contacts provided a linear I–V curve at applied voltages up to ± 100 V. Some of the samples were γ-irradiated in the cooled channel of a Co^{60} source at doses of 10^2–10^6 Gy. The measurements were preceded by thermal cycling of the samples in the range 120–350 K.

The temperature dependence of the dark conductivity of unirradiated *a*-GeTe films prepared by both methods (Fig. 1) is typical of the conductivity at the mobility edge E_v in the range $T = 140$–350 K and corresponds to the relation

$$\sigma = \sigma_{\min} \exp[-(E_f - E_v)/kT], \tag{1}$$

where E_f is the energy of the Fermi level, and $\sigma_{\min}$ is the minimum metallic conductivity.

If $E_f - E_v$ depends linearly on the temperature, it follows from Ref. 11 that

$$E_f - E_v = (E_f - E_v)_0 - \gamma T \tag{2}$$

and, as a result,

$$\sigma = \sigma_0 \exp(\Delta E_\sigma / kT), \tag{3}$$

where $\Delta E_\sigma = (E_f - E_v)_0$ is the temperature-independent activation energy, and

$$\sigma_0 = \sigma_{\min} \exp(\gamma/k) \tag{4}$$

is a pre-exponential factor.

The thermopower of the unirradiated *a*-GeTe films (Fig. 2 is described by the relation

$$S = -k/e[-(E_f - E_v)/kT - \gamma/k + A_v], \tag{5}$$

where γ is the temperature coefficient of the activation energy; as in Eq. (2), A_v is a parameter that depends on the carrier scattering mechanisms (according to Ref. 11, $A_v = 1$–3).

The values obtained for the thermal activation energies of conduction from the slopes of the $\log \sigma = (10^3/T)$ and $S = f(10^3/T)$ curves agree (see Table I). The parameter γ determined from thermopower measurements is equal to 2.25×10^{-4}eV^{-1}K, which gives $\exp(\gamma/k) = 14$. It then follows from relation (4) that $\sigma_{\min} = 115$ S/cm. Using the relation
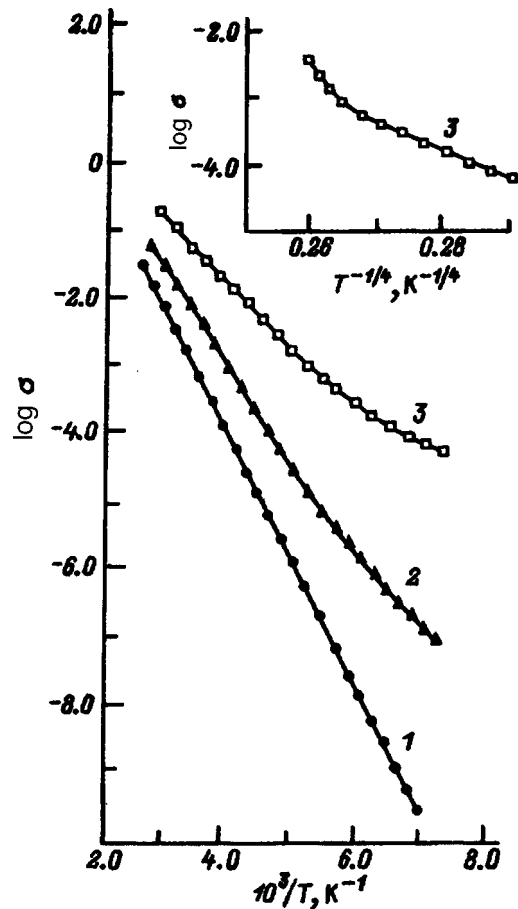
$$\sigma_{\min} = 0.026e^2/a_E h, \tag{6}$$

FIG. 1. Electrical conductivity of *a*-GeTe films versus temperature before and after γ irradiation. *1)* $D = 0$; *2)* $D = 10^4$ Gy; *3)* $D = 10^6$ Gy.
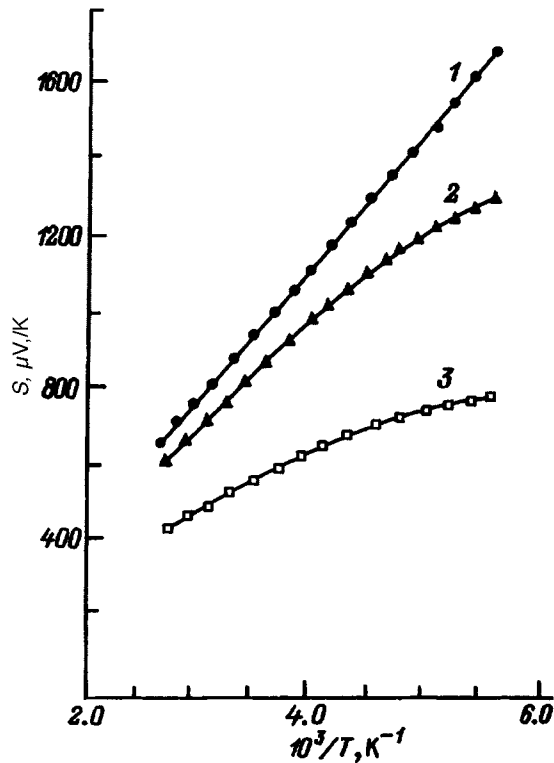
FIG. 2. Thermopower of a-GeTe films versus temperature before and after γ irradiation. *1*) $D=0$; *2*) $D=10^4$ Gy; *3*) $D=10^6$ Gy.

where e is the electron charge, a_E is the damping parameter of the wave function, and h is Planck's constant, we find $a_E = 4$ Å. Mott's equation[11] for a_E has the form

$$a_E/a = [N_v(E_m)/N_v(E_v)], \tag{7}$$

where $N_v(E_v)$ is the density of states at the level E_v, $N_v(E_m)$ is the density of states in the middle of the valence band, and a is the interatomic distance.

Assuming the values $N_v(E_m) = 10^{21} \text{eV}^{-1}\text{cm}^{-3}$ and $a = 2.5$ Å, we obtain $N_v(E_v) = 2.4 \times 10^{20} \text{eV}^{-1}\text{cm}^{-3}$.

At an exposure dose $D = 10^2$ Gy the temperature curves of σ and S do not exhibit any changes. At $D = 10^4$ Gy, in the range $T < 250$ K they deviate from linear and are no longer described by a single conduction activation energy (Figs. 1 and 2). It is evident from the inset to Fig. 1 that for an exposure dose $D = 10^6$ Gy at low temperatures a linear temperature dependence is observed in Mott coordinates, and at high temperatures it exhibits the same variation as in samples irradiated to $D = 10^4$ Gy. The thermopower of the films (Fig.

2) decreases considerably, and its temperature dependence is not described by a single conduction activation energy in the investigated temperature range.

The parameters characterizing the electrical conductivity and thermopower of the unirradiated and irradiated a-GeTe films are given in Table I.

We know[2,11] that the changes in the conductivities of glassy and amorphous semiconductors after irradiation can be attributed to radiation-induced defect formation. According to the models of Matt and Davis[11] and Kastner et al.,[12] a narrow band of localized states exists near the Fermi level in the band gap of an amorphous semiconductor. The origin of this band is associated with various (intrinsic) defects of the substance. Owing to the formation of radiation defects, the conductivity of amorphous layers can change only when their concentration becomes commensurate with (or exceeds) that of the existing intrinsic defects.

An analysis of the results shows that the total conductivity of a-GeTe samples irradiated to a dose of 10^4 Gy is the sum of two components:

$$\sigma = \sigma_v + \sigma_b, \tag{8}$$

where $\sigma_v = \sigma_{01} \exp[-(E_f - E_v)/kT]$ is the conductivity via extended states, $\sigma_b = \sigma_{02} \exp[-(E_f - E_b + W)/kT]$ is the conductivity at the level E_b in the tail of localized states of the valence band, and W is the polaron term.

The thermopower is described as the sum of two conductivity-weighted terms:

$$S = (\sigma_v/\sigma)S_v + (\sigma_b/\sigma)S_b, \tag{9}$$

where

$$S_v = -k/e[-(E_f - E_v)/kT + A_v], \tag{10}$$

$$S_b = -k/e[-(E_f - E_b)/kT + A_b]. \tag{11}$$

For samples irradiated to a dose of 10^6 Gy, at low temperatures we observe a hopping conduction mechanism among localized states near the Fermi level, and at high temperatures conduction takes place by carrier transport to the level E_b. The total conductivity of these samples can be written in the form

$$\sigma_{\text{tot}} = \sigma_{02} \exp[-(E_f - E_b + W)/kT + \sigma_{03} \exp(B/T^{-1/4})]. \tag{12}$$

Invoking the expressions[11]

$$B = 1.66[a^3/k(N_f)], \tag{13}$$

$$R = 3^{1/4}[2\pi a N(E_f)kT]^{-1/4} \tag{14}$$

and assuming, as in Ref. 9, that the reciprocal of the damping parameter of the wave function $a^{-1} = 0.8$ nm, we find that the calculated radiation-induced density of localized states at the Fermi level is $N(E_f) = 2.6 \times 10^{18} \text{eV}^{-1}\text{cm}^{-3}$, and the carrier hopping length is $R = 90$ Å. The thermopower is also described by expressions analogous to (9)–(11).

For a-GeTe films irradiated to a dose of 10^4 Gy, better agreement between the calculated (using the given models) values and the experimental data is observed for $E_f - E_b = 0.17$ eV and $W = 0.04$ eV. This fact suggests that when a-GeTe is irradiated to 10^4 Gy, the tail of the valence

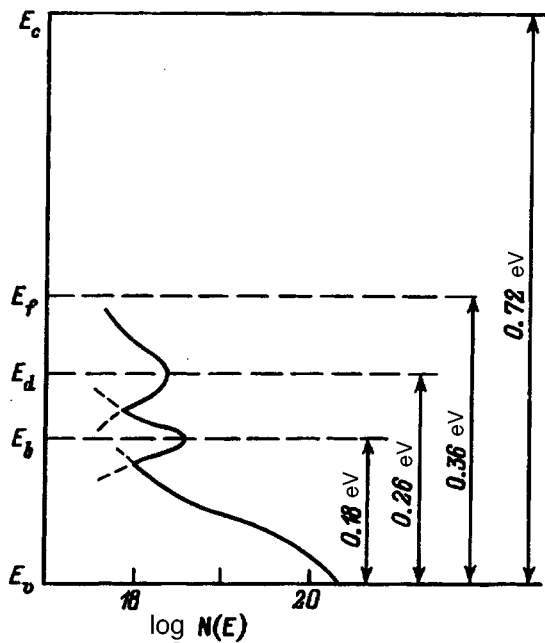TABLE I. Values of the pre-exponential factors (σ_{01}) and thermal activation energies of conduction (ΔE_σ and ΔE_s) for γ-irradiated a-GeTe films.

Exposure dose, Gy	$T > 250$ K			$T < 250$ K		
	σ_{0i}, S/cm	ΔE_σ, eV	ΔE_s, eV	σ_{0i}, S/cm	ΔE_σ, eV	ΔE_s, eV
$0 - 10^2$	1600	0.36	0.36	1600	0.36	0.36
10^4	800	0.32	0.30	5	0.22	0.18
10^6	200	0.20	0.15	–	–	–

FIG. 3. Energy diagram of states below E_f for γ-irradiated a-GeTe films.

films independently of the technique by which they are prepared; these changes are irreversible, because thermal cycling of the samples in the investigated temperature range does not produce any changes in the behavior of the kinetic parameters of the irradiated samples. We assume that the observed changes in the conduction mechanism of a-GeTe is caused mainly by the formation of radiation defects, which we denote, in accordance with Ref. 12, as elementary $C_1^-(\text{Te})$ and $T_3^-(\text{Ge})$, along with complexes of the type $T_3^-(\text{Ge})-C_1^-(\text{Te})$. In view of data[13] on the relative molar energies of the chemical bonds in the system Ge–Te, it is entirely possible that the degree of destruction of certain bonds will depend on the absorption of energy from the incident radiation, creating a definite energy spectrum of localized states in the band gap of the condensates.

band acquires a diffuse level of defects, which governs the hopping transport mechanism in the low-temperature range. At low temperatures an increase of the exposure dose to 10^6 Gy is accompanied by transition to a hopping transport mechanism controlled by a defect level E_d situated 0.1 eV below the level E_f. Hopping transport via the defect level E_b is observed at high temperatures. In Fig. 3 we show the corresponding energy diagram of localized levels induced by γ irradiation in the band gap of a-GeTe films in (the width of the optical gap is taken from data in Refs. 7 and 10).

We also note that the influence of γ irradiation produces major changes in the dark-conduction mechanism of a-GeTe

[1] A. Madan and M. Shaw, *The Physics and Applications of Amorphous Semiconductors* (Academic Press, Boston, 1988) [Russian translation, Mir, Moscow, 1991].

[2] K. K. Shvarts, *Physics of Optical Recording in Dielectrics and Semiconductors* [in Russian], Zinatne, Riga (1986).

[3] L. I. Tamarinova, *Structure of Liquids and Amorphous Solids* [in Russian], Moscow (1983).

[4] T. G. Fowler and S. R. Elliott, J. Non-Cryst. Solids **53**, 43 (1982).

[5] K. L. Chopra, S. K. Barthwal, and A. K. Pandya, Phys. Status Solidi A **35**, 761 (1976).

[6] K. L. Chopra and S. K. Bahl, Thin Solid Films **12**, 211 (1972).

[7] R. Tsu and W. E. Howard, Appl. Phys. Lett. **16**, 380 (1970).

[8] K. P. Scharnhorst, J. Non-Cryst. Solids **23**, 435 (1977).

[9] A. G. Mikolaĭchuk, I. S. Dutsyak, V. S. Frenchko et al., Ukr. Fiz. Zh. **32**, 1534 (1987).

[10] S. V. Makarenko, I. S. Dutsyak, O. G. Mykolaychuk, and A. Z. Pavlyshyn, Phys. Status Solidi A **143**, 353 (1994).

[11] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Clarendon Press, Oxford, 1971) [Russian translation, Mir, Moscow, 1974].

[12] M. Kastner, D. Adler, and H. Fritzsche, Phys. Rev. Lett. **22**, 1504 (1976).

[13] A. Feltz, *Amorphous Inorganic Materials and Glasses* (Weinheim, New York, 1993) [Russian translation, Mir, Moscow, 1986].

Translated by James S. Wood

Structure and nonlinear optical properties of zinc selenide films

A. V. Khomchenko

Institute of Applied Optics, Academy of Sciences of Belarus, 212793 Mogilev, Belarus

The linear and nonlinear properties of polycrystalline zinc selenide films are investigated as a function of their deposition conditions. It is shown that the complex nonlinear refractive index correlates with the crystallite dimensions in the deposited film. This correlation suggests the localization of electrons in surface states of the crystallites as a possible mechanism of optical nonlinearity in zinc selenide films excited in the transparency band at a wavelength of 633 nm. © *1997 American Institute of Physics.* [S1063-7842(97)01209-9]

INTRODUCTION

The preparation and investigation of the properties of zinc selenide (ZnSe) thin films are intriguing by virtue of the potential of this material for the efficient nonlinear conversion of optical signals in information processors.[1] Optical nonlinearity in polycrystalline semiconductor films is usually observed when the films are excited in or at the edge of the strong absorption region. In a number of cases, however, appreciable light-induced changes are observed[2] in the optical properties of semiconductor thin-film structures excited by 633-nm radiation with a power density ~ 1 W/cm^2. Since the emergence of nonlinear properties in polycrystalline films is linked to their structural characteristics,[1] it is important in this regard to investigate the optical nonlinearity as a function of the structure and degree of crystallinity of the deposited films.

EXPERIMENTAL PROCEDURE AND RESULTS

Here we give the results of an investigation of the optical properties and structure of zinc selenide films prepared by rf sputtering of a polycrystalline ZnSe target. Films having a thickness up to 1 μm on substrates of K8 optical glass, fused quartz, and leucosapphire [α-Al$_2$O$_3$, (0001) plane] were grown at substrate temperatures ranging from 350 K to 550 K in an argon atmosphere at a pressure of 0.01–0.03 Pa. The rate of deposition was calculated from the measured thickness $d(\delta d = 10$ nm).

For a given type of substrate, its temperature and the deposition rate are the main parameters governing the properties of the film. The dependence of the film deposition rate V on the substrate temperature T is represented by curve *1* in Fig. 1 (fused quartz substrate). The curve is typical of the growth of films of II–VI compounds in a quasi-closed volume.[3] The decrease in the deposition rate of the films in the range of substrate temperatures below 460 K can be attributed to an increase in the desorption flux from the substrate surface, and the increase in the deposition rate at substrate temperatures above $T_k = 460$ K is accompanied by an increase in the surface diffusion coefficient of the deposited components, which causes the film deposition rate to increase.[3] Consequently, the minimum of this curve is the result of two opposing processes, viz., the escalation of sur-

face diffusion and the activation of desorption processes, as the substrate temperature rises, which govern the structure of the deposited film.

The crystal structure of the films has been analyzed by means of an x-ray diffractometer with a wavelength $\lambda = 1.54178$ Å. A typical diffraction spectrum of ZnSe films is shown in Fig. 2a. Diffraction patterns of films deposited at various substrate temperatures are shown in Fig. 2b. The films are polycrystalline, and in every case the crystallites have a cubic structure oriented predominantly with (022) parallel to the substrate. No other structures have been observed in identifying the diffraction patterns.

The optical transmittance was measured by means of a spectrophotometer in the wavelength range 300–1000 nm. The results for films deposited at various substrate temperatures with allowance for reflection and interference effects are shown in Fig. 3.

All the films are characterized by substantial variation of the absorption near ~ 450 nm, which can be interpreted as the fundamental absorption edge. The presence of steps on the absorption-versus-wavelength curves is associated with size effects due to the polycrystalline structure of the film.[1]

The film refractive index n and the absorption coefficient k at a wavelength of 633 nm were determined by a waveguide technique.[4] The dependence of n on the substrate temperature T is represented by curve *2* in Fig. 1. The absorption coefficient k fluctuated very little for different films and had values from 2×10^{-4} to 10^{-3}. The nonlinear optical constant
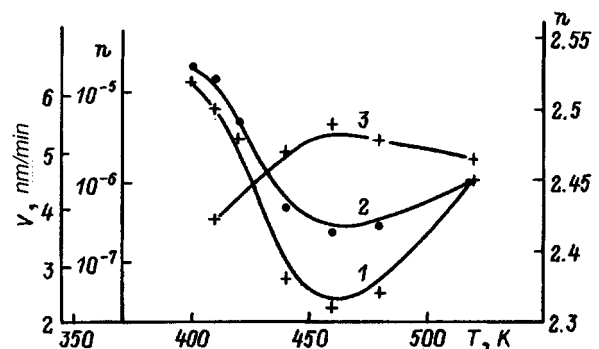

FIG. 1. Film deposition rate (*1*), refractive index (*2*), and nonlinear optical constant n_2 (*3*) versus the substrate temperature.
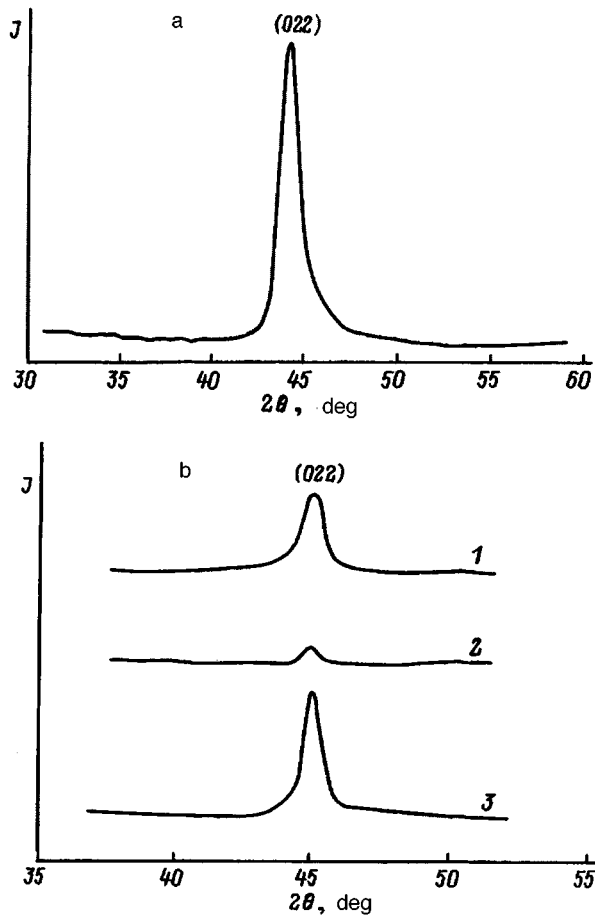
FIG. 2. X-ray diffraction patterns of films deposited on a quartz glass substrate at various temperatures. a: 440 K. b: *1*) 410 K; *2*) 460 K; *3*) 520 K.
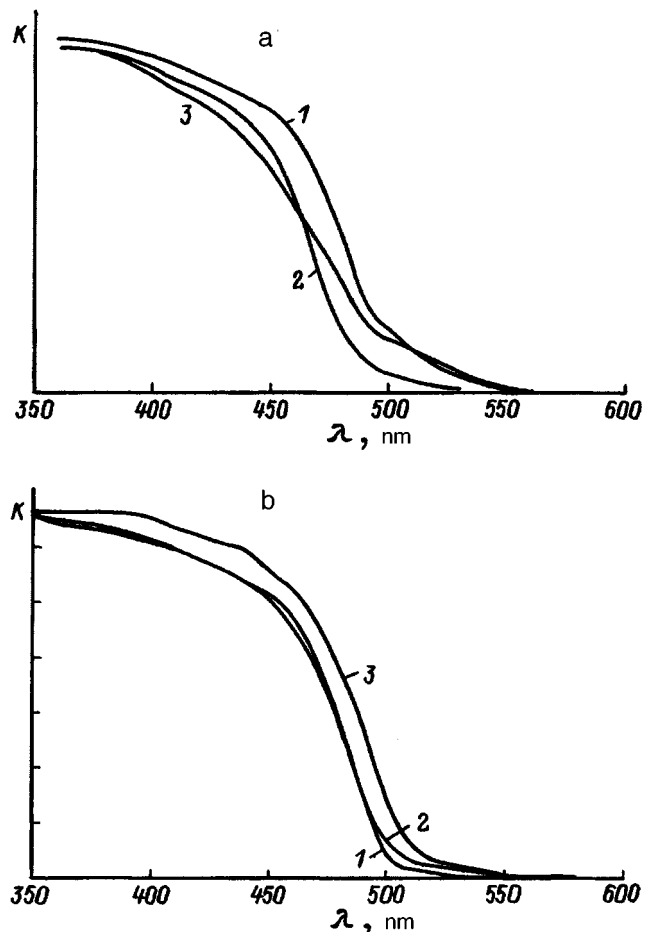

FIG. 3. Absorption spectra of films deposited at various temperatures on a fused quartz substrate (a) and on different substrates at 440 K (b); film thickness $\sim 0.25\ \mu$m. a: *1*) 410 K; *2*) 460 K; *3*) 520 K. b: *1*) Fused quartz; *2*) K8 glass; *3*) sapphire.

n_2 and the nonlinear absorption coefficient k_2 were determined by a procedure described in Ref. 2, i.e., from the variation of the intensity distribution in the Fourier spectrum of a light beam reflected from the film as the power of the incident beam was varied under self-excitation conditions at a radiation wavelength of 633 nm. The dependence of the nonlinear constant n_2 of the film on its deposition temperature is represented by curve *3* in Fig. 1. The value of the nonlinear coefficient k_2 varies from 10^{-5} to 6×10^{-7} (light-induced transmission enhancement of the film is observed) and is also a maximum for films grown at $T = 460$ K. The power density of the sensing beam does not exceed 10 W/cm^2 ($\lambda = 633$ nm). The thermal nonlinearity estimated in accordance with Ref. 2 is negligible in this case.

DISCUSSION OF THE RESULTS

Our investigations of the film properties have shown that at deposition rates above 5.0 nm/min the films have a low refractive index and a diffuse fundamental absorption edge in the absorption spectra, which is shifted toward the long-wavelength end of the spectrum. This characteristic is probably attributable to imperfections of the film and a high concentration of defects due to a departure of the deposited material from stoichiometric composition as the growth rate increases. At deposition rates below 5.0 nm/min the refrac-

tive index of the films is close to the index of single-crystal ZnSe (2.52–2.58; Ref. 5). The absorption rises steeply in the spectrum near ~ 450 nm. Data from an x-ray structural analysis are indicative of the deposition of an oriented zinc selenide film. The film parameters (refractive index, width of the band gap, and structure) scarcely change at deposition rates below 4 nm/min. Judging from the quality of the films and the technological amenability of their deposition, we can regard deposition rates of 4.5–4.0 nm/min as optimal. All the measurement results discussed below have been obtained for films prepared at this deposition rate.

The minimum of the refractive index as a function of the substrate temperature (curve *2* in Fig. 1) and the observed high-frequency shift of the fundamental absorption edge in the absorption edge of films grown at $T = 460$ K (Fig. 3a) are possible evidence of the growth of finely disperse crystallites in the film. Here the film material has a refractive index of 2.47, and the diffraction patterns exhibit a broad peak at 44°, which is characteristic of zinc selenide and corresponds to (022) orientation.

Yodo *et al.*[6] have also noted significant degradation in the quality of ZnSe films grown at temperatures below 470 K in an investigation of the properties of films deposited on

zinc selenide substrates. In the same paper, however, the authors mention a considerable improvement in the quality of the layers as the growth temperature is lowered. In the present study, on the other hand, we have observed an improvement in the crystal quality of a film on a substrate of a dissimilar material at substrate temperatures below 460 K. A growth temperature ~460 K can be called critical from the standpoint of film quality. It is interesting to note that films obtained at growth temperatures of 410 K have a better crystal quality than films grown above 470 K. This statement is supported by x-ray structural analysis data (Fig. 2).

The properties of films deposited on substrates made of different materials differ considerably. For example, as the substrate is changed in the order sapphire–K8 glass–fused quartz, the lattice constant estimated from x-ray spectra decreases from 5.71 Å to 5.66 Å, a high-frequency shift of the fundamental absorption edge is observed in the absorption spectra (Fig. 3b), and the refractive index of the film varies over the interval 2.51–2.48 ($T = 440$ K). These differences are not as pronounced for thicker films. The reason for this is not altogether clear. However, we are inclined to agree with Kalishkin et al.[3] in their hypothesis that the influence of the substrate subsides when the thickness of the layers exceeds certain values, and one can expect a transition to the growth of films with a thermodynamically stable cubic modification. This fact can be exploited to produce oriented films on various substrates at energetically more favorable low temperatures.

Films grown under critical conditions, i.e., films deposited on a fused quartz substrate at a substrate temperature ~460 K, show the greatest promise in relation to studying the nonlinear optical properties. Since the nonlinear optical properties of polycrystalline films are related to size effects due to the influence of the grain boundaries,[1] we have attempted to estimate the dimensions of the crystallites in the films. They are estimated by two methods, which give sufficiently well-correlated results. The grain diameter is determined from the broadening of the x-ray lines[7] and varies from 19 nm to 7 nm as the substrate temperature varies in the investigated range. Here the minimum grain diameter is found for films obtained at a substrate temperature of 460 K. On the other hand, the size of the crystallites can be estimated from the blue shift of the fundamental absorption edge. Such estimates have been made using the effective-mass approach[8] with allowance for the influence of the variance of the crystallite dimensions.[9] The width of the band gap is calculated from the absorption spectra on the assump-tion of direct transitions. Although the application of the given model is not entirely correct for our case — it does not altogether adequately represent the synthesized structures — and although these calculations are approximate in nature, the crystallites are found to have dimensions ~12–7 nm, which are close to the data of x-ray structural measurements. The minimum grain diameter is obtained for films prepared at the critical temperature. A decrease in the crystallite dimensions leads to an increase in the density of surface states in the bulk of the film, and this causes the complex nonlinear constant to increase.

CONCLUSION

We have prepared oriented, polycrystalline zinc selenide films on an amorphous substrate and in the process have discovered the existence of a critical temperature from the standpoint of the quality of the deposited film.

We have investigated the behavior of the linear and nonlinear properties of the films as their deposition conditions are varied. We have shown that the complex nonlinear constant correlates with the dimensions of the crystallites in the deposited film. This correlation suggests that the localization of electrons in surface states of the crystallites can be viewed as a possible mechanism of optical nonlinearity in zinc selenide films excited in the transparency band for radiation with a wavelength of 633 nm.

[1] S. P. Apanasevich, O. V. Goncharova, F. V. Karpushko, and G. V. Sinitsin, Zh. Prikl. Spektrosk. **47**, 200 (1987).

[2] A. B. Sotskiĭ, A. V. Khomchenko, and L. I. Sotskaya, Pis'ma Zh. Tekh. Fiz. **20**(16), 49 (1994) [Tech. Phys. Lett. **20**, 667 (1994)].

[3] I. P. Kalinkin, V. B. Aleskovskiĭ, and A. V. Simashkevich, *Epitaxial Films of II–VI Compounds* [in Russian], Izd. LGU, Leningrad (1978).

[4] V. P. Red'ko, A. A. Romanenko, A. B. Sotskiĭ, and A. V. Khomchenko, Pis'ma Zh. Tekh. Fiz. **18**(4), 14 (1992) [Sov. Tech. Phys. Lett. **18**, 100 (1992)].

[5] *Crystalline Optical Materials: Catalog*, edited by G. T. Petrovskiĭ [in Russian], Dom Optiki, Moscow (1982), pp. 27–28.

[6] T. Yodo, T. Koyoma, H. Ueda, and K. Yamashita, J. Appl. Phys. **65**, 2728 (1989).

[7] Ya. S. Umanskiĭ, Yu. A. Skanov, A. N. Shanov, and L. N. Rastorguev, *Crystallography, X-Ray Diffractometry, and Electron Microscopy* [in Russian], Metallurgiya, Moscow (1982).

[8] L. Drus, IEEE J. Quantum Electron. **QE-22**, 1909 (1986).

[9] Al. L. Éfros and A. A. Éfros, Fiz. Tekh. Poluprovodn. **16**, 1209 (1982) [Sov. Phys. Semicond. **16**, 825 (1982)].

Translated by James S. Wood

Growth of periodic $Hg_{1-x}Mn_xTe$ structures by liquid-phase epitaxy

S. V. Kletskiĭ

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, 252028 Kiev, Ukraine

A procedure for the calculation of phase equilibria in the tellurium corner of the phase diagram of the ternary system Mn–Hg–Te is described within the framework of the model of regular associated solutions. The initial concentrations of the nonstoichiometric fluxed melts and the temperature regimes of their cooling for the controlled liquid-phase epitaxial growth of inhomogeneous $Hg_{1-x}Mn_xTe$ structures with prescribed modulation of the composition along their thickness are determined by numerical solution of a nonlinear inverse Stefan problem. © *1997 American Institute of Physics.* [S1063-7842(97)01309-3]

We have previously[1] investigated the problem of determining the temperature and concentration conditions for the growth of $Hg_{1-x}Cd_xTe$ epitaxial layers with a specified composition profile. To calculate the phase equilibria in the ternary system Cd–Hg–Te, we used the quasi-binary CdTe–HgTe tie line of the state diagram of this substance, an approach that enabled us to describe the crystallization of the ternary alloy by means of a relatively simple Stefan diffusion problem consisting of one diffusion equation in the liquid phase and appropriate boundary conditions on the moving and stationary boundaries. The determination of all the unknowns characterizing the moving phase interface was reduced to the numerical solution of a single transcendental equation representing a finite-difference analog of the Stefan condition of equal diffusion fluxes.

The description of the phase equilibria in the ternary system Mn–Hg–Te poses a far more complex problem[2-4] and does not allow the approach of Ref. 1 to be transferred formally to the case of liquid-phase epitaxy from a nonstoichiometric $(Hg_{1-z}Mn_z)_{1-y}Te_y$ melt. The low accuracy of some experimental procedures (above all, differential thermal analysis), the existence of metastable states of the melt, the strong diffusion of cadmium from the substrates at the growth temperatures, and other factors severely compound the difficulty of plotting an accurate phase diagram of this substance. Inhomogeneous $Hg_{1-x}Mn_xTe$ films suitable for the fabrication of infrared photodetectors are usually grown by liquid-phase epitaxy from a fluxed melt with a low manganese content and a high tellurium content. It has been shown previously[2] that the model of regular associated solutions is applicable to this range of compositions. The solid phase is regarded as a simple regular solution of the pseudocomponents HgTe and MnTe, and the liquid phase $(Hg_{1-z}Mn_z)_{1-y}Te_y$ is regarded as a completely associated solution. The phase equilibrium equations for the tellurium corner of the Mn–Hg–Te state diagram is then conveniently written in the form

$$x = \frac{z(1-y)}{y}$$
$$\times \exp\left[-\frac{W(1-x^2) - \Delta S_1(T_1 - T) + \beta_0 - \beta_1 x}{RT}\right], \quad (1)$$

$$T = \frac{\Delta S_2 T_2 + \alpha_0 - Wx^2}{\Delta S_2 - \alpha_1 + R\ln\left[\dfrac{y(1-x)}{(1-y)(1-z)}\right]}. \quad (2)$$

Here T is the absolute temperature, ΔS_i and T_i are the entropies and melting points of the binary compounds ($i = HgTe, MnTe$), R is the universal gas constant, and α_i, β_i, and W are adjustable parameters. The parameters α_0 and α_i are used to renormalize the entropy and the melting point of the binary alloy HgTe, β_0 and β_i serve a similar purpose for the alloy MnTe, and the interaction parameter W of the components in the solid phase is chosen to match the calculated and experimental data along the liquidus and solidus curves ($0 < x < 0.1$) for the quasi-binary HgTe–MnTe tie line. The numerical values of all quantities in the phase equilibrium equation are given in Table I. The details of the procedure for determining these values and, in general, the procedure for calculating the phase equilibria in the given system are described in Ref. 2 and in the literature cited therein.

The system of transcendental equations (1) and (2) are amenable only to numerical solution, for example, by the Gauss–Seidel method. However, the thermodynamic theory of phase equilibria is inadequate for calculations of transient processes of liquid-phase epitaxy of inhomogeneous solid solutions. The growth of the epitaxial layer is determined by the diffusion influx of the components toward the moving phase interface from the volume of the fluxed melt and by the temperature variation of the growth system. As a rule, the temperature distribution in liquid-phase epitaxy is small enough that it can disregarded. The controlling time variation of the temperature $T(t)$ to ensure the production of inhomogeneous structures with the required composition-thickness profile $x(r)$ can then be determined by the numerical solution of two parabolic equations describing the diffusion of manganese and mercury in the liquid phase:

$$\frac{\partial c_i}{\partial t} = D_i \frac{\partial^2 c_i}{\partial r^2}, \quad t > 0, \quad h(t) < r < H, \quad i = Mn, Hg, \quad (3)$$

subject to the initial conditions

$$h(t) = 0, \quad T(t) = T_0, \quad c_i(r,t) = c_i^0, \quad t = 0, \quad (4)$$

TABLE I. Thermodynamic parameters for calculating phase equilibria in the system Hg–Mn–Te.

System	Compound	ΔS_i, J/mole·K	T_i, K	W, J/mole	α_0, J/mole	α_1, J/mole·K	β_0, J/mole	β_1, J/mole
Hg–Te	HgTe	38.5	943	–	1633	3.3	–	–
Mn–Te	MnTe	41.9	1430	–	–	–	28890	51920
HgTe–MnTe		–	–	14236	–	–	–	–

Stefan-type conditions at the moving interface:

$$D_i \frac{\partial c_i}{\partial r} = (c_i^s - c_i) \frac{\partial h}{\partial t}, \quad t > 0, \quad r = h(t), \tag{5}$$

and the condition of zero diffusion flux at the second, fixed boundary:

$$\frac{\partial c_i}{\partial r} = 0, \quad t > 0, \quad r = H. \tag{6}$$

Here $c_1 = z(1-y)$ and $c_2 = (1-z)(1-y)$ are the concentrations of manganese and mercury in the liquid phase, respectively, $c_1^s = 0.5x$ and $c_2^s = 0.5(1-x)$ are the same in the solid phase, r is the coordinate, t is the time, $h(t)$ is the instantaneous coordinate of the moving phase interface, D is the diffusion coefficient in the liquid phase ($D_{Hg} = 1.5 \times 10^{-5}$ cm^2/s, $D_{Mn} = 1.15 \times 10^{-5}$ cm^2/s), and H is the initial thickness of the liquid-phase layer ($H = 0.1$ cm). The equilibrium values of c_i^s and c_i at the crystallization front are related to the instantaneous temperature by the phase diagram equations (1) and (2).

In the direct Stefan problem $T(t)$ is given, and the unknowns are x, y, z, and $h(t)$. In the inverse problem the required composition profile $x(r)$ of the growing solid phase is given, and the unknowns are $y, z, h(t)$, and $T(t)$. The direct and inverse problems with the same set of remaining parameters are related in such a way that the initial data for one problem are the results of solving the other problem.

A special modification of the general forward/backward elimination method has been developed for the numerical solution of problem (1)–(6); the main distinguishing feature of this modification is the technique used to determine all the unknowns associated with the moving phase interface. Following Ref. 5, the ''best'' finite-difference scheme (implicit, monotonic, conservative schemes with second-order accuracy in the coordinate and first-order accuracy in time) are constructed by global interpolation for the solution of equations of the type (3) on a nonequilibrium Eulerian grid $\omega = \{\tau = j \cdot \delta t, r_f^{j+1} = h, r_{f+1}^{j+1}, r_{f+2}^{j+1}, \ldots, r_n^{j+1} = H\}$. The discrete analogs of the Stefan conditions at the moving interface are written in the form (we drop the superscript $j+1$)

$$D_1 \frac{\alpha_f^1 z(1-y) + \beta_f^1 - z(1-y)}{\Delta r_f} = [0.5x - z(1-y)]V,$$

$$D_2 \frac{\alpha_f^2 (1-z)(1-y) + \beta_f^2 - (1-z)(1-y)}{\Delta r_f}$$

$$= [0.5(1-x) - (1-z)(1-y)]V,$$

where $\Delta r_f = r_{f+1} - r_f$, $V = (h_f^{j+1} - h_f^j)/\delta t$, and α and β are adjustable coefficients.[5]

These equations can be rewritten in the form

$$z(1-y) = a_1 x + b_1, \tag{7}$$

$$(1-z)(1-y) = a_2 x + b_2, \tag{8}$$

where

$$a_1 = \frac{0.5V(\Delta r_f/D_1)}{\alpha_f^1 - 1 + V(\Delta r_f/D_1)}, \quad b_1 = -\frac{\beta_f^1}{\alpha_f^1 - 1 + V(\Delta r_f/D_1)},$$

$$a_2 = -\frac{0.5V(\Delta r_f/D_2)}{\alpha_f^2 - 1 + V(\Delta r_f/D_2)}, \quad b_2 = \frac{0.5V(\Delta r_f/D_2) - \beta_f^2}{\alpha_f^1 - 1 + V(\Delta r_f/D_2)}.$$

All the unknown parameters associated with the phase interface can be determined in each time step by solving numerically the system of transcendental equations (1), (2), (7), and (8) by the Gauss–Seidel method, where for small values of δt a physically reasonable initial approximation is known from the preceding time step. When the space and time intervals are properly chosen, the indicated system is well-posed, and its solution does not require any kind of regularization algorithms. The concentration profiles at the remaining nodes of the grid are calculated from the usual backward sweep of the left elimination process. The advancement of the phase front leads to successive diminution of the total number of nodes $N = n - f$ and variation of the coordinate step $\Delta r_f = r_{f+1} - r_f$. Consequently, in each time step a search is executed within prescribed error limits for the required set of parameters, including the composition of the growing solid phase, as a function of the coordinate of the phase interface.

The method has been tested on a series of control examples. For example, at small times we have $h(t) \sim \sqrt{t}$, and the error of the solution is estimated by comparing the numerical results with the known self-similar solution for a half space; at large times $h(t) \cong H$, and grids with a multiple number of nodes are used (Runge principle[5]). In each time step the total quantity of each component in the liquid and solid phases is monitored:

$$M(t) = \int_0^{h(t)} c_s(r)dr + \int_{h(t)}^H c_l(r)dr = \text{const.} \tag{9}$$

Once the inverse problem has been solved and the required dependence $T(t)$ has been determined, the direct problem is solved using this dependence as a known quantity for determining the concentrations of all components at the phase interface. In every case the composition profile $x(r)$ calculated in the direct problem coincides with the specified profile $x_*(r)$ within the computational error limits.

The temperature–concentration regimes of epitaxial growth of inhomogeneous Hg$_{1-x}$Mn$_x$Te layers with linear
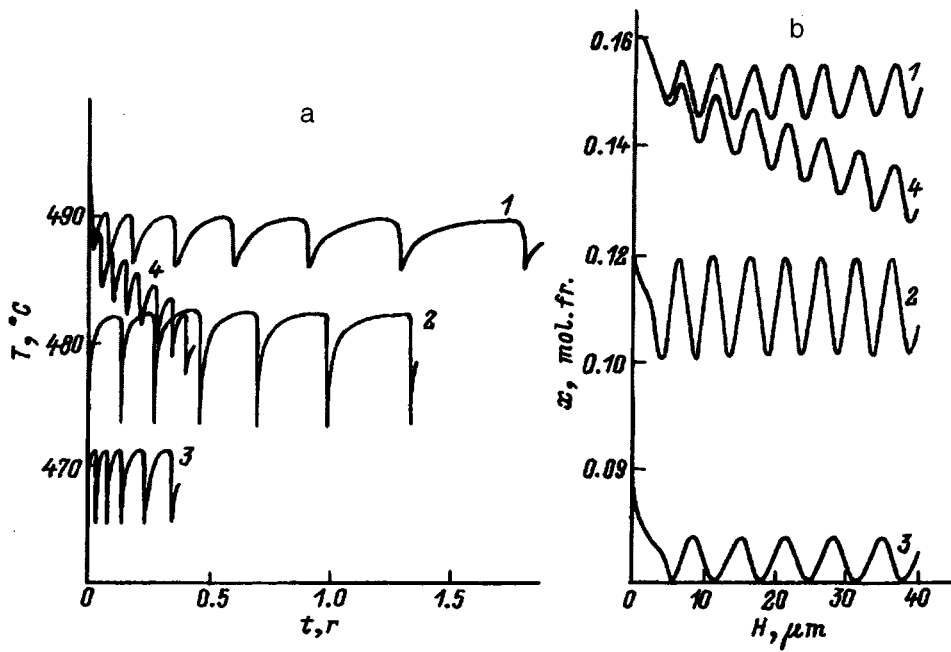
FIG. 1. Time variation of the temperature (a) to ensure the growth of a an epitaxial $Hg_{1-x}Mn_xTe$ layer with prescribed modulation of the composition (b), $y=0.8$. *1,4*) $x=0.17$, $z=0.045$, $T=497$ °C; *2*) $x=0.13$, $z=0.037$, $T=488$ °C; *3*) $x=0.076$, $z=0.025$, $T=478$ °C.

composition profiles have been determined previously. Figure 1a shows the temperature curves $T_i(t)$ ensuring the growth of epitaxial layers with a thickness profile of the composition in the form

$$x(r)=A\,\sin(\omega r)+Br+C, \qquad (10)$$

where A,B,C, and ω are given constants.

These composition profiles with the corresponding curve numbers are shown in Fig. 1b. The small initial segments of these curves are associated with the need to create a certain supersaturation of the growth solution to prevent the epitaxial layer from dissolving as the temperature periodically increases. It is seen at once that the temperature curves $T_i(t)$ in Fig. 1a are deformed by the analogs of the concentration profiles (Fig. 1b). The growth of regions of the epitaxial layer with high values of the concentration $x(r)$ is accompanied by a decrease in the subcooling temperature difference, while the growth of zones with low values of $x(r)$ is accompanied by increased subcooling. This process leads to contraction of the lower parts of the $T_i(t)$ curves and elongation of the upper parts. The gradual manganese and mercury enrichment of the liquid phase has the effect of increasing the time intervals between successive maximum of the temperature curves. Notice that this behavior is not observed for curves of the type *4*, because the enrichment of the melt in this case is offset by a gradual transition into the low-temperature region. The same cause accounts for the decrease in the total growth time of layer *3* in comparison with layers *1* and *2*. These results, of course, are in no way submitted as an exhaustive description of the problem, but merely to demonstrate the possibilities of the approach.

In summary, the mathematical modeling of liquid-phase epitaxy of the semimagnetic semiconductor $Hg_{1-x}Mn_xTe$ provides a means for determining the initial compositions of the growth solutions, their required degree of supersaturation, and the law specifying the time variation of the temperature of the growth cell in order to grow inhomogeneous structures with specified modulation of the composition. The given approach is easily generalized to the growth of single crystals and to the liquid-phase epitaxy and liquid-phase electroepitaxy of other multicomponent systems.

[1] G. I. Zhovnir and S. V. Kletskiĭ, Zh. Tekh. Fiz. **57**, 2247 (1987) [Sov. Phys. Tech. Phys. **32**, 1358 (1987)].

[2] G. I. Zhovnir, S. V. Kletskiĭ, N. V. Sochinskiĭ, and V. M. Frasunyak, Izv. Akad. Nauk SSSR Neorg. Mater. **25**, 1216 (1989).

[3] M. A. Danilov, A. M. Litvak, and K. E. Mironov, Izv. Akad. Nauk SSSR Neorg. Mater. **28**, 1860 (1992).

[4] T. I. Koneshova and E. N. Kholina, Neorg. Mater. **30**, 1101 (1994).

[5] A. A. Samarskiĭ, *Theory of Differencing Schemes* [in Russian], Nauka, Moscow (1983).

[6] G. I. Zhovnir, S. V. Kletskii, and N. V. Sochinskii, Phys. Status Solidi A **115**, K31 (1989).

Translated by James S. Wood

Thermal and Auger processes in $p-n$ junctions based on GaInAs/InAs and InAsSbP/InAs heterostructures

G. A. Sukach, P. F. Oleksenko, A. B. Bogoslovskaya, Yu. Yu. Bilinets,[†)]
and V. N. Kabatsiĭ

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

A study is made of the excess-energy relaxation processes and the mechanisms responsible for overheating of the active zone of infrared emitters made from nonisoperiodic structures with stressed InGaAs layers and from nearly isoperiodic InAsSbP structures and emitting in the wavelength range $\lambda = 2.5-5.0$ μm are investigated. The relationship between the overheat ΔT of the active zone of the structure and Auger processes is established for $In_{1-x}Ga_xAs$ infrared emitters. It is shown that the efficiency of Auger recombination decreases as x increases in the interval 0–0.09, promoting a sharp reduction in ΔT. At $x > 0.09$ the efficiency of CHHS Auger processes decreases exponentially, but an increase in the density of dislocations due to the appreciable value ($\sim 6.9\%$) of the lattice mismatch parameter causes ΔT to increase, but slowly. © *1997 American Institute of Physics.* [S1063-7842(97)01409-8]

INTRODUCTION

The development and investigation of semiconductor infrared emitters operating in the wavelength range $\lambda = 2.5-5.0$ μm, which spans the absorption bands of most industrial (noxious, toxic, and explosive) gases, unquestionably poses a timely problem.

The occurrence of severe current overheating has been established experimentally[1,2] in single and double heterostructures made from isoperiodic InGaAsSb/InSb compounds emitting in the range $\lambda = 1.7-2.4$ μm. This phenomenon naturally produces unwanted changes in the radiative, threshold, and other characteristics of devices. The authors have also analyzed the interrelationship between recombination processes (including Auger recombination) and thermal processes, along with the role of energy barriers in the latter.

Similar research has not been performed on excess-energy relaxation processes and optimization of the overheating temperatures of the active zone of IR emitters utilizing nonisoperiodic compounds with stressed InGaAs/InAs layers and almost-isoperiodic emitters with unstressed InAsSbP/InAs layers operating in the range $\lambda = 2.5-5.0$ μm.

SAMPLES AND EXPERIMENTAL RESULTS

We have investigated nonisoperiodic heterostructures with stressed layers made from $In_{1-x}Ga_xAs$ ($0.02 < x < 0.25$) and emitting in the range $\lambda = 2.5-3.6$ μm, and also structures made from $InAs_{1-x-y}Sb_xP_y$ ($x = 0.07-0.12$, $y = 0-0.1$), $\lambda = 3.8-5.0$ μm. Layers with $y \Rightarrow 0$ were grown to decrease the width of the band gap E_g and to obtain unstressed structures nearly isoperiodic with the InAs substrates.

Layers of n-InGaAs and n-InAsSbP on (111)-oriented n-InAs substrates were grown in a hydrogen flow from fluxed melts containing In and Ga with weighed portions of InAs (InGaAs) and containing In and Sb with weighed portions of InAs and InP (InAsSbP); p-type layers were obtained by introducing elements Mn and Zn into the melt. The density of equilibrium carriers in the n-InGaAs was $n_0 \sim 3-5 \times 10^{17}$cm^{-3} (due to the excess of In and Ga), in the p-region $p_0 \sim 9 \times 10^{16}-4 \times 10^{17}$cm^{-3}, and in the InAsSbP the carrier densities were approximately an order of magnitude lower. The upper p layer had a thickness in the range 8–15 μm. The investigated emitters had an area of 800×800 μm with a point contact on the emitting surface (the contact area comprised $\sim 20\%$ of the area of the emitting surface) and a full-coverage Ohmic contact on the back surface. The Ohmic contact was formed by bonding an In−Mn−Au compound to the p-InGaAs and the p-InAsSbP, and an In−Sn−Au compound to the n-InAs. In the InGaAs IR emitters the output radiation exited through the upper p-type epitaxial layer, and for the InAsSbP compounds it exited through the substrate.

Figure 1 shows the dependence of the overheat of the active zone of the diode $\Delta T = T_{p-n} - T_{amb}$ (T_{p-n} is the temperature of the $p-n$ junction, and T_{amb} is the ambient temperature) on the amplitude of the sinusoidal current, determined by a procedure described in Refs. 3 and 1. A characteristic feature of the experimental curves is their grouping into two series: curve *1* for InAsSbP heterostructures and curves *2–4* for $In_{1-x}Ga_xAs$ heterostructures with various concentrations x. Each of the last three curves can be partitioned into two segments $\Delta T \sim I^k$, depending on the current and type of heterostructure. In the first segment ($I < 50$ mA) the exponent $k = 0.71$ for all types of heterostructures; in the second segment ($I > 50$ mA) $k = 0.78-0.83$. The maximum exponent occurs for $In_{1-x}Ga_xAs$ heterostructures with $x = 0.026$, and the minimum for $x = 0.09$.

DISCUSSION OF THE EXPERIMENTAL RESULTS

We analyze the difference in the $\Delta T = f(I)$ curves for p-$In_{1-x}Ga_xAs$ heterostructures with different concentrations x, drawing on processes in narrow-gap materials with a high majority carrier density in the active zone of $p-n$ junction IR emitters.
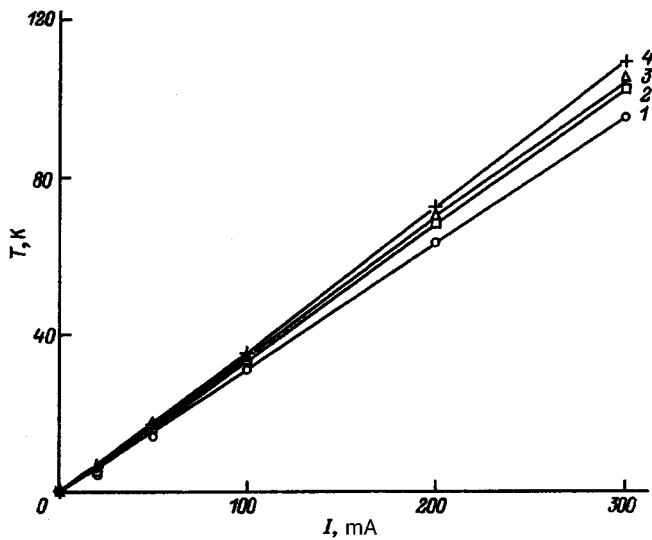
FIG. 1. Overheat of the active zone of IR emitters versus amplitude of the sinusoidal current at $T = 298$ K. *1*) InAsSbP emitter, $\lambda_{max} = 4.31$ μm; *2–4*) In$_{1-x}$Ga$_x$As: *2*) $x = 0.09$; *3*) 0.17; *4*) 0.026.

FIG. 2. Increment of the overheat of the active zone of IR emitters with a specified composition of the In$_{1-x}$Ga$_x$As solid solution above the overheat for IR emitters with the optimum composition ($x = 0.09$) at $T = 298$ K and $I = 300$ mA.

First of all, we note that the effective current $I/\sqrt{2}$ must be used in estimating the thermal effect of the sinusoidal current and to calculate the slope k. The inclusion of this factor brings the slopes of the $\Delta T(I)$ curves in the present study close to those in Ref. 2. The equal slopes of the $\Delta T = f(I)$ curves at $I < 50$ mA (taking into account the thermal effect of the current, $I_{eff} < 35$ mA) for all three In$_{1-x}$Ga$_x$As structures indicates that they have the same mechanism of relaxation of the excess energy. The heating mechanism operates as follows[2]: At low to moderate currents the temperature of the electron gas T_e, whose steady-state value is established by electron–electron and electron–hole interaction, scarcely differs from the lattice temperature $T_p = T_{p-n}$ (only lattice heating of the active zone due to thermal expansion of the lattice constant is encountered here; the contribution of this mechanism does not exceed 25%). At $I > 50$ mA the steady-state temperature of the electron gas exceeds the lattice temperature (the relaxation time to this temperature is $\sim 10^{-14}$–10^{13} s). The excess energy of the electron gas is transferred to the lattice both by interaction with optical phonons (initially with those having a long wavelength and eventually with all of them) and by the interaction of nonequilibrium and nonequilibrium phonons. In this current range the $\Delta T = f(I)$ curves are observed to have different slopes for different In$_{1-x}$Ga$_x$As compositions.

The difference in the slopes k for In$_{1-x}$Ga$_x$As structures with different concentrations x at $I > 50$ mA is conducive to changes both in the energy dissipation mechanism (interaction of carriers with all excited optical and acoustical phonons and phonon–phonon interaction, which causes energy to be transferred from the electronic system to the lattice) and in the recombination mechanism (increasing value of τ_r) and variation of the relation between the monomolecular mechanism of nonradiative recombination and higher-order mechanisms).

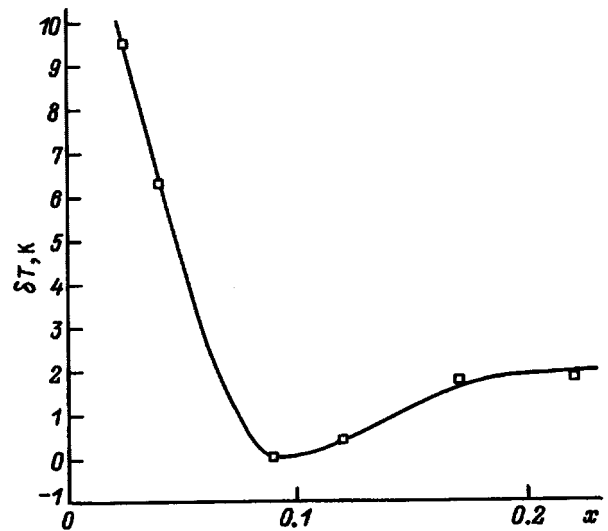Competition between radiative and Auger processes[4–5] is known to provide a definite contribution to recombination processes in In$_{1-x}$Ga$_x$As compounds in the temperature interval 200–400 K. The contributions are comparable at $T \sim 130$–160 K.

The behavior of the temperature $\delta T(x)$ due to the contribution of only Auger and other recombination processes is shown in Fig. 2 as determined from the difference in the overheats of the active zone of emitters with various values of x in the interval 0–0.22 and with $x = 0.09$ (minimum ΔT). It is evident that at maximum amplitude of the sinusoidal current ($I \approx 300$ mA) this difference δT does not exceed 10 K. It has been noted earlier[4] that in the four-band Kane model the classical Auger recombination channel of the CHHV type (recombination of one electron and two holes, with the energy transferred to one of these holes in the same band) in p-InAs materials and in similarly constituted p-In$_{1-x}$Ga$_x$As compounds does not play a significant role. Here the principal mechanism of recombination of nonequilibrium carriers at $T > 100$ K is the CHHS Auger process, i.e., a process involving one electron and two holes, with one of them (the heavy hole $-m_h$) ejected into the spin–orbit split-off band (the rate of this process is $\sim R_A p_0^2 \Delta n$, and its time constant is $\tau_A = R_A^{-1} p_0^{-2}$, where R_A is the rate of Auger recombination by the CHHS mechanism). Here the laws of conservation of energy and momentum do not impose any restrictions on the energies of the recombining particles.

The values of $\tau = 5$–55 ns determined in this study are close to the interband Auger recombination times measured experimentally (at 300 K) by techniques based on the photoconductivity and the photomagnetic effect, $\tau = 5$–10 ns (Ref. 6), and calculated theoretically, $\tau = 1145$ ns (Ref. 5), for p-In$_{0.94}$Ga$_{0.06}$As layers with a carrier density $\sim (1$–$2) \times 10^{17}$cm^{-3}. For p-InAs and p-In$_{1-x}$Ga$_x$As compositions with $x \Rightarrow 0$ the efficiency of the Auger processes is a maximum, and the internal quantum yield η_1 is a minimum. This situation is brought about by the proximity of E_g for these materials to the spin–orbit splitting energy Δ. The lat-

ter consideration is conducive to extinction of the luminescence of such materials and to maximum overheating of structures based on them.[2] According to Ref. 5, upon satisfaction of the condition

$$\frac{E_g - \Delta}{T} = 2\frac{m_h}{m_{so}} \qquad (1)$$

the Auger recombination coefficient is a maximum (since a compromise is struck between the requirement $E_g - \Delta \Rightarrow 0$ and the requirement of a high density of final states for the third charge carrier), does not depend on the density of majority carriers (in the nondegenerate case), and has the value $R_A = 2.2 \times 10^{-27}$ cm^6/s. In Eq. (1) $m_h = 0.41 - 0.42 m_0$ and $m_{so} = 0.14 m_0$ (Ref. 5) are the effective masses of the heavy hole in the spin–orbit split-off band and of the light hole, respectively, and m_0 is the electron mass.

As x is varied in the interval 0–0.09, the efficiency of the nonradiative CHHS Auger recombination channel decreases because the difference $E_g - \Delta$ increases ($E_g - \Delta = 40$ meV at $x = 0.082$ on account of the increase in x, where E_g changes by 40 meV, whereas Δ changes only by 0.2 meV; Ref. 4). The indicated disparity naturally has the effect of increasing the recombination flux through the alternative emitting channel and of decreasing ΔT (see Fig. 2, and also Fig. 1, where ΔT is a minimum for $x = 0.09$); this result correlates with the sharp increase of the radiation power in the active zone of p-InGaAs as $E_g - \Delta$ varies in the interval 30–45 meV [at $T = 77$ K ($x = 0.082$) it is 20 times the value for p-InAs; Ref. 4].

With a further increase in x, even though the density of final states $(m_{so}T)^{3/2}$ also increases, the probability of the CHHS Auger process diminishes exponentially, because only carriers from the tail of the distribution function participate in recombination, on account of the increase of E_g and, in particular, of the difference $E_g - \Delta$ (Refs. 7 and 5). The temperature difference ΔT should decrease even more in this case. In the given range of compositions, however, alternative nonradiative recombination channels are activated as a result of the abrupt increase in the density of inclined dislocations due to the appreciable value ($\sim 6.9\%$) of the lattice mismatch parameter. Viewed from the upper epitaxial layer of p-InGaAs, the dislocation density determined by counting etch pits is observed to increase as x increases, attaining 5×10^6 cm^{-2} at $x = 0$. In actual operation dislocations serve as a source for the formation of a powerful, stable channel of nonradiative recombination, causing $\delta T(x)$ to increase, in the range of compositions where the probability of CHHS Auger processes falls off drastically. They can be Shockley–Read processes through deep levels in the bulk and interface regions of the heterostructure due to the presence of dislocations and their dislocation frames (with an average density of point defects $\sim 5 \times 10^{17}$ cm^{-3}) (Ref. 8), together with the inclusion of other types of Auger recombination having a lower probability under ordinary conditions, for example, classical CHCC and CHHV Auger processes, or of processes of the type CHHL (involving an electron and two heavy holes, one of which is ejected into the light-hole band), where the probability of these processes increases against the background of diminishing probability of CHHS Auger pro-
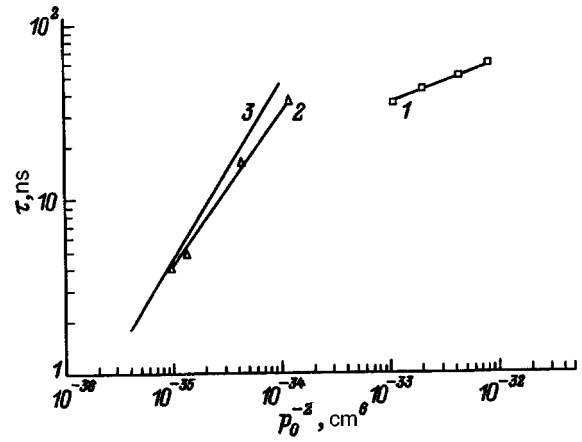

FIG. 3. Experimentally measured lifetime in the compounds p-InAsSbP (1) and p-InGaAs (2) and theoretically calculated time of the CHHS Auger process for InGasAs structures (3). $I = 120$ mA, $T = 298$ K.

cesses in the presence of a large number of phonons and impurities, which remove the constraints imposed on the given processes by the momentum conservation law.[9] This tends to reduce both τ and η_i and, therefore, to increase ΔT, albeit slowly.

The emergence of these various types of nonradiative recombination in the thermal processes of IR emitters based on InGaAs/InAs and InAsSbP/InAs heterostructures is illustrated in Fig. 3, which shows the dependence of τ_{eff} on the density of majority carriers p_0. Clearly, at $p_0 < 3 \times 10^{16}$ cm^{-3} (InAsSbP heterostructures) radiative recombination rather than Shockley–Read recombination contributes predominantly to the dependence $\tau_{eff}(p_0)$ (of course, Shockley–Read nonradiative recombination does contribute to δT). In the latter case τ_{eff} would be observed to be independent of p_0. For $p_0 > 9 \times 10^{16}$ cm^{-3} (InGaAs heterostructures) the main contribution is from Auger processes. This fact is evinced by the closeness of the theoretical and experimental curves for τ_{eff}, which exhibit the same quadratic behavior. The decline of the experimental curve below the theoretical can be attributed to the contribution of other recombination mechanisms (in particular, radiative and Shockley–Read recombination) to the dependence $\tau_{eff}(p_0)$.

CONCLUSIONS

1. We have investigated the overheating temperatures of the active zones of IR diodes of various compositions based on In$_{1-x}$Ga$_x$As ($0.02 < x < 0.25$), emitting in the wavelength range $\lambda = 2.5$–3.6 μm, and InAs$_{1-x-y}$Sb$_x$P$_y$ ($x = 0.07$–0.12, $y = 0$–0.1), $\lambda = 3.8$–5.0 μm.

2. We have established that in In$_{1-x}$Ga$_x$As emitters the main contribution to overheating of the active zone of structures with $x = 0$–0.09 is from CHHS Auger processes (in the limit $x \to 0$ for $I = 300$ mA we have $\delta T \sim 10$ K); for structures with $x > 0.09$ the overheating is due to Shockley–Read recombination mechanisms acting through deep centers.

†)Deceased.

---

[1] N. M. Kolchanova, A. A. Popov, A. B. Bogoslovskaya, and G. A. Sukach, Pis'ma Zh. Tekh. Fiz. **19**(21), 61 (1993) [Tech. Phys. Lett. **19**, 690 (1993)].

[2] N. M. Kolchanova, A. A. Popov, G. A. Sukach, and A. B. Bogoslovskaya, Fiz. Tekh. Poluprovodn. **28**, 2065 (1994) [Semiconductors **28**, 1137 (1994)].

[3] S. V. Svechnikov, G. A. Sukach, N. I. Sypko, and V. V. Nikolaenko, Zh. Tekh. Fiz. **55**, 2265 (1985) [Sov. Phys. Tech. Phys. **30**, 1343 (1985)].

[4] N. V. Zotova and I. N. Yassievich, Fiz. Tekh. Poluprovodn. **11**, 1882 (1977) [Sov. Phys. Semicond. **11**, 1102 (1977)].

[5] B. L. Gel'mont, Z. N. Sokolova, and I. N. Yassievich, Fiz. Tekh. Polu-provodn. **16**, 592 (1982) [Sov. Phys. Semicond. **16**, 382 (1982)].

[6] A. I. Andrushko, Kh. M. Salikov, S. V. Slobodchikov *et al.*, Fiz. Tekh. Poluprovodn. **20**, 537 (1986) [Sov. Phys. Semicond. **20**, 337 (1986)].

[7] M. Aĭdaraliev, N. V. Zotova, S. A. Karandashev, and N. M. Stus', Fiz. Tekh. Poluprovodn. **23**, 592 (1989) [Sov. Phys. Semicond. **23**, 371 (1989)].

[8] B. L. Sharma and R. K. Purohit, *Semiconductor Heterojunctions* (Perga-mon Press, Oxford–New York, 1974) [Russian translation, Sov. Radio, Moscow, 1979].

[9] M. Sh. Aĭdaraliev, G. G. Zegrya, N. V. Zotova *et al.*, Fiz. Tekh. Poluprovodn. **26**, 246 (1992) [Sov. Phys. Semicond. **26**, 138 (1992)].

Translated by James S. Wood

Multiple reflections in backscattering by a single scatterer near a perfectly reflecting surface

F. M. Ismagilova and F. M. Ismagilov

Elabuga State Pedagogical Institute, 423630 Elabuga, Russia

The scattering of a plane electromagnetic wave by a solitary scatterer randomly positioned near a perfectly reflecting surface is investigated in the dipole approximation taking into account infinite multiplicity of scattering. It is shown that the presence of multiply reflected waves increases the ''effective'' polarizability of the particle and adds another component of the dipole moment along the normal to the surface. As a result, the backscattering enhancement effect becomes stronger, and the angle of incidence at which amplification of the *p*-polarized wave vanishes becomes smaller. The influence of multiply scattered waves increases as the particle approaches the boundary in this case. © *1997 American Institute of Physics.*
[S1063-7842(97)01509-2]

The scattering by a system of randomly arrayed particles is accompanied by the formation of coherent channels, which lead to the enhancement of backscattering.[1-3] This phenomenon is observed even in scattering by a single particle located near an interface. The problem has been investigated[4] in the example of scalar (acoustic) waves, where an estimate has been obtained for the enhancement factor in scattering above a perfectly reflecting surface. The vector nature of electromagnetic waves introduces new aspects into this phenomenon.

The scattering of *s*-polarized light by slender rods (i.e., two-dimensional scattering) above a dielectric surface has been investigated.[5,6] The enhancement factor for *p*-polarized light is found to be suppressed as a result of the Brewster effect.[7] Ismagilov[8] has derived simple equations for the backscattering enhancement factor in the three-dimensional scattering of electromagnetic waves by a particle sufficiently far from an interface, when the average distance $\langle \rho \rangle$ from the interface is much greater than the wavelength λ of the incident radiation: $\langle \rho \rangle \gg \lambda$. This condition enables us to disregard the influence of multiply reflected waves, i.e., waves that are scattered by the particle, are incident on the interface, are then reflected by the interface, and are once again incident on the particle. As the particle is brought into closer proximity with the interface, the influence of multiply reflected waves increases significantly, altering the polarization relations obtained for backscattering enhancement.[8]

Here we consider a point scatterer *S* situated near the interface of two media (Fig. 1). A wave from the source *0* arrives at the scatterer *S* by two channels: 1) by the direct path; 2) by reflection from the interface. There are two more channels *1′* and *2′*, by which the singly scattered field arrives at the observation point *0′*. When the positions of the receiver and the source coincide ($0=0′$) the paths $(12′)$ and $(21′)$ are identical, and the corresponding fields become coherent: $u(12′)=u(21′)$ (coherent Watson channels). As a result of statistical averaging over the scatterer positions, the foregoing situation leads to backscattering enhancement. We note that the other two channels $(11′)$ and $(22′)$ do not

contribute to the backscattering enhancement effect and can therefore be designated as idle.

We carry out a wave analysis to determine the polarization effects. Let a scatterer be located at the point $\mathbf{r}_0 = (0, 0, \rho)$ at a distance ρ from a perfectly reflecting surface. We place the coordinate origin at the interface, and direct the *z* axis along the normal at a point such that it passes through the scatterer (Fig. 2a). We also assume that a monochromatic plane wave polarized along the unit vector **e** — $\mathbf{E}_{\text{inc}} = \mathbf{e}E_0 \exp(i\mathbf{k} \cdot \mathbf{r})$, where $\mathbf{k} = (k \sin \theta, 0, -k \cos \theta)$ is the wave vector — propagates from the source and is incident on the particle. Two polarizations of the incident wave relative to the plane of incidence, i.e., the *xz* plane, are possible: 1) *s*-polarization with the polarization vector **e** perpendicular to the plane of incidence, $\mathbf{e}_s = (0,1,0)$; 2) *p*-polarization with the polarization vector **e** lying in the plane of incidence, $\mathbf{e}_p = (\cos \theta, 0, \sin \theta)$.

Let α be the polarizability of the particle. The dipole moment induced on the scatterer consists of three terms:

$$\mathbf{p} = (p_x, p_y, p_z) = \alpha \mathbf{E}_{\text{inc}} + \alpha \mathbf{E}_r + \alpha \mathbf{E}_p, \quad (1)$$
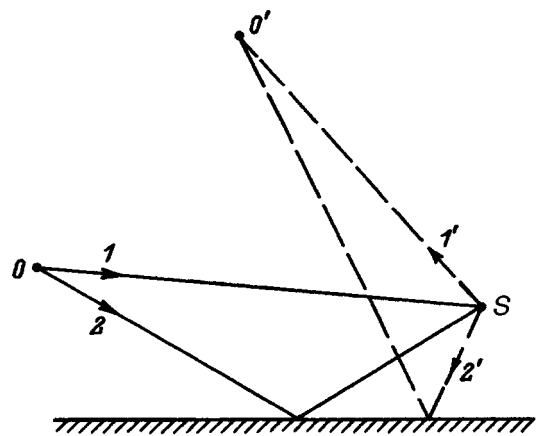

FIG. 1. Ray paths in scattering by a particle *S* near a perfectly reflecting surface. Channels 12′ and 21′ become coherent when the positions of the receiver *0′* and the source *0* coincide ($0=0′$).
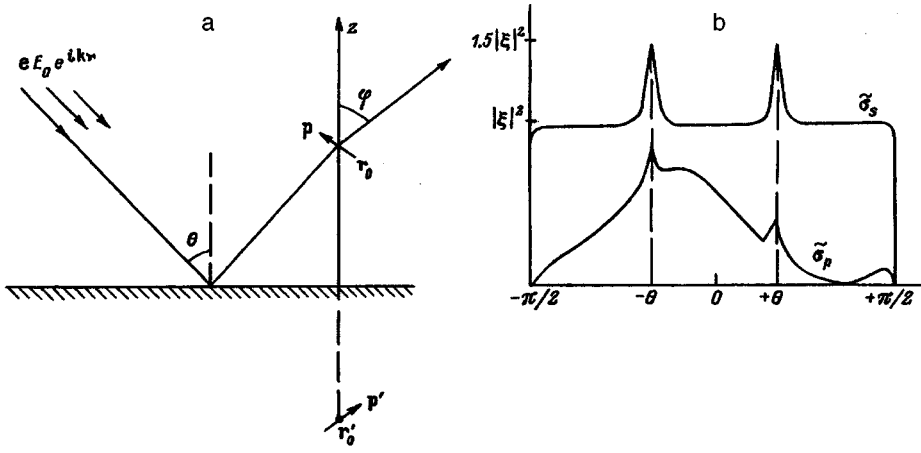
FIG. 2. Coordinate system (a) and angular dependence (b) of the normalized scattering cross sections $\tilde{\sigma}_s(\varphi)$ and $\tilde{\sigma}_p(\varphi)$ for a solitary scatterer randomly positioned near a perfectly reflecting surface.

where $\mathbf{E}_{\text{inc}}$ is the primary incident wave, $\mathbf{E}_r$ is the primary wave reflected by the interface, $\mathbf{E}_p = \hat{g}$, and $\mathbf{p}'$ is the secondary reflected wave.

The tensor Green's function $\hat{g}$ describes the field generated at the position of the dipole $\mathbf{p}$ (point $\mathbf{r}_0$) by its mirror image $\mathbf{p}' = (-p_x, -p_y, -p_z)$ situated at the point $\mathbf{r}_0' = (0, 0, -\rho)$ (Ref. 9):

$$\hat{g}\mathbf{p} = \gamma\mathbf{p}' + \beta\mathbf{e}_z(\mathbf{e}_z \cdot \mathbf{p}'), \tag{2}$$

where $\mathbf{e}_z = (0, 0, 1)$, $\gamma = (k^2/2\rho + ik/(2\rho)^2 - 1/(2\rho)^3) \times \exp(2ik\rho)$, and $\gamma = (-k^2/2\rho + 3ik/(2\rho)^2 + 3/(2\rho)^3) \times \exp(2ik\rho)$.

Equation (1) can be used to find the dipole moment $\mathbf{p}$ with allowance for infinite multiplicity of scattering:

$$\mathbf{p} = \alpha\{\xi(\mathbf{E}_{\text{inc}} + \mathbf{E}_r) + \varkappa\mathbf{e}_z(\mathbf{e}_z \cdot (\mathbf{E}_{\text{inc}} + \mathbf{E}_r))\}, \tag{3}$$

where $\xi = (1 + \alpha\gamma)^{-1}$, and $\varkappa = \alpha\xi(\beta + 2\gamma)(1 - \alpha\beta - \alpha\gamma)^{-1}$.

Equation (3) can be written in the form

$$\mathbf{p} = \alpha_{\text{eff}}(\mathbf{E}_{\text{inc}} + \mathbf{E}_r) + \alpha_{\text{eff}}^z\mathbf{e}_z(\mathbf{e}_z \cdot \mathbf{E}_{\text{inc}} + \mathbf{E}_r)), \tag{4}$$

from which it is clear that multiply scattered waves lead to an "effective" change in the polarizability $\alpha_{\text{eff}} = \alpha\xi$ and an additional variation of the z-component with $\alpha_{\text{eff}}^z = \alpha\varkappa$.

Expanding (3) in $|\alpha\gamma|$ and $|\alpha\beta|$, we obtain an expression for the induced dipole moment in powers of the scattering multiplicity. The case $|\alpha\gamma| = |\alpha\beta| = 0$ ($\xi = 1$, $\varkappa = 0$) corresponds to the absence of multiply reflected waves.

Knowing the induced dipole moment (3) and the laws of reflection of a plane wave from a perfectly reflecting surface,[9] we readily find the intensity of the scattered field at a distant point:

$$\mathbf{r} = r\mathbf{n} = r(\sin\varphi, 0, \cos\varphi) \quad (r \gg \rho, \ r \gg \lambda)$$

$$\mathbf{E}_s = 2\mathbf{e}_s\alpha\xi\frac{k^2}{r}E_0 e^{ikr}[\cos(\beta^+ k\rho) - \cos(\beta^- k\rho)],$$

$$\mathbf{E}_p = 2\alpha\frac{k^2}{r}E_0 e^{ikr}[\xi[\mathbf{n}\cdot[\mathbf{e}_p \times \mathbf{n}]]\cos(\beta^+ k\rho)$$

$$+ \xi[\mathbf{n}\times[\mathbf{e}_p \times \mathbf{n}]]\cos(\beta^- k\rho) + \varkappa[\mathbf{n}\times[\mathbf{e}_z \times \mathbf{n}]]$$

$$\times (\mathbf{e}\cdot\mathbf{e}_z)(\cos(\beta^+ k\rho) + \cos(\beta^- k\rho))], \tag{5}$$

where $\beta^\pm = \cos\theta \pm \cos\varphi$, $\mathbf{e}_p = (\cos\theta, 0, -\sin\theta)$ is the polarization vector of the incident field after reflection from the interface.

It is evident from Eqs. (5) that the inclusion of multiply reflected waves does not lead to depolarization, i.e., transition from one polarization to another. Assuming that the receiver has the same polarization as the source, we find the scattering cross section $\sigma = |(\mathbf{E}\cdot\mathbf{e})|^2 r^2/E_0^2$:

$$\sigma_s = 2\sigma_0|\xi|^2[2 + \cos(2\beta^+ k\rho) + \cos(2\beta^- k\rho)$$

$$+ 2\cos(\beta^+ + \beta^-)k\rho + 2\cos(\beta^+ - \beta^-)k\rho],$$

$$\sigma_p = 2\sigma_0\cos^2(\theta + \varphi)[|C_+|^2(1 + \cos(2\beta^+ k\rho))$$

$$+ |C_-|^2(1 + \cos(2\beta^- k\rho)) - 2\,\text{Re}(C_+^* C_-)$$

$$\times (\cos(\beta^+ + \beta^-)k\rho + \cos(\beta^+ - \beta^-)k\rho)], \tag{6}$$

where

$$C_+ = \left(\xi + \frac{\varkappa}{2}\right)\cos(\theta + \varphi) - \frac{\varkappa}{2}\cos(\theta - \varphi),$$

$$C_- = \frac{\varkappa}{2}\cos(\theta + \varphi) - \left(\xi + \frac{\varkappa}{2}\right)\cos(\theta - \varphi),$$

and $\sigma_0 = \alpha^2 k^4$ is the single-particle scattering cross section.

We assume that the scatterer S is randomly positioned in a volume V_s enclosing a sufficient number of interference bands of the primary field. To find the averaged scattering cross section, it is therefore necessary to integrate Eqs. 6 with a probability density function $w(\rho)$ such that the standard deviation $\sigma_\rho = \sqrt{\langle\rho^2\rangle - \langle\rho\rangle^2}$ will be much greater than the wavelength of the incident radiation λ,

$$\sigma_\rho \ll \lambda. \tag{7}$$

The quantity σ_ρ denotes a characteristic linear dimension of the volume V_s, $\sigma_\rho^3 \approx V$, where V is the volume of the region V_s.

The model density function $w(\rho)$ can be an exponential function

$$w(\rho) = \exp(-\rho/\rho_0)/\rho_0. \tag{8}$$

By virtue of inequality (7) the quantity $\rho_0 = \langle\rho\rangle = \sigma_\rho$ in Eq. (8) must satisfy the condition $\rho_0 \gg \lambda$, so that the average value of the function $\cos(\beta k\rho)$ appearing in[6],

$$\langle\cos(\beta k\rho)\rangle = F(\beta)$$

$$= \int_0^\infty \cos(\beta\rho)w(\rho)d\rho = \frac{1}{1+\beta^2 k^2 \rho_0^2}, \qquad (9)$$

has a resonance dependence on β in the neighborhood of zero $[F(0)=0, \ F(\beta\to\infty)=0]$ with a small half-width $\Delta\beta = (k\rho_0)^{-1} \ll 1$.

The result can be generalized to any density function satisfying condition (7). Of all the maxima of the function $\cos(\beta k\rho)$: $\beta_{\max}^{(n)} = n\pi/k\rho$ ($n = 0, 1, 2, \dots$), only one remains fixed in the same position as ρ varies: $\beta^{(0)} = 0$, $\cos(\beta^{(0)}k\rho) = 1$. Consequently, for a sufficiently broad range of variation of ρ, as a result of averaging, the function $F(\beta) = \langle\cos(\beta k\rho)\rangle$ essentially vanishes outside a narrow interval $\Delta\beta = 1/k\langle\rho\rangle \ll 1$, attaining its maximum value $F(\beta) = 1$ at $\beta = 0$.

The function $\cos(\beta k\rho)$ enters into (6) for four values of the parameter β: $\beta_1 = \beta^+$, $\beta_2 = \beta^-$, $\beta_3 = \beta^+ + \beta^-$, and $\beta_4 = \beta^+ - \beta^-$. The average scattering cross section therefore has maxima under the conditions: a) $\beta^+ = \cos\theta + \cos\varphi = 0$; b) $\beta^- = \cos\theta - \cos\varphi = 0$; c) $\beta^+ + \beta^- = 2\cos\theta = 0$; d) $\beta^+ - \beta^- = 2\cos\varphi = 0$, corresponding to the scattering directions: a) $\varphi = \pi \pm \theta$; b) $\varphi = \pm\theta$; c) $\theta = \pm\pi/2$; d) $\varphi = \pm\pi/2$. Assuming that the angle of incidence θ and the scattering angle φ are not grazing angles ($\theta = \pi/2$, $\varphi = \pi/2$), we find that the scattering cross section has two maxima at $\varphi = \pm\theta$.

The behavior of the normalized scattering cross sections $\widetilde{\sigma}(\varphi) = \sigma(\varphi)/4\sigma_0$, averaged over the scatterer positions, is illustrated schematically in Fig. 2b. It is evident from Fig. 2 that for any polarization of the primary field two peaks of width $\Delta\varphi = (k\langle\rho\rangle\sin\theta)^{-1}$ are observed, corresponding to specular ($\varphi = \theta$) and antispecular ($\varphi = -\theta$) scattering directions. For p-polarization the field vanishes at the scattering angle $\varphi = \pi/2 - \theta$ as a result of the anisotropy of the directivity pattern of the dipole radiation [the factor $\cos^2(\theta + \varphi)$].

The backscattering enhancement effect is manifested in the fact that the scattering cross section in the strictly backward direction

$$\sigma_s^{\mathrm{bsc}} = \sigma_s^{\mathrm{sep}} + 2|\xi|^2\sigma_0,$$

$$\sigma_p^{\mathrm{bsc}} = \sigma_p^{\mathrm{sep}} + 2\left|\frac{\varkappa}{2} - \left(\xi + \frac{\varkappa}{2}\right)\cos 2\theta\right|^2\sigma_0 \qquad (10a)$$

is greater than the scattering cross section in the almost-backward direction

$$\sigma_s^{\mathrm{sep}} = 4|\xi|^2\sigma_0,$$

$$\sigma_p^{\mathrm{sep}} = 2\left[|\xi + \varkappa\sin^2\theta|^2 + \left|\frac{\varkappa}{2} - \left(\xi + \frac{\varkappa}{2}\right)\cos 2\theta\right|^2\right]\sigma_0, \qquad (10b)$$

when the scattering angle deviates from $\varphi = -\theta$ by $\Delta\varphi \approx (k\langle\rho\rangle\sin\theta)^{-1}$.

It is evident from relations (10) that for s-polarization of the incident field the effective backscattering cross section of

a small object situated near a perfectly conducting surface is, on the average, 1.5 times greater than in two-station observation. This simple effect is directly related to the existence of coherent Watson channels. We note in this regard that the backscattering enhancement effect in p-polarization is suppressed, $\Delta\sigma_p^{\mathrm{bsc}} = \sigma_p^{\mathrm{bsc}} - \sigma_p^{\mathrm{sep}} < \Delta\sigma_s^{\mathrm{bsc}} = \sigma_s^{\mathrm{bsc}} - \sigma_s^{\mathrm{sep}}$, and if the angle of incidence θ_0 satisfies the condition

$$\cos 2\theta_0 = \frac{\varkappa}{2\xi + \varkappa} = \frac{\alpha\beta + 2\alpha\gamma}{2 - \alpha\gamma}, \qquad (11)$$

the effect vanishes altogether. Although the problem has been solved taking into account infinite multiplicity of scattering, it must be borne in mind that in the dipole approximation the particle diameter d needs to be sufficiently small: $d \ll \rho$, $d \ll \lambda$. For this reason the parameters $\alpha\beta$ and $\alpha\gamma$ are small: $|\alpha\beta| \ll 1$, $|\alpha\gamma| \ll 1$, so that the influence of multiply reflected waves is not very strong.

To obtain numerical estimates, we consider two limiting cases: a) $\langle\rho\rangle \gg \lambda$; b) $\langle\rho\rangle \ll \lambda$. If the particle is at a large distance $\langle\rho\rangle \gg \lambda$ from the interface, multiply reflected waves can be disregarded: $\xi = 1$, $\varkappa = 0$, and Eqs. (10) assume the form described in Ref. 8. The enhancement factor in the p-polarized channel $\Delta\sigma_p = 2\sigma_0\cos^2 2\theta$ is suppressed relative to the enhancement in the s-polarized channel $\Delta\sigma_s = 2\sigma_0$, and for $\theta = \theta_0 = \pi/4 = 45°$ enhancement is totally absent.

We now consider the opposite limiting case of a particle in close proximity ($\langle\rho\rangle \ll \lambda$). Now $\gamma = -\beta/3 = -(2\rho)^{-3}$. Table I gives the normalized enhancement of the backscattering cross section $\Delta\widetilde{\sigma}^{\mathrm{bsc}} = (\sigma^{\mathrm{bsc}} - \sigma^{\mathrm{sep}})/4\sigma_0$ and the angle θ_0 for three values of the parameter $\alpha\gamma = -0.05, -0.1, -0.2$.

It is evident that the backscattering enhancement factor increases as the particle approaches the boundary, and the angle θ_0 decreases.

It is important to note that averaging over the particle positions can be replaced by averaging over a finite frequency interval ($\omega_1, \omega_1 + \Delta\omega$). It is only necessary that sufficiently many interference bands ΔN pass through the scatterer as the frequency varies from ω_1 to $\omega_1 + \Delta\omega$. If the condition $\Delta N \gg 1$ is satisfied, the backscattering enhancement effect can be observed in a single measurement using a wideband signal.

The results of the foregoing analysis can be used in analyzing experimental data on backscattering by dilute suspensions. In the presence of a dielectric or metal plate (the role of which can be filled, for example, by the walls or bottom of a vessel containing the investigated solution) additional backscattering enhancement can occur besides that associated with scattering by particles. The magnitude of the additional enhancement and the angle θ_0 can serve as a test of the

TABLE I.

$\alpha\gamma$	ξ	$\varkappa$	$\Delta\widetilde{\sigma}_s$	$\Delta\widetilde{\sigma}_p$	θ_0, deg
-0.01	1.01	0.01	0.51	$|0.005 - 1.015\cos 2\theta|^2/2$	44.7
-0.1	1.10	0.14	0.605	$|0.07 - 1.18\cos 2\theta|^2/2$	43.3
-0.2	1.25	0.42	0.78	$|0.2 - 1.4\cos 2\theta|^2/2$	40.9

proximity of particles to the boundary. Similar phenomena occur in the scattering of waves by vegetation,[10] the ground functioning as the interface.

These results should also be useful in the analysis of backscattering by rough surfaces, where the roughness can be regarded as scatterers close to a boundary.[11] The influence of multiply reflected waves cannot be ignored in this case.

[1] K. M. Watson, J. Math. Phys. (N.Y.) **10**, 688 (1969).
[2] Yu. M. Barabanenkov, Usp. Fiz. Nauk **117**, 49 (1975) [Sov. Phys. Usp. **18**, 673 (1975)].
[3] Yu. N. Barabanenkov, Yu. A. Kravtsov, V. D. Ozrin, and A. I. Saichev, in *Progress in Optics*, edited by E. Wolf, (North-Holland, Amsterdam, 1991), Vol. 29, pp. 67–197.
[4] K. G. Akhunov and Yu. A. Kravtsov, KOF (FIAN), No. 8, 8 (1983).
[5] J. J. Greffet, Opt. Commun. **72**, 274 (1989).
[6] J. J. Greffet, Waves Random Med. **1**, 565 (1991).
[7] J. J. Greffet and A. Sentenac, in *Wave Propagation and Scattering in Varied Media II*, edited by V. K. Varadan [Proc. SPIE **1558** (1991)].
[8] F. M. Ismagilov, Waves Random Med. **5**, 27 (1995).
[9] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th English ed. (Pergamon Press, Oxford–New York, 1975) [Russian edition, Nauka, Moscow, 1988].
[10] R. H. Lang, Radio Sci. **16**, 15 (1981).
[11] F. Moreno, J. M. Saiz, P. J. Valle, and F. Gonzalez, Waves Random Med., No. 1, 73 (1995).

Translated by James S. Wood

Formation of x-ray images by the action of an optical image on a diffracting lithium niobate crystal

V. N. Trushin, A. A. Zholudev, M. A. Faddeev, E. V. Chuprunov, and A. F. Khokhlov

Nizhniĭ Novgorod State University, 603600 Nizhniĭ Novgorod, Russia

A possible mechanism for the formation and processing of an x-ray image using a corresponding optical image as a template is considered. The method is based on the thermal influence of light on the x-ray diffraction parameters of an LiNbO$_3$ crystal. A mechanism for the influence of the gradients of the temperature fields formed in the crystal on the x-ray structural parameters of the crystal is proposed. Information which permits the evaluation of some practical aspects of the application of the observed effect in technology is given. © *1997 American Institute of Physics.* [S1063-7842(97)01609-7]

The intensity of the x-ray diffraction peaks for compact ideal crystals differs from the intensity of the corresponding peaks for slightly distorted crystals because of dynamic x-ray scattering effects.[1] This makes it possible to use external influences that alter the structure of a diffracting crystal to control its x-ray diffraction parameters. Thus, many problems of current interest in x-ray optics, particularly the creation of spatially inhomogeneous x-ray beams of variable intensity (x-ray images), can be solved.

The spatial distribution of the intensity in a diffracted x-ray beam in a compact, slightly distorted crystal is determined both by the departure of local regions of the diffracting volume from the reflecting position and by the dynamic scattering effects. An external influence acting on a crystal can alter the influence of both factors on the x-ray diffraction parameters. As a result, the spatial structure of the x-ray beams can correspond to the external signal under certain conditions.

One of the ways to affect a diffracting crystal is to create a nonuniform strain field caused by a nonuniform temperature field in it.[2] The influence of a nonuniform temperature field caused by low-power laser radiation on the x-ray diffraction parameters was considered in Refs. 3–5. The results of the experiments confirmed the possibility of obtaining x-ray images by such a method.

In the present work we investigated the features of the spatial structure of an x-ray diffraction beam from a lithium niobate (LiNbO$_3$) crystal when the diffracting crystal is irradiated by a spatially modulated laser beam.

The experiment was performed on a two-crystal x-ray spectrometer according to the $(n, -n)$ scheme with high angular resolution. A Ge crystal (the 511 reflection) served as a monochromator. The single-domain lithium niobate crystal investigated, which measured $30 \times 20 \times 2$ mm, was oriented in the reflecting position (Fig. 1a) for the (066) reflection. The x-ray beam was incident upon ''blackened'' surface *2* of the *Y*-cut crystal. The surface *2* of the crystal was simultaneously illuminated by the output of continuous-wave YAG:Nd laser *4* (the wavelength was 1.06 μm, and the radiant flux density was 0.02 W·mm^2) through transparency *5*. An optical image in the form of one rectangular (0.7×1.6 mm) and two square (0.6×0.6 mm) light spots was formed by the transparency on the surface of the crystal. When optical radiation passed through the square opening which is hatched in the figure, it was weakened by 50% by a neutral filter. Polished surface *1* of the LiNbO$_3$ crystal opposite the irradiated surface was pressed against the polished surface of solid metal holder *7*, which served to remove heat. Beam *3* of CuK_α x radiation with a 4×15 mm^2 aperture was incident upon the crystal in the region of the optical image. The structure of the diffraction peak of the crystal was recorded on photographic plate *6* in the form of a topogram. The time of a single experiment was 45 min. Figure 1b shows the distribution of the intensity I_{ph} of the optical image formed on the surface of the crystal normalized to the radiant flux density of the light incident upon the transparency. A typical topogram (x-ray image) is presented in Fig. 1c. The spatial structure of the diffracted x-ray beam was characterized by the parameter $\Delta I_{(x,y)} = I_{xy} - I_0$, where I_0 is the intensity of the diffraction peak from the unilluminated surface of the sample. The distribution of the degree of blackening ΔI of photographic plate *6*, which corresponds to the intensity in a cross section of the x-ray beam, is shown in Fig. 1d.

The intensity of the optical radiation was selected such that its influence would be completely reversible. This was monitored by observing the form of an additional topogram obtained after the optical illumination was switched off.

The x-ray image obtained (Figs. 1c and 1d) is, on the whole, the inverse of the optical image. Regions of both increased and diminished intensity of the x radiation are observed in it. The regions with a diminished intensity, which correspond to the illuminated portions of the surface, are surrounded by regions of increased intensity located within the geometric shadow. The maximum contrast of the observed image was ~60%.

We associate the mechanism by which the x-ray image is formed with changes in the diffraction conditions of the x radiation as a consequence of structural distortions caused by inhomogeneous heat-induced strains, whose distribution is determined by the spatiotemporal structure of the optical image. The distortions appearing lead to local alteration of the x-ray diffraction parameters of the crystal.

To account for the details of the observed x-ray image, we mathematically simulated the temperature field within the
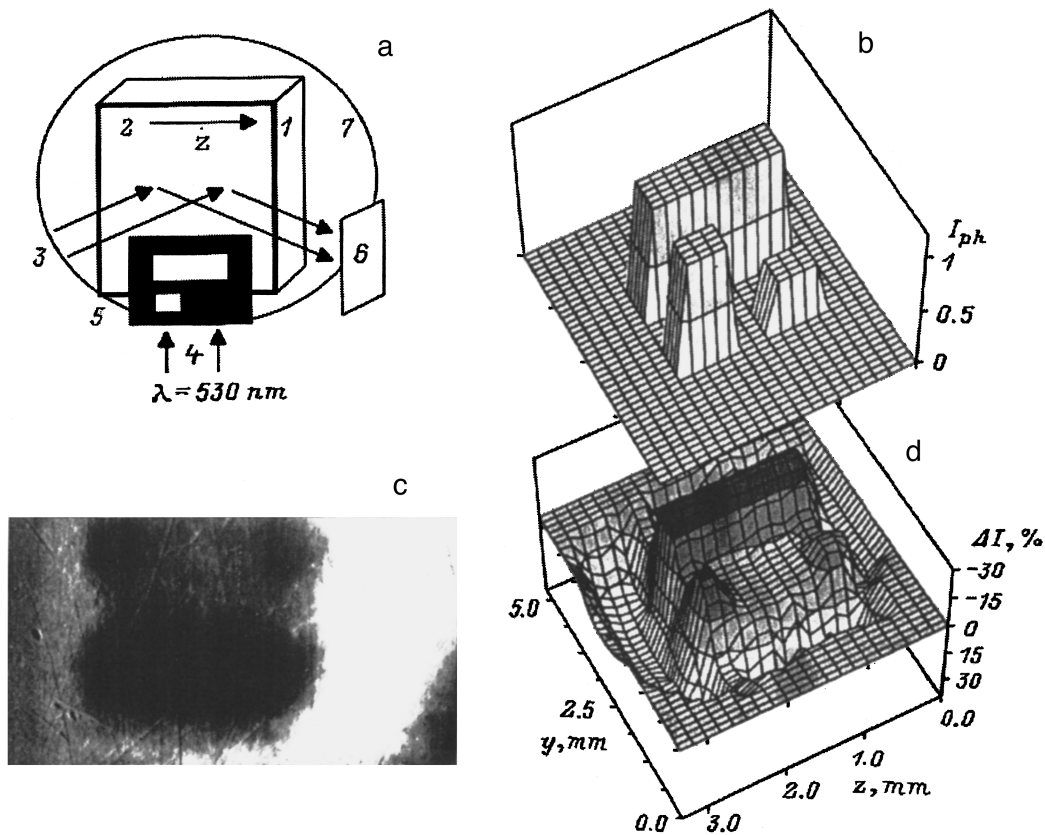
FIG. 1. Diagram (a) and results of the formation of an x-ray image by illuminating a ''blackened'' surface of a Y-cut LiNbO$_3$ crystal with an optical image (b); c — photograph, d — pseudo-three-dimensional distribution of the energy in the x-ray image.

crystal near a light–dark boundary during illumination of the crystal surface. The temporal heat-conduction equation for the LiNbO$_3$ crystal was solved numerically with consideration of the temperature dependence of the components of the thermal conductivity and specific heat tensors.

Figure 2 presents the results of the simulation of the steady-state temperature field T (a) and the x component of the gradient $\nabla_x T$ (b) in the crystal after 5 s of illumination. The temperature field in the vicinity of the light–dark boundary is nonuniform in a band with a width of approximately 6 mm. The field of $\nabla_x T$ in this region has the form of a sharp maximum with a width at half-height of about 1 mm. The position and width of the $\nabla_x T$ peak coincide with the transitional region of the x-ray image corresponding to the edges of the optical image on the transparency. The $\nabla_x T$ maximum is asymmetric relative to the light–dark boundary and has a greater width in the illuminated region. The experimentally measured (Fig. 1) width of the region of increased intensity of the diffracting x radiation in the shadow (≈ 1.2 mm) is significantly greater than the simulated region of temperature variation (≈ 4 mm) and practically coincides with the region of variation of $\nabla_x T$. This makes it possible to presume that the formation of the region with increased intensity of the x radiation is most probably caused by structural distortions initiated by the nonuniform field of the temperature gradient $\nabla_x T$.

Simulation of the heating of the crystal showed that the time dependences of the temperature and its gradient are approximated by exponential functions of the form $A \exp(-t/\tau)$, where A is the steady-state value of the temperature or the gradient component at a given point in the crystal, t is the illumination time of the crystal surface, and τ is the characteristic time (the relaxation time). The spatial dependences of the relaxation times of the temperature and $\nabla_x T$ on both sides of the light–dark boundary (± 2 mm) at different sample depths are presented in Fig. 2c. Plots of the dependence of the relaxation time $\tau_{(y)}$ on the depth at different distances from the light–dark boundary are presented in Fig. 2d.

At a given density of the optical radiation, stationary distributions of the temperature and the components of its gradient were achieved after times shorter than 4 and 2 s, respectively, and they varied subsequently by no more than 5%. The mean temperature of the sample varied by no more than 0.5 °C. The time for establishment of the temperature gradient at all points in the crystal is appreciably smaller than the time for establishment of the temperature at the same points. The shortest establishment time is observed in the region adjacent to the light–dark boundary and near the crystal surface, providing for the initial rapid formation of contours in the x-ray image. The times for formation of the x-ray image after the illumination is switched on and its relaxation after the illumination is switched off do not exceed 2 and 0.3 s, respectively.

The formation of regions with reduced and increased intensities of the diffracted x radiation is determined, in our opinion, by at least two factors. One of them is associated
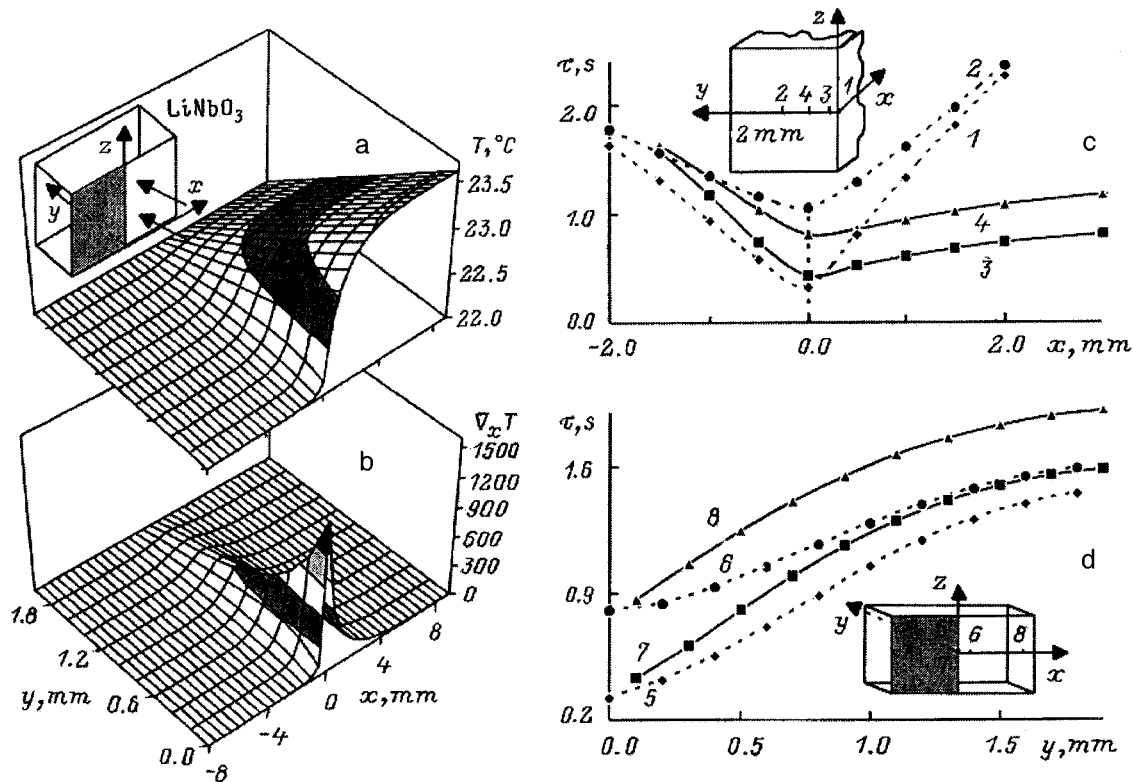
FIG. 2. Two-dimensional distribution of the temperature (a) and its gradient (b) in the plane of a Y-cut LiNbO$_3$ crystal when the XZ surface of a semi-infinite crystal of thickness 2 mm is illuminated by an optical image in the form of a light–dark boundary ($x=0$) with a power density of 0.02 W/mm^2 for 5 s. A diagram of the illumination of the sample is shown in the upper left-hand corner, where the region of the geometric shadow is indicated by gray shading. Spatial dependences of the characteristic times for the establishment of the temperature (3, 4, 7, and 8) and the component of the temperature gradient (1, 2, 5, and 6) in the direction orthogonal to the light–dark boundary (c) and into the depth of the sample (d), respectively. y, mm: 1, 5, 7 — 0; 2 — 1; 3 — 0.1; 4, 6 — 0.5; 8 — 5.

with the variation of the crystal lattice parameters caused by thermal straining of the region where the light acts. This leads to departure of the crystal from the reflecting position and, consequently, to a decrease in the intensity of the diffraction peak. The other factor is determined by dynamic effects, whose manifestation depends on the degree of perfection of the crystal lattice and the magnitude and direction of the temperature gradient. The contribution of each of these factors to the shaping of the image is determined by the distribution of the temperature gradient in the diffraction region of the crystal.

Thus, the increase in the intensity of the x-ray peak near the boundary between light and shadow can be qualitatively explained in the following manner. The LiNbO$_3$ crystal used in the experiment had a high degree of perfection (the width of the rocking curve was about 7″); therefore, the dynamic x-ray scattering effects are manifested in this case. Inhomogeneous thermal straining alters the conditions of the dynamic diffraction of x rays in the crystal with a resultant increase in the intensity of the diffraction peak[2] and causes

"outlining" of the image boundaries. The resolving power of the x-ray image obtained using a lithium niobate crystal can be determined from Fig. 1d and amounts to at least 5 mm[4].

Thus, the intensity distribution of the x radiation diffracted from the crystal under the action of spatially inhomogeneous laser radiation is functionally dependent on the intensity distribution in the laser beam, permitting the formation of x-ray images.

[1] R. W. James, *The Optical Principles of the Diffraction of X-Rays*, Bell and Sons, London (1948) [Russian trans., IL, Moscow (1950)].
[2] M. A. Navasardyan, R. K. Karakhanyan, and P. A. Bezirganyan, Kristallografiya **15**, 235 (1970) [Sov. Phys. Crystallogr. **15**, 197 (1970)].
[3] V. N. Trushin, A. A. Zholudev, M. A. Faddeev *et al.*, Pis'ma Zh. Tekh. Fiz. **21**(9), 72 (1995) [Tech. Phys. Lett. **21**, 348 (1995)].
[4] V. N. Trushin, E. V. Chuprunov, and A. F. Khokhlov, Pis'ma Zh. Tekh. Fiz. **14**, 1749 (1988) [Sov. Tech. Phys. Lett. **14**, 759 (1988)].
[5] V. N. Trushin, T. M. Ryzhkova, M. A. Faddeev *et al.*, Kristallografiya **38**(4), 213 (1993) [Crystallogr. Rep. **38**, 542 (1993)].

Translated by P. Shelnitz

Intermediate regime for the diffraction of light on ultrasound in gyrotropic anisotropic crystals

G. V. Kulak

Mozyr' State Pedagogic Institute, 247760 Mozyr', Belarus

An intermediate regime for the diffraction of light on ultrasound in gyrotropic anisotropic and cubic crystals in an external electric field is considered. A system of equations of the coupled waves, which describes acoustooptic diffraction in gyrotropic anisotropic crystals with consideration of the electrically induced optical anisotropy for a strong interaction between light and ultrasound, is presented. An intermediate regime for the diffraction of light on ultrasound in gyrotropic anisotropic crystals which is close to the Bragg regime for a weak acoustooptic interaction is studied. It is shown that the diffracted light is elliptically polarized and that the ellipticity and polarization azimuth of the diffracted wave depend on the anisotropy of the photoelasticity, the electrically induced anisotropy of the crystal in the external electric field, the gyrotropy, and the asymmetry of the diffraction structure. © *1997 American Institute of Physics.* [S1063-7842(97)01709-1]

The diffraction of light on ultrasonic waves in gyrotropic crystals in the Raman–Nath regime was examined in the approximation of assigned (circular) polarization of the diffracted waves in Refs. 1–4. A system of equations of the coupled waves, which describes the features of Raman–Nath acoustooptic (AO) diffraction in gyrotropic cubic crystals for a strong interaction between light and ultrasound, was presented in Ref. 5. The influence of electrically induced optical anisotropy on the features of Bragg AO diffraction in uniaxial and gyrotropic quartz crystals was investigated in Ref. 6. The AO interaction in the intermediate diffraction regime for a nongyrotropic insulator was considered in Ref. 7 and 8. It was noted there that AO information-processing devices (modulators, deflectors, filters, and processors) exhibit their best characteristics specifically in the intermediate diffraction regime. While in the case of AO interactions in uniaxial and biaxial crystals optical gyrotropy is manifested only for light propagation directions close to the optical axes, in an isotropic medium and in cubic crystals it must be taken into account for any geometry of the interaction between light and ultrasound.[3–6]

In the present work we used the material equations for a gyrotropic insulator[9,10] and the method of slowly varying amplitudes to examine the features of an intermediate regime for the diffraction of light on ultrasound in gyrotropic anisotropic and cubic crystals in an external electric field.

When light propagates near an optical axis of a gyrotropic anisotropic crystal, as in a cubic crystal, the AO diffraction regime is specified by the wave parameter $Q \approx \lambda_0 l/n\Lambda^2$, where n is the refractive index in the propagation direction of the incident light wave, l is the length of the AO interaction region, λ_0 is the wavelength of the light in a vacuum, and Λ is the wavelength of the ultrasound.[7] The intermediate regime for an AO interaction is observed for $0.5 < Q \leq 1.5$. Four diffraction orders should then be taken into account for a light wave incident at the Bragg angle $\varphi_6 \approx \lambda_0/2n\Lambda$ (instead of the two for the Bragg refraction regime). The arrangement of the wave vector $\mathbf{K}$ of the ultrasound and the wave vectors $\mathbf{k}_m (m = -1, 0, 1, 2)$ of the light waves is shown in Fig. 1.

An ultrasonic wave with a displacement vector $\mathbf{U} = \mathbf{U}_0 \exp[i(Kx - \Omega t)]$ ($K = \Omega/v$, where Ω is the angular frequency and v is the phase velocity of the ultrasonic wave) occupies the space between the $z = 0$ and $z = l$ planes. A longitudinal or shear ultrasonic wave induces a spatial and temporal periodic variation of the dielectric tensor $\Delta\hat{\varepsilon}_{ik}$, which is related to the elastic strains $\hat{U}_{ik} = (1/2) \times (\nabla_k U_i + \nabla_i U_k)$ and the photoelastic constants $p_{\hat{i}jkl}$ by the expressions $\Delta\hat{\varepsilon}_{ik} = -\hat{\varepsilon}_{il}\hat{\varepsilon}_{jk}\hat{p}_{ljmn}\hat{U}_{mn}$, where $\hat{\varepsilon} = \hat{\varepsilon}^0 + \Delta\hat{\varepsilon}^e$ is the dielectric tensor of the crystal not perturbed by ultrasound in the external electric field $\mathbf{E}^e$. The components of the dielectric constant perturbation tensor $\Delta\hat{\varepsilon}^e$ of the crystal $\Delta\hat{\varepsilon}_{il}^e = -\hat{\varepsilon}_{ik}^0\hat{\varepsilon}_{jl}^0\hat{r}_{kjt}E_t^e$, where the $\hat{r}_{kjt}$ are the components of the electro-optic tensor and the $\hat{\varepsilon}_{ij}^0$ are the components of the dielectric tensor of the unperturbed crystal.[11]

The wave equation for the light field strength E in the region occupied by the ultrasound follows from Maxwell's equations and the material equations[9,10] for a gyrotropic insulator (see, for example, Ref. 6). We shall seek a solution of the wave equation in the form

$$\mathbf{E} = \sum_{m=-\infty}^{+\infty} \mathbf{E}_m(z) \exp[i(\mathbf{k}_m\mathbf{r} - \omega_m t)], \qquad (1)$$

where $\mathbf{E}_m = A_m(z)\mathbf{e}_m + B_m(z)\mathbf{e}_2$; $\mathbf{k}_m = (\omega_m/c)\sqrt{\bar{\varepsilon}} \times (\sin\varphi_m, 0, \cos\varphi_m)$ denotes the wave vectors of the diffracted waves, ω_m denotes the angular frequencies of the diffracted waves, $\bar{\varepsilon} = 1/3 \, \mathrm{Tr}\,\hat{\varepsilon}$, the influence of the ultrasound and gyrotropy is confined to variation of the wave vectors $\mathbf{E}_m$ in comparison with the case of a nongyrotropic medium, $\mathbf{e}_m = [\mathbf{e}_2/\mathbf{k}_m]/|[\mathbf{e}_2\mathbf{k}_m]|$ denotes the polarization unit vectors lying in the XZ diffraction plane, and $\mathbf{e}_2$ is the unit vector of the polarization orthogonal to the XZ plane.

Substituting (1) into the wave equation for the field strength E (Ref. 6) and using the method of slowing varying
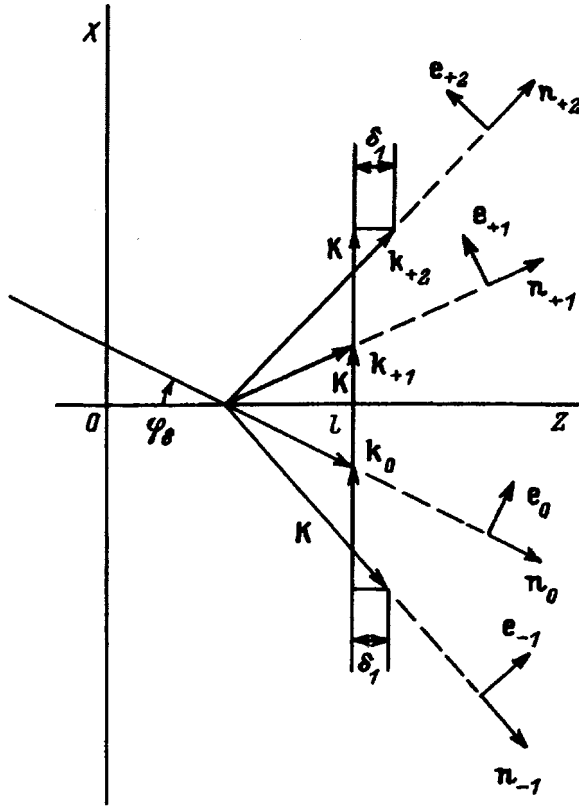
FIG. 1. Geometry of the acoustooptic interaction in gyrotropic crystals in the Bragg-like intermediate diffraction regime: $\mathbf{k}_0$ is the wave vector of the refracted light wave; $\mathbf{k}_{\mp 1,2}$ are the wave vectors of the diffracted waves; $\mathbf{K}$ is the wave vector of the ultrasound; $\mathbf{e}_0$ and $\mathbf{e}_{\mp 1,2}$ are the polarization unit vectors; δ_1 is the detuning from phase synchronism; $\mathbf{n}_{\mp 1,2}=\mathbf{k}_{\mp 1,2}/|\mathbf{k}_{\mp 1,2}|$; and $\mathbf{n}_0=\mathbf{k}_0/|\mathbf{k}_0|$.

amplitudes, we obtain the system of differential equations of the coupled waves

$$\frac{dA_m}{dz}=i(\Delta_{ma}^{an}+\Delta_{ma}^{ae})A_m+(\rho_m+i\Delta_{mb}^{e})B_m$$

$$+i\chi_{m,m+1}^{a,a}A_{m+1}\exp(-i\delta_m z)$$

$$+i\chi_{m,m+1}^{a,b}B_{m+1}\exp(-i\delta_m z)$$

$$+i\chi_{m,m-1}^{a,a}A_{m-1}\exp(-i\delta_{m-1}z)$$

$$+i\chi_{m,m-1}^{a,b}B_{m-1}\exp(i\delta_{m-1}z),$$

$$\frac{dB_m}{dz}=i(\Delta_{mb}^{an}+\Delta_{mb}^{ae})B_m+(\rho_m+i\Delta_{ma}^{e})A_m$$

$$+i\chi_{m,m+1}^{b,a}A_{m+1}\exp(-i\delta_m z)$$

$$+i\chi_{m,m+1}^{b,b}B_{m+1}\exp(-i\delta_m z)$$

$$+i\chi_{m,m-1}^{b,a}A_{m-1}\exp(-i\delta_{m-1}z)$$

$$+i\chi_{m,m-1}^{b,b}B_{m-1}\exp(i\delta_{m-1}z). \tag{2}$$

Here we have introduced the notation $\Delta_{ma}^{an}=q_m[\mathbf{e}_m(\hat{\varepsilon}-\bar{\varepsilon})\mathbf{e}_m]$, $\Delta_{mb}^{e}=q_m(\mathbf{e}_m\Delta\hat{\varepsilon}_e\mathbf{e}_2)$, $\Delta_{ma}^{ae}=q_m(\mathbf{e}_m\Delta\hat{\varepsilon}_e\mathbf{e}_m)$, $\Delta_{mb}^{an}=q_m[\mathbf{e}_2(\hat{\varepsilon}-\bar{\varepsilon})\mathbf{e}_2]$, $\Delta_{ma}^{e}=q_m(\mathbf{e}_2\Delta\hat{\varepsilon}_e\mathbf{e}_m)$, $\Delta_{mb}^{ae}=q_m(\mathbf{e}_2\Delta\hat{\varepsilon}_e\mathbf{e}_2)$,

$\chi_{m,m+1}^{a,a}=q_{m+1}(\mathbf{e}_m\Delta\hat{\varepsilon}_e\mathbf{e}_{m+1})$, $\chi_{m,m+1}^{a,b}=q_{m+1}(\mathbf{e}_m\Delta\hat{\varepsilon}\mathbf{e}_2)$, $\chi_{m,m-1}^{a,a}=q_{m-1}(\mathbf{e}_m\Delta\hat{\varepsilon}\mathbf{e}_{m-1})$, $\chi_{m,m-1}^{a,b}=q_{m-1}(\mathbf{e}_m\Delta\hat{\varepsilon}^*\mathbf{e}_2)$, $\chi_{m,m+1}^{b,a}=q_{m+1}(\mathbf{e}_2\Delta\hat{\varepsilon}^*\mathbf{e}_{m+1})$, $\chi_{m,m+1}^{b,b}=q_{m+1}(\mathbf{e}_2\Delta\hat{\varepsilon}^*\mathbf{e}_2)$, $\chi_{m,m-1}^{b,a}=q_{m-1}(\mathbf{e}_2\Delta\hat{\varepsilon}^*\mathbf{e}_{m-1})$, $\chi_{m,m-1}^{b,b}=q_{m-1}(\mathbf{e}_2\Delta\hat{\varepsilon}\mathbf{e}_2)$, and $\rho_m=(\mathbf{G}_m\mathbf{k}_m)/|\mathbf{k}_m|$, where $\mathbf{G}_m$ is the gyration vector in the propagation direction of the diffracted wave,[10] $q_m=\omega_m/2c\sqrt{\varepsilon}\,\cos\varphi_m$ (φ_m is the diffraction angle), $\delta_m=|k_{mz}-k_{m+1,z}|$, and an asterisk denotes complex conjugation. The solution of the system of equations (2) describing different AO interaction regimes in gyrotropic anisotropic crystals in an external electric field (the Bragg, Raman–Nath, and intermediate regimes) should be sought using the boundary conditions $A_0(0)=A\cos\Psi$, $B_0(0)=A\sin\Psi$, and $A_m(0)=B_m(0)=0$ for $m\neq 0$, where $\Psi(A)$ is the polarization azimuth (amplitude) of the incident light wave. The interaction of light and ultrasonic waves in gyrotropic anisotropic media must be considered in a principal plane of the crystal.[12] In the case of gyrotropic cubic crystals there is no natural anisotropy, and we should set $\Delta_{ma}^{an}=\Delta_{mb}^{an}=0$ in system of equations (2).

Let us consider the Bragg-like intermediate regime for the diffraction of light on ultrasound.[7,8] It is assumed that light is incident upon the AO interaction region at the Bragg angle ($\varphi=\varphi_B$). In this case we should set $\delta_m=mK^2/k_0$ for $k_0\gg K$ and confine ourselves to the four diffraction orders $A_{-1}(B_{-1})$, $A_0(B_0)$, $A_1(B_1)$, and $A_2(B_2)$.

In the approximation of a weak AO interaction ($|\chi_{q,j}^{s,t}|\ll\rho$ for s, $t=a$, b and q, $j=m$, $m\pm 1$; $|\delta_m|$, $|\delta_{m\pm 1}|\ll\rho$), which can be used to design most AO information-processing devices,[7] the expression for the complex vector amplitude of the diffracted wave $\mathbf{E}_{-1}$ on the exit face $z=l$ of the AO interaction region is given by the relation

$$\mathbf{E}_{-1}=As\{[(\rho\cos\chi l-r\chi\sin\chi l)+i\Delta_e\cos\chi l]$$

$$-\rho_+e^{i\delta_1 l}\}\mathbf{e}_{-1}+Ap\{[(\rho\cos\chi l-r^{-1}\chi\sin\chi l)$$

$$-i\Delta_e\cos\chi l]-\rho_+e^{i\delta_1 l}\}\mathbf{e}_2. \tag{3}$$

Here ρ is the specific rotation in the propagation direction of the diffracted wave; $\rho_{\pm}=(\rho\pm i\Delta_e)$; $\chi=(\rho^2+\Delta_e^2)^{1/2}$, where $\Delta_e=\Delta_{-1a}^{e}\approx\Delta_{-1b}^{e}$; $p=(q_{\parallel}/\chi^2)\exp[-i(\Delta_{-1a}^{an}+\Delta_{-1a}^{ae})l]$; $s=(q_{\perp}/\chi^2)\exp[-i(\Delta_{-1b}^{an}+\Delta_{-1b}^{ae})l]$; and $r=R/Q$, where $Q=(q_{\perp}\sin\Psi+q_0\cos\Psi)$, $R=(q_{\parallel}\cos\Psi+q_0\sin\Psi)$, $q_{\parallel}=iq_{-1}(\mathbf{e}_{-1}\Delta\hat{\varepsilon}^*\mathbf{e}_{-1})$, $q_{\perp}=iq_{-1}(\mathbf{e}_2\Delta\hat{\varepsilon}^*\mathbf{e}_2)$, and $q_0=iq_{-1}(\mathbf{e}_{-1}\Delta\hat{\varepsilon}^*\mathbf{e}_2)$.

It follows from Eq. (3) that the diffracted wave is elliptically polarized.[10] The ellipticity of the diffracted light (the ratio between the minor and major axis of the polarization ellipse) and the polarization azimuth of the diffracted wave (the orientation of the major axis of the polarization ellipse) are determined by the anisotropy of the photoelasticity ($q_{\parallel}\neq q_{\perp}$), the gyrotropy, the electrically induced anisotropy of the crystal in the external electric field ($\Delta_e\neq 0$), and the asymmetry of the diffraction structure (see Fig. 1), which leads to the appearance of an additional phase shift in the diffracted waves ($\delta_1\neq 0$). Using (3), we can easily show that the relative intensity of the diffracted light $\eta_{-1}=|E_{-1}|/|A|^2$

decreases as the specific rotation ρ and the strength of the external electric field increase.

It should be noted that important scientific results on AO diffraction in gyrotropic anisotropic crystals in an external electric field can be obtained in the regime with a strong interaction between light and ultrasound.[6] The number of diffraction orders taken into account in solving system of equations (3) should then satisfy the relation $m \geq 2$. We propose investigating this case of an AO interaction in the intermediate diffraction regime using numerical methods in our forthcoming research.

[1] S. V. Bogdanov, D. V. Petrov, and I. B. Yakovkin, Opt. Spektrosk. **40**, 558 (1976) [Opt. Spectrosc. (USSR) **40**, 316 (1976)].

[2] N. Wakatsuki, N. Chubachi, and Y. Kukuchi, Electron. Commun. Jpn. **58**, 117 (1975).

[3] V. V. Soroka, Fiz. Tverd. Tela **19**, 3327 (1977) [Sov. Phys. Solid State **19**, 1942 (1977)].

[4] V. N. Belyĭ, G. V. Kulak, and V. V. Shepelevich, Opt. Spektrosk. **65**, 636 (1988) [Opt. Spectrosc. (USSR) **65**, 378 (1988)].

[5] G. V. Kulak, Opt. Spektrosk. **78**, 292 (1995) [Opt. Spectrosc. **78**, 261 (1995)].

[6] G. V. Kulak, Opt. Spektrosk. **76**, 1027 (1994) [Opt. Spectrosc. **76**, 920 (1994)].

[7] V. I. Balakshiĭ, V. N. Parygin, and L. E. Chirkov, *Physical Principles of Acoustooptics* [in Russian], Moscow (1985).

[8] V. N. Balakshiĭ, T. G. Kulish, and F. A. Khasan, Opt. Spektrosk. **74**, 1175 (1993) [Opt. Spectrosc. **74**, 696 (1993)].

[9] B. V. Bokut', A. N. Serdyukov, and F. I. Fedorov, Kristallografiya **15**, 1002 (1970) [Sov. Phys. Crystallogr. **15**, 871 (1971)].

[10] F. I. Fedorov, *The Theory of Gyrotropy* [in Russian], Minsk (1976).

[11] Yu. I. Sirotin and M. P. Shaskol'skaya, *Fundamentals of Crystal Physics*, Mir, Moscow (1982).

[12] V. N. Belyĭ, S. N. Kurilkina, and A. G. Khatkevich, Izv. Akad. Nauk Belarusi, Ser. Fiz.-Mat. Nauk., No. 2, 54 (1992).

Translated by P. Shelnitz

Probability model for the scattering of centimeter waves by an object located near a disturbed ocean surface

V. V. Leont'ev

St. Petersburg State University of Electrical Engineering, 197376 St. Petersburg, Russia

A two-dimensional (joint) probability density function is obtained for the amplitude and phase of the backscattered field from a point isotropic reflector located near a statistically rough interface between two media. It is shown that the power of the scattered signal may be enhanced by a factor of more than 16 compared to free space. The probability that fluctuations in the effective scattering area may overshoot a fixed level is estimated. © *1997 American Institute of Physics.* [S1063-7842(97)01809-6]

In the design of data-processing electronic systems (e.g., radar systems) for use in marine environments, much attention is devoted to the development of mathematical models which will provide an adequate description of the interaction of radio waves with various bodies (buoys, beacons, landmarks, ships, etc.) and with the surface of the ocean. This is necessary because the scattering characteristics of objects located near the interface of two bodies differ substantially from those of the same objects in free space. For example, it is well known[1] that various specific fluctuation effects are observed owing to multipath mechanisms for the propagation of electromagnetic waves. In particular, wave backscattering from bodies located near flat[2] or statistically rough[3,4] interfaces may be enhanced. The random character of the underlying surface requires that the methods of statistical electronics be used to analyze the wave processes.

Up to now a large number of probability models have been developed[5] for fluctuations in the amplitude of reflected radar signals, but these models cannot be used to establish a relationship among the characteristics of the fluctuations and the parameters of the probe signal, object, and interface. There are essentially no models which provide adequate descriptions of the fluctuations in the phase shifts of signals reflected from objects over water.

The purpose of this paper is to develop a probability model for the scattering of electromagnetic waves from an object located over water which takes into account the geometric and electrodynamic characteristics of the problem.

The location of an object near a statistically rough ocean surface and the geometric characteristics of this problem are illustrated in Fig. 1. We shall assume that an isotropic source of monochromatic electromagnetic waves is located at point A and that the object being modeled is a fixed isotropic point reflector located at point B. Taking the receiver to be coincident with the radiation source, we consider only the case of monostatic scattering. The distribution of the height (ordinate) H of the ocean waves is Gaussian with zero expectation value ($\langle H \rangle = 0$) and a root-mean-square deviation σ_H. The following relations of the parameters are characteristic for ocean waves: $R_0 \gg h$, $h \gg \lambda$, and $\alpha = (\sigma_H/\lambda)\sin\theta \ll 1$, where α is the generalized Rayleigh parameter.

Using a ray model for the radio wave propagation and relying on the reciprocity theorem (which is valid both for the fields themselves and for the rays of geometric optics), we can write the field backscattered by the reflector (including multipath propagation) in the form

$$\dot{E}_r = \dot{E}_1 + \dot{E}_2 + 2\dot{E}_3 \tag{1}$$

where $\dot{E}_1$ is the field created by the reflector when the waves propagate in free space along the path ABA (Fig. 1), $\dot{E}_2$ is the field propagating along the source–surface–reflector–surface–receiver path with double reflection at the surface, and $\dot{E}_3$ is the field propagating along the source–reflector–surface–receiver or source–surface–reflector–receiver paths with single reflection at the interface.

Eq. (1) has been written in scalar form and, therefore, refers to an arbitrarily chosen linear polarization which is the same for all components. Other wave propagation trajectories can exist; however, their contribution to the resultant field (1) is quite small.

In many problems it is more convenient to consider not the scattered field itself, but a relative quantity associated with it, the complex scattering coefficient which is defined[2] as follows:

$$\dot{A} = 2\sqrt{\pi}R_0(\dot{E}_r/\dot{E}_i)\exp(-ikR_0), \tag{2}$$

where $\dot{E}_i$ is the incident field and $k = 2\pi/\lambda$ is the wave number in free space.

Then, given the geometry of Fig. 1, it can be shown, using Eqs. (1) and (2), that

$$\dot{A} = \sqrt{\sigma_0}[\exp(-i2kh\sin\Theta)$$
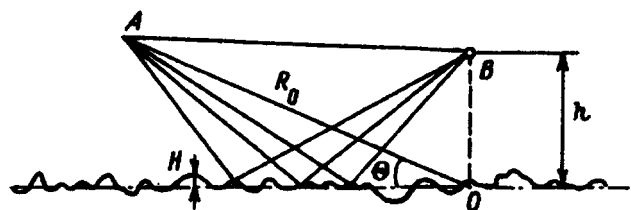$$+ \dot{\Gamma}^2\exp(i2kh\sin\Theta) + 2\dot{\Gamma}], \tag{3}$$

FIG. 1. The geometry of the problem.

where σ_0 is the effective scattering area of the point reflector in free space, h is the height of the reflector above the mean sea level, Θ is the grazing angle, and $\dot{\Gamma}$ is the complex reflection coefficient of the electromagnetic waves at the statistically rough ocean surface.

In accordance with the great number of possible shapes of the ocean surface, the complex reflection coefficient

$$\dot{\Gamma}=x+iy=\rho\,\exp(i\nu) \tag{4}$$

is a random real quantity and, therefore, the complex scattering coefficient (3) is also a random quantity.

In the wave field reflected from a statistically rough ocean surface one can distinguish determinate (or coherent) and random (or incoherent) components,[6–8] each of which can be specified by the generalized parameter α. Then the x and y projections of the complex reflection coefficient (4) are defined as follows:

$$x=x_c+\varepsilon_x, \quad y=y_c+\varepsilon_y,$$

where x_c and y_c are the projections of the coherent complex reflection coefficient $\dot{\Gamma}_c$ and $\varepsilon_x\equiv N(0,\eta)$ and $\varepsilon_y\equiv N(0,\eta)$ are the projections of the incoherent complex reflection coefficient, which are described by two independent Gaussian random variables with zero expectation values and the same root-mean-square deviations η.

The joint (two-dimensional) probability density function of the magnitude ρ and phase shift ν of the complex reflection coefficient (4) at a statistically rough surface with a Gaussian wave-height distribution is given by

$$W(\rho,\nu)=\rho\,\frac{1}{\eta\sqrt{2\pi}}\exp\left[-\frac{(\rho\cos\nu-x_c)^2}{2\eta^2}\right]$$
$$\times\frac{1}{\eta\sqrt{2\pi}}\exp\left[-\frac{(\rho\sin\nu-y_c)^2}{2\eta^2}\right]. \tag{5}$$

In this case the amplitude of the field incident on the object will have a generalized Rayleigh distribution (or Rice distribution).

Introducing the notation

$$\dot{A}/\sqrt{\sigma_0}=\zeta\,\exp(i\varphi) \tag{6}$$

and using the rule[9] for transforming the distributions of two functionally related (Eq. (3)) variables $\dot{A}$ and $\dot{\Gamma}$, we can determine the two dimensional distribution function of the random variables ζ and φ. Because the inverse functions $\rho(\zeta,\varphi)$ and $\nu(\zeta,\varphi)$ are nonunique and have two branches, we have

$$W(\zeta,\varphi)=\sum_{i=1}^{2}W[\rho_i(\zeta,\varphi),\nu_i(\zeta,\varphi)]|J_i(\zeta,\varphi)|, \tag{7}$$

where

$$\rho_{1,2}(\zeta,\varphi)=\sqrt{1+\zeta\mp2\sqrt{\zeta}\cos\left(\frac{\varphi}{2}+\frac{\Phi}{2}\right)},$$

$$\Phi=2kh\sin\Theta,$$

$$\nu_{1,2}(\zeta,\varphi)=\text{arctg}\left[\frac{\sin\Phi\pm\sqrt{\zeta}\sin(\frac{\varphi}{2}-\frac{\Phi}{2})}{-\cos\Phi\pm\sqrt{\zeta}\cos(\frac{\varphi}{2}-\frac{\Phi}{2})}\right],$$

and

$$J_{1,2}(\zeta,\varphi)=\frac{1}{4\sqrt{1+\zeta\mp2\sqrt{\zeta}\cos(\frac{\varphi}{2}+\frac{\Phi}{2})}}$$

are the Jacobians of the transformation.

Substituting Eq. (5) in Eq. (7), we obtain

$$W(\zeta,\varphi)=\frac{1}{8\pi\eta^2}\sum_{i=1}^{2}\exp\left\{\frac{\rho_i^2(\zeta,\varphi)+|\dot{\Gamma}_c|^2-2|\dot{\Gamma}_c|\rho_i(\zeta,\varphi)\cos[\nu_i(\zeta,\varphi)-\varphi_c]}{2\eta^2}\right\}, \tag{8}$$

where $|\dot{\Gamma}_c|^2=x_c^2+y_c^2$ and $\varphi_c=\arctan(y_c/x_c)$.

Equation (8) can be simplified. After some transformations, we have

$$W(\zeta,\varphi)=\frac{\exp\left(-\dfrac{\zeta+\Gamma_0^2}{2\eta^2}\right)}{8\pi\eta^2}$$
$$\times\sum_m\exp\left[\frac{m2\sqrt{\zeta}\cos(\frac{\varphi}{2}+\frac{\Phi}{2})+m2|\dot{\Gamma}_c|\sqrt{\zeta}\cos(\frac{\varphi}{2}-\frac{\Phi}{2}-\varphi_c)}{2\eta^2}\right], \tag{9}$$

where $\Gamma_0=\sqrt{1+|\dot{\Gamma}_c|^2+2|\dot{\Gamma}_c|\cos(\Phi+\varphi_c)}$, with $m=\pm1$.

It is clear that Eq. (9) can be used to study the influence of various parameters of the probe signal source, object, and interface on the fluctuations in the field strength of scattered

electromagnetic waves. Figure 2 shows a plot of the joint probability density function of the magnitude and phase shift of the normalized complex scattering coefficient (6) calculated using Eq. (9) for a horizontally polarized field, $\lambda=3$ cm, $\Theta=1°$, $h/\lambda=125$, and $\alpha=0.18$.

The simultaneous probability density functions for the magnitude and phase shift of the normalized complex scattering coefficient are given by the standard relations

$$W(\zeta)=\int_{-\pi}^{\pi}W(\zeta,\varphi)d\varphi, \tag{10}$$

and

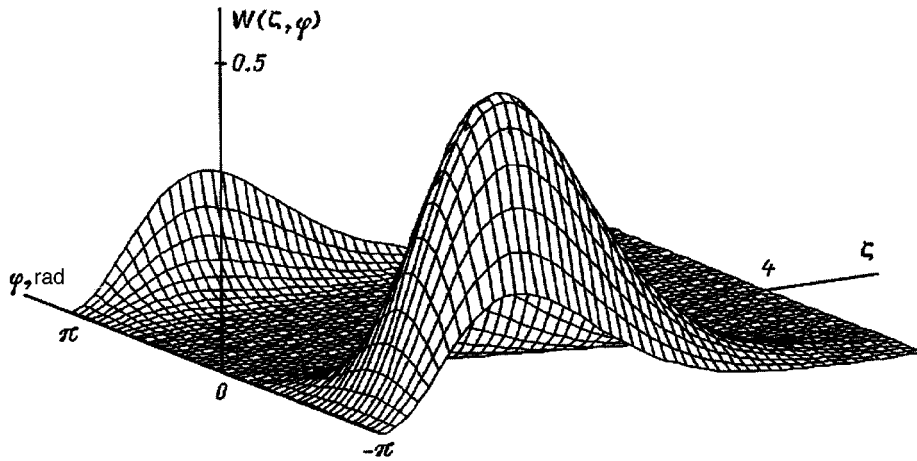$$W(\varphi)=\int_{0}^{\infty}W(\zeta,\varphi)d\zeta. \tag{11}$$

FIG. 2. A two dimensional probability density function of the normalized complex scattering coefficient (6).

Substituting Eq. (9) in Eq. (10) and making some transformations, we obtain

$$W(\zeta) = \frac{1}{2\eta^2} \exp\left(-\frac{\zeta + \Gamma_0^2}{2\eta^2}\right) I_0\left(\frac{2\Gamma_0\sqrt{\zeta}}{2\eta^2}\right), \tag{12}$$

where $I_0(x)$ is the modified Bessel function of zeroth order.

The distribution (12) is known and is referred to as the Rice square.[10,11] The hypothesis that the field reflected from a point target near a statistically rough surface for monochromatic signals is described by this probability density function was apparently first formulated in Ref. 10. However, the distribution (12) has been obtained in the earlier work[10,11] from a simpler mathematical model than in our case. As in the present paper, the authors assume that the incident field at the target has a Rice distribution. As opposed to our model (1), however, there[10,11] the field at the receiver antenna is determined from a two-ray model for the propagation of the reflected wave (with target–receiver antenna and target–surface–receiver antenna rays). The question arises of why the two different models yield the same result (12). This happens because in both cases the target is modeled by a point isotropic reflector for which the intrinsic monostatic and bistatic complex scattering coefficients are the same. It is known[2] that four-ray and two-ray models also yield the same result in the case of a smooth, flat interface between two media.

In some cases it is sufficient to know just the moments of the distribution (12). It can be shown that the kth moment about the origin of the Rice square distribution is given by

$$m_k = (2\eta^2)^k \Gamma(k+1) \exp\left(-\frac{\Gamma_0^2}{2\eta^2}\right) {}_1F_1\left(k+1,1,\frac{\Gamma_0^2}{2\eta^2}\right), \tag{13}$$

where $\Gamma(k+1)$ is the gamma function and ${}_1F_1(k+1,1,\Gamma_0^2/2\eta^2)$ is the confluent hypergeometric or Kammer function.

Using the identity ${}_1F_1(1,2,z) = (z+1)\exp(z)$ and Eq. (13), it is easy to obtain the expectation value

$$m_1 = 2\eta^2\left(\frac{\Gamma_0^2}{2\eta^2} + 1\right) \tag{14}$$

and the second moment about the origin

$$m_2 = (2\eta^2)^2\left[\left(\frac{\Gamma_0^2}{2\eta^2+3}\right)\left(\frac{\Gamma_0^2}{2\eta^2+1}\right) - 1\right] \tag{15}$$

of the absolute value of the normalized complex scattering coefficient (6). We use the notation

$$2\eta^2/\Gamma_0^2 = p. \tag{16}$$

The numerator of Eq. (16) is determined by the dispersion of the incoherent component of the complex reflection coefficient at a rough sea surface and characterizes the fluctuations ("noise") in the irradiating field at the target. The denominator of Eq. (16) is proportional to the constant component of the Rice distribution or the constant component of an irradiating field

$$E = E_0 \sqrt{1 + |\dot{\Gamma}_c|^2 + 2|\dot{\Gamma}_c|\cos(\Phi + \varphi_c)},$$

that includes the "forward" ray and the ray coherently reflected from the interface. In this regard, Eq. (16) can be referred to as the noise/signal ratio. Then, it follows from Eqs. (15) and (14) that

$$\frac{m_2}{m_1^2} = 2 - \frac{1}{(1+p)^2}.$$

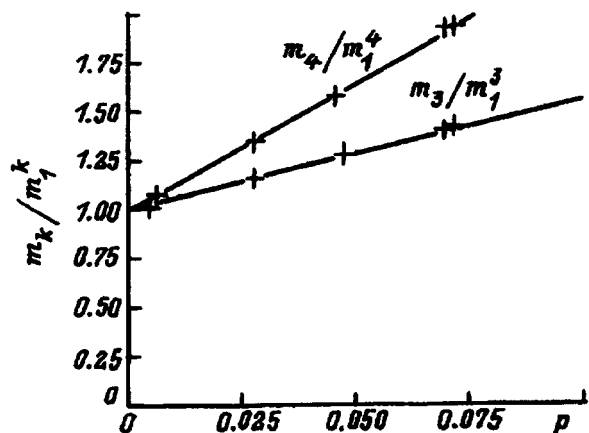In Fig. 3 the smooth curves denote the normalized fourth and

FIG. 3. Comparisons of the normalized third and fourth moments of the Rice-square distribution with experimental data.

third moments of the Rice square distribution (calculated using Eq. (13)) as functions of the noise/signal ratio. In the same figure, the crosses denote the results of laboratory measurements in which a spherical reflector was used as a target.

In the experiment the noise/signal ratio was calculated using the formula

$$p = \sqrt{\frac{1}{2 - (m_2/m_1^2)} - 1},$$

and the moments m_2 and m_1 were estimated from the measurement data. A comparison of the theoretical and experimental results shows that they are in good agreement.

We determine the probability density function for the phase shift of the normalized complex scattering coefficient (6) by substituting Eq. (9) in Eq. (11):

$$W(\varphi) = \frac{\exp\left(-\dfrac{\Gamma_0^2}{2\eta^2}\right)}{2\pi} \left\{ 1 + \frac{\sqrt{\pi}\Gamma_0 \cos(\frac{\varphi}{2} - \beta)}{\sqrt{2\eta^2}} \right.$$

$$\left. \times \exp\left[\frac{\Gamma_0^2 \cos^2(\frac{\varphi}{2} - \beta)}{2\eta^2}\right] f\left[\frac{\Gamma_0 \cos(\frac{\varphi}{2} - \beta)}{\sqrt{2\eta^2}}\right] \right\}, \quad (17)$$

where

$$f(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$$

is the probability integral or error function, and

$$\beta = \arcsin\left[\frac{-\sin(\Phi/2) + |\dot{\Gamma}_c|\sin(\varphi_c + \Phi/2)}{\Gamma_0}\right].$$

Figure 4 shows the probability density functions of the magnitude and phase shift of the normalized complex scattering coefficient (6) of a point reflector located above a rough ocean surface calculated using Eqs. (12) and (17) for different heights above the mean sea level and for different parameters α.

It can be shown that the probability density function of the effective scattering area σ of a reflector near a statistically rough surface is given by the expression

$$W(\sigma) = \frac{1}{4\eta^2\sqrt{\sigma\sigma a_0}} \exp\left(-\frac{\Gamma_0^2}{2\eta^2}\right)$$

$$\times \exp\left(\frac{-\sqrt{\sigma}}{2\eta^2\sqrt{\sigma_0}}\right) I_0\left(\frac{\Gamma_0 \sqrt[4]{\sigma}}{\eta^2 \sqrt[4]{\sigma_0}}\right). \quad (18)$$

Figure 5 shows probability density functions of the effective scattering area of a reflector calculated using Eq. (18) for different values of the parameter α. In the calculations it has been assumed that the effective scattering area of the reflector itself (in free space) is $\sigma_0 = 10$ m^2, the wavelength of the irradiating field is $\lambda = 3$ cm, the polarization is horizontal, and the grazing angle is $\Theta = 1°$. The parameter α was changed by varying the root-mean-square deviation σ_H of the wave height, and an increase in α corresponds to enhanced sea swell. The height of the reflector above the level of the ocean was $h = 3.87$ m, this value having been chosen so that for the case of a smooth ocean surface the effective
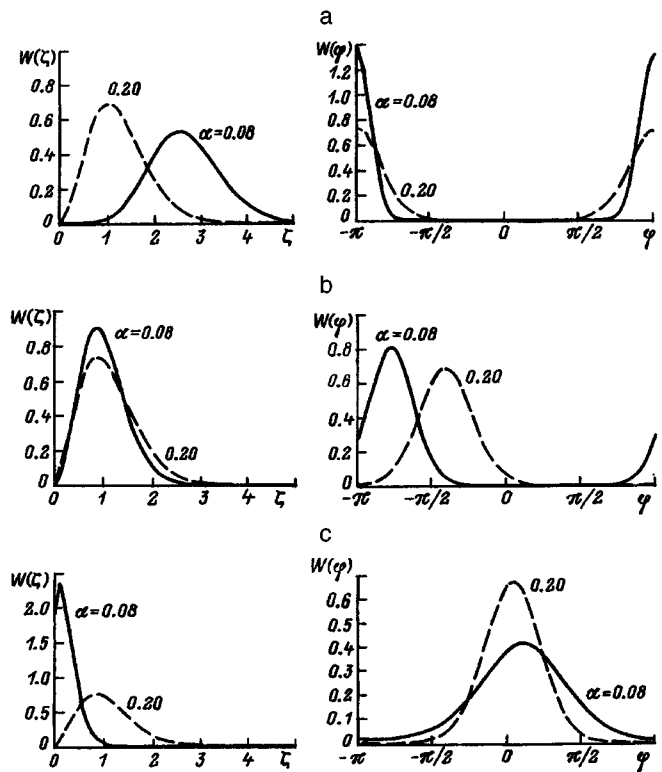
FIG. 4. Probability density functions of the magnitude and phase shift of the normalized complex scattering coefficient (6). $h/\lambda = 129$ (a), 120 (b), 114 (c).

scattering area of the reflector (relative to its effective scattering area in free space) would be increased by a factor of 16 (12 dB).[2] Thus, for a small sea swell ($\alpha = 0.02$) the probability density function of the effective scattering area is concentrated near $\sigma = 160$ m^2. As the sea swell increases, the coherent component of the complex reflection coefficient in the wave field reflected from a statistically rough ocean surface decreases and this leads to a reduction in the effective scattering area of the reflector.
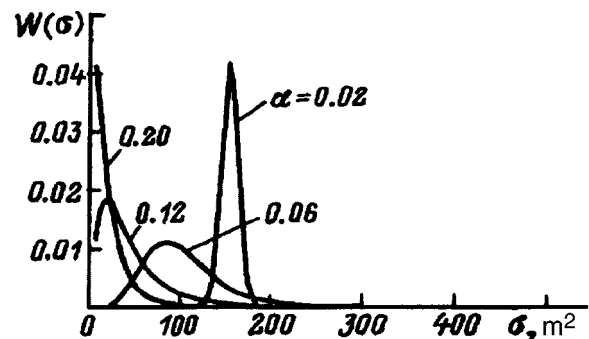
The moments of the distribution (18) are given by

FIG. 5. Variation in the probability density function of the effective scattering area of a reflector near the ocean surface.
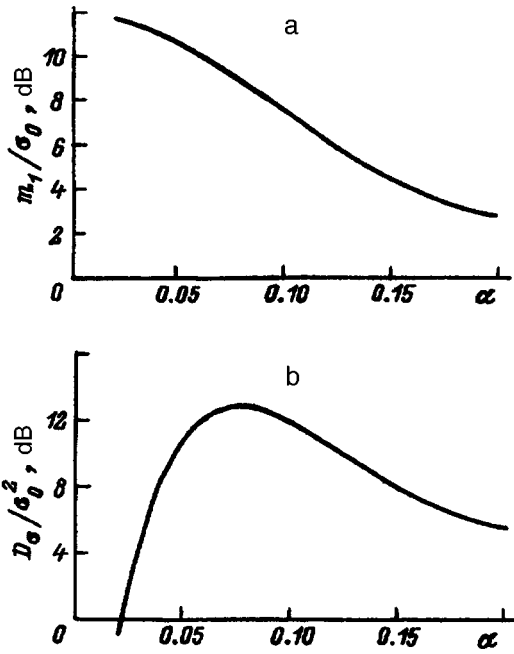
FIG. 6. Normalized mathematical expectation (a) and dispersion (b) of the effective scattering area of a reflector located near the ocean surface.
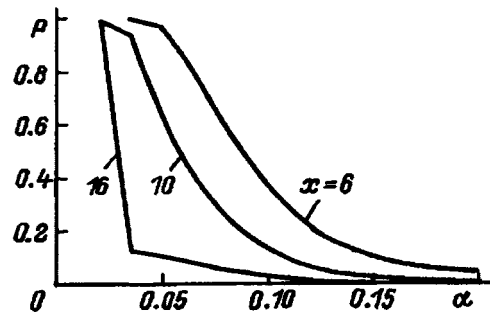

FIG. 7. The change in the probability that the fluctuations in the effective scattering area of a reflector will exceed a fixed level $\sigma = x\sigma_0$ as the generalized parameter α is varied.

The results of a numerical integration of Eq. (23) are shown in Fig. 7. There is a special case in which the integral in Eq. (23) can be calculated analytically.[12] When $\Gamma_0 /\sqrt[4]{x} = 1$, Eq. (23) takes the form

$$P = 1 - \frac{1}{2} \exp\left(-\frac{\Gamma_0^2 + \sqrt{x}}{2\eta^2} \right)\left[\exp\left(\frac{\sqrt{x}}{\eta^2} \right) - I_0\left(\frac{\sqrt{x}}{\eta^2} \right) \right].$$

The fluctuations in the radar scattering characteristics of a point isotropic reflector near a statistically rough interface of two media are, therefore, non-Gaussian. The appearance of non-Gaussian reflections in the backward direction can be explained by the complicated interaction of an electromagnetic wave with a ''reflector plus interface'' system.

$$m_k = (2\eta^2)^{2k}\Gamma(2k+1)\exp\left(-\frac{\Gamma_0^2}{2\eta^2} \right)\sigma_0^k \, {}_1F_1$$

$$\times \left(2k+1, 1, \frac{\Gamma_0^2}{2\eta^2} \right). \tag{19}$$

Equation (19) implies that the mathematical expectation for the effective scattering area is

$$m_1 = (2\eta^2)^2(z^2 + 4z + 2)\sigma_0 \tag{20}$$

and the dispersion in the effective scattering area is

$$D_\sigma = (2\eta^2)^4(8x^3 + 52z^2 + 80z + 20)\sigma_0^2, \tag{21}$$

where $z = \Gamma_0^2/(2\eta^2)$.

The normalized mathematical expectation (20) and dispersion (21) of the effective scattering area of the reflector are shown as functions of the generalized parameter α in Fig. 6.

It is clear from Fig. 5 that the effective scattering area of a reflector near a statistically rough surface can be substantially greater than its intrinsic effective scattering area measured in free space. As a quantitative characteristic of the backscattering enhancement it is appropriate to estimate the probability that the fluctuations in the effective scattering area of the reflector will exceed a fixed level $\sigma = x\sigma_0$,

$$P = \int_{x\sigma_0}^{\infty} W(\sigma)d\sigma. \tag{22}$$

Substituting Eq. (18) in Eq. (22), we obtain

$$P = 1 - \frac{\sqrt{x}}{\eta^2} \exp\left(-\frac{\Gamma_0^2}{2\eta^2} \right)\int_0^1 t \exp(-\tau t^2)I_0\left(2\frac{\Gamma_0}{\sqrt[4]{x}}\tau t \right)dt, \tag{23}$$

where $\tau = \sqrt{x}/(2\eta^2)$.

[1] Yu. A. Kravtsov and A. I. Saichev, Ukr. Fiz. Zhur. **137**, 501 (1982).

[2] A. Yu. Andreev, L. I. Bogin, V. O. Kobak, and V. V. Leont'ev, Radiotekh. Elektron. **35**, 734 (1990).

[3] A. G. Vinogradov, A. S. Gurvich, S. S. Kashkarov et al., Invention (Certificate No. 359), Byull. Izobret. (1989), No. 21.

[4] V. U. Zavorotnyĭ and V. I. Tatarskiĭ, Dokl. Akad. Nauk SSSR **265**, 608 (1982).

[5] V. M. Shlyakhin, Radiotekh. Elektron. **32**, 1793–1817 (1987).

[6] C. I. Beard, IRE Trans. Antennas Propag. **AP-9**, 470–483 (1961).

[7] C. I. Beard, I. Katz, and L. M. Spetner, IRE Trans. Antennas Propag. **AP-4**, 162–167 (1956).

[8] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon, New York (1963).

[9] B. R. Levin, *Theoretical Foundations of Statistical Electronics* [in Russian], Sov. Radio, Moscow (1974), Vol. 1.

[10] C. J. Baker and K. D. Ward, AGARD Conf. No. 419 (1987), pp. 25-1 – 25-4.

[11] K. D. Ward, C. J. Baker, and S. Watts, IEE Proc. F (Radar and Signal Processing) **137** (2), 51–62 (1990).

[12] I. S. Gradshtein and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, New York (1965) [Russian original, Nauka, Moscow, 1971].

Translated by D. H. McNeill

Transition radiation measurements at distances from the transition point comparable to the formation length

A. V. Serov and B. M. Bolotovskiĭ

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

The spatial distribution of the electromagnetic field excited by a relativistic particle crossing the surface of a metal is studied. It is shown that the field of the uniformly moving charge must also be taken into account during measurements at distances comparable to the path length for formation of the radiation. Expressions describing the effect of the self-field of the charge on the transition radiation field are derived. © *1997 American Institute of Physics.*
[S1063-7842(97)01909-0]

INTRODUCTION

The simplest type of transition radiation arises when a particle in uniform and straight-line motion crosses the interface between two media and has been studied in some detail, both theoretically and experimentally.[1,2] A special case of this type of radiation is the emission generated by a relativistic charged particle escaping from or incident on a conductor. In this case the spectrum of the transition radiation is flat over a wide band from radio to well beyond optical frequencies. The angular distribution of the radiation is independent of the distance between the point where the particle escapes the metal (or is incident on the metal) and the radiation detector. At arbitrary distances from the interface, the energy emitted by an escaping particle is equal to zero in the direction of the particle's velocity and is greatest at an angle $\theta = \gamma^{-1}$ to the velocity, where γ is the relative energy of the particle. When a relativistic particle is incident on a metal there is no emission at an angle of $\theta = \pi$ and the emission is greatest at an angle of $\theta = \pi - \gamma^{-1}$.

The electromagnetic field generated by a charge escaping from a metal or incident on it is the sum of the field of the uniformly moving charge and the radiation field. The intrinsic field of the charge propagates at the speed of the particle and the radiation field, at the speed of light in the given medium. Usually the energy of the transition radiation is calculated as the energy of the radiation field. A relativistic particle moving in a straight line at a constant velocity $v = c$, however, produces electric and magnetic fields at the observation point that are almost equal and mutually perpendicular. In terms of their structure, these fields are indistinguishable from the emission (radiation) fields. Thus, the emitted energy is determined by separating the radiation field and the intrinsic field of the charge. Separation of the fields is taken to mean a temporal separation[1] such that a pulse of radiation arrives at the observation point much earlier or much later than the time of flight of the charge by the radiation detector; then the field dragged along with the particle and the emitted field will also be spatially separated from one another.

If we are measuring the field generated by a particle incident on a metal (Fig. 1a) and the radiation detector P is positioned closer to the particle's trajectory than to the point where the particle strikes the metal, then the self-field of the particle will be detected before the emission field. The delay between the arrival times at the observation point of the peak in the intrinsic field of the moving charge and of the pulse of electromagnetic radiation is roughly

$$\Delta t \simeq \frac{2L}{c}, \tag{1}$$

where L is the distance from the observation point to the metal.

If we are measuring the field generated by a particle escaping from a metal (Fig. 1b), then the radiation field arrives at the detector before the self-field of the particle. Then the delay is

$$\Delta t \simeq \frac{L}{2\gamma^2 v}. \tag{2}$$

In this case, the time delay arises as the particle moves over a distance from the point where it leaves the metal that is equal to the formation length.[3] The formation length l for radiation with wavelength λ is

$$l \simeq \frac{\lambda \gamma^2}{1 + \gamma^2 \theta^2}. \tag{3}$$

The conditions for separation (especially condition (2)) cannot always be fulfilled in practice. When a relativistic particle leaves a metal the time delay is short, much shorter than the time for the particle to move from the interface to the observation point. (For $L = 1$ m, $\Delta t = 3 \times 10^{-9}/2\gamma^2$.) Thus, if the resolving time of the detector is not short enough, then its readout will be determined by the field of the charge, as well as by the electromagnetic radiation field. This is even more true of the case when a spectral device is used for detection. Then the readout is always determined by the spectrum of the self-field of the particle, as well as by the spectrum of the radiation field.

In this paper we examine the effect of the self-field of the charge on the spectral and angular characteristics of the field generated by a relativistic charged particle escaping from a metal.
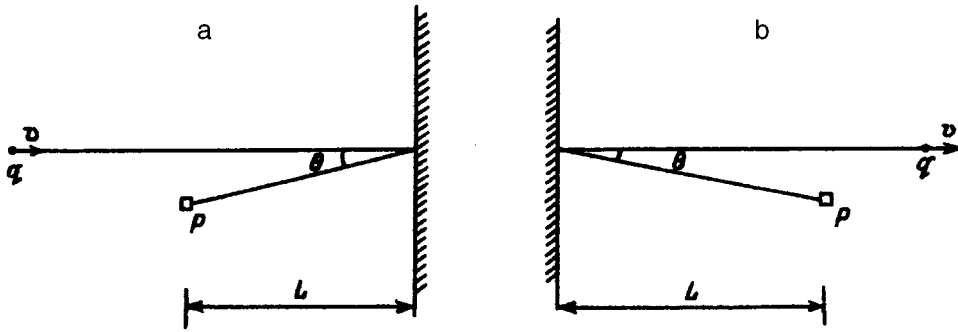
FIG. 1. Generation of transition radiation by a charge incident on a metal (a) and escaping from a metal (b).

BASIC EQUATIONS

Let us consider the field which develops when a particle escapes from a metal perpendicular to its surface. We assume that the metal is an ideal conductor. The particle escapes from the point $x=0$, $y=0$ in the positive x direction with velocity v. The generated field can be represented as the superposition of the fields of two instantaneously launched charged particles (Fig. 2). One of the particles is the real charge q, while the second is the image of this charge. The image has a charge opposite in sign and equal in magnitude to that of the escaping particle. The velocity of the image is equal in magnitude and opposite in sign to the velocity of the escaping particle, so that the position of the image is given by $x=-vt$. Evidently, if a plane is drawn through the point $x=0$ perpendicular to the x axis, then the field lines of the total electric field created by the charge and the image are perpendicular to this plane. Thus, the same boundary conditions hold at the plane $x=0$ as at a metal. Consequently, in this case the transition radiation problem reduces to finding the radiation emitted upon the instantaneous launching of two charges of equal magnitude and opposite sign from a single point in opposite directions.[4]

In this case the field has the following spatial–temporal structure. Let us consider the hemisphere lying in the half space $x>0$ with its center located at the point where the charge escapes and with a radius of $r=ct$. Outside this hemisphere the field equals zero. Inside the hemisphere, the field is given by the superposition of the fields of the uniformly
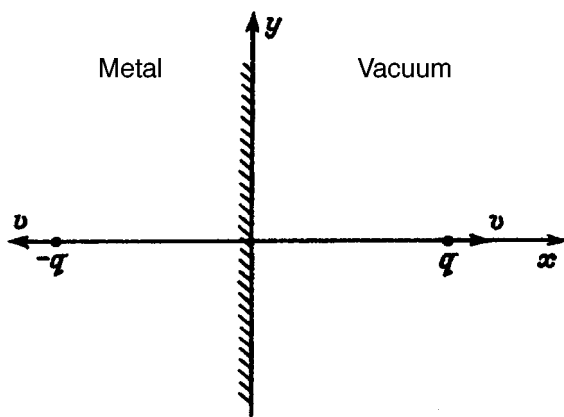

FIG. 2. Representing the field formed by a charge escaping from a metal as the superposition of the fields from two instantaneously launched charged particles.

moving charge and its image. The field lines lying on the surface of the sphere determine the radiation field.

Let us assume that the radiation detector P lies at the point x_p, y_p. The radiation field generated by instantaneous launch of the charge and the image charge is given by[3]

$$E_x^r = \frac{q}{r_p}\,\delta(r_p-ct)\left\{\frac{\beta\,\sin^2\theta}{1-\beta\,\cos\theta}+\frac{\beta\,\sin^2\theta}{1+\beta\,\cos\theta}\right\}$$
$$= \frac{q}{r_p}\,\delta(r_p-ct)\frac{2\beta\,\sin^2\theta}{1-\beta^2\cos^2\theta}, \tag{4}$$

and

$$E_y^r = \frac{q}{r_p}\,\delta(r_p-ct)\left\{\frac{\beta\,\sin\theta\,\cos\theta}{1-\beta\,\cos\theta}+\frac{\beta\,\sin\theta\,\cos\theta}{1+\beta\,\cos\theta}\right\}$$
$$= \frac{q}{r_p}\,\delta(r_p-ct)\frac{2\beta\,\sin\theta\,\cos\theta}{1-\beta^2\cos^2\theta}, \tag{5}$$

where $r_p=\sqrt{x_p^2+y_p^2}$ and $\beta=v/c$ is the relative velocity of the particle.

The delta function with argument (r_p-ct) accounts for the fact that the radiation field differs from zero only on a sphere $r=ct$ propagating at the speed of light.

The field created by the uniformly moving particle and its image is given by

$$E_x^q = q(1-\beta^2)\left[\frac{x_p-vt}{[(1-\beta^2)y_p^2+(x_p-vt)^2]^{3/2}}\right.$$
$$\left.+\frac{x_p+vt}{[(1-\beta^2)y_p^2+(x_p+vt)^2]^{3/2}}\right], \tag{6}$$

and

$$E_y^q = q(1-\beta^2)\left[\frac{y_p}{[(1-\beta^2)y_p^2+(x_p-vt)^2]^{3/2}}\right.$$
$$\left.+\frac{y_p}{[(1-\beta^2)y_p^2+(x_p+vt)^2]^{3/2}}\right]. \tag{7}$$

As it expands, the sphere $r=ct$ passes through the observation point. At a time $t=\sqrt{x_p^2+y_p^2}/c$ the field at the observation point changes from zero to the radiation field given by Eqs. (4) and (5). The spectral expansions of the components of the radiation field have the form

$$E_x^r(\omega)=\frac{q}{\pi c r_p}\frac{\beta\,\sin^2\theta}{1-\beta^2\cos^2\theta}\exp\left(i\,\frac{\omega}{c}\,r_p\right), \tag{8}$$

$$E_y^r(\omega) = \frac{q}{\pi c r_p} \frac{\beta \sin\theta \cos\theta}{1 - \beta^2 \cos^2\theta} \exp\left(i\,\frac{\omega}{c}\, r_p \right). \qquad (9)$$

At later times ($t > \sqrt{x_p^2 + y_p^2}/c$) the field at the observation point equals the sum of the fields of two charges moving uniformly along the x axis, one of which is the real q moving at velocity v from the point with coordinates $x = \beta\sqrt{x_p^2 + y_p^2}$, $y = 0$ while the other is its image $-q$ moving at velocity $-v$ from the point with coordinates $x = -\beta\sqrt{x_p^2 + y_p^2}$, $y = 0$. The time dependence of the field is given by Eqs. (6) and (7). In the following we shall consider the case where the charged particle escapes from the metal and the observation point is in the vacuum a short distance from its trajectory. Then it is possible to neglect the field of the image and just include the first terms in Eqs. (6) and (7).

It is clear that the E_x component of the field changes sign at the time when the charge appears at the point $x = x_p$, $y = 0$. The time integral of this field component is zero. The field pulse E_x at $t \cong x_p/v$ is close to sinusoidal in shape with a frequency $\omega \sim \gamma v/y_p$; thus, its spectrum consists of a narrow range of frequencies near $\omega = \gamma v/y_p$. The E_y component of the electromagnetic field consists of a bell-shaped pulse with amplitude $E_y \sim q\gamma/y_p^2$ and characteristic width $\tau \sim y_p/\gamma v$. The spectrum of the pulse contains all frequencies up to $\omega \sim 1/\tau$. The Fourier transforms of the components of the electromagnetic field have the form[5]

$$E_x^q(\omega) = -i\,\frac{q\omega}{\pi v^2 \gamma^2}\, K_0\left(\frac{\omega y_p}{\gamma v} \right) \qquad (10)$$

and

$$E_y^q(\omega) = \frac{q\omega}{\pi v^2 \gamma}\, K_1\left(\frac{\omega y_p}{\gamma v} \right), \qquad (11)$$

where K_0 and K_1 are the modified Bessel functions of imaginary arguments (MacDonald functions).

The above equations describe the electromagnetic fields and their spectra when the radiation and intrinsic fields of a uniformly moving charge are completely separated in time. The concept of complete separation of the fields during transition radiation is definitely an idealization. In practice, the pulses formed by the radiation field and the self-field of the particle overlap to a greater or lesser extent, so that the spatial–temporal and spectral–angular structures of the field depend significantly on the distance between the receiver and the metallic surface. A practical case is closer to the ideal when the radiation detector lies further from the transition point of the charge ($L \gg 1$) and the angle at which the observation is made is smaller ($\theta \ll 1/\gamma$).

THE TRANSITION RADIATION FIELD

Figure 3 shows the time variations in the E_y components of the electric field at different points in space. These curves were constructed using Eqs. (5) and (7) for an electron escaping a metal at an energy $\gamma = 150$. The distance from the metal surface to the measurement plane is $x_p = 3$ m, while the transverse distances from the detector to the particle trajectory are $y_p = 5$, 10, and 15 mm (curves 1–3, respectively). The initial jump in the electric field, which is described by
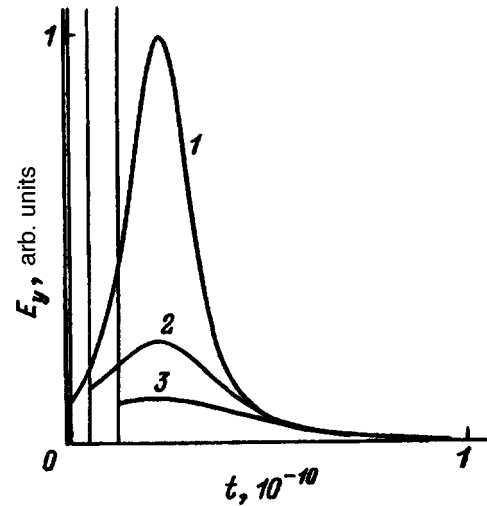

FIG. 3. Time dependence of the E_y component of the electric field. $\gamma = 150$, $x_p = 3$ m; y_p (mm): 5 (1), 10 (2), 15 (3).

the delta function in Eq. (5), corresponds to the moment the radiation field of the electron arrives at the observation point. At later times the variation is determined by the self-field of the uniformly moving particle. In Fig. 3 the time $t = 0$ corresponds to the time of arrival of the radiation field at a point lying on the intersection of the particle trajectory and the measurement plane $x = L$ and has coordinates $x_p = 3$ m and $y_p = 0$. It is clear from the figure that increasing the transverse coordinate of the receiver will cause both a drop in the amplitude of the pulse formed by the self-field of the uniformly moving electron and a change in the pulse shape, reducing its duration and increasing its asymmetry. It can be shown that at a receiver located at an angle $\theta = \gamma^{-1}$ the self-field will form only half of the bell shaped pulse, with the left half cut off.

The spectrum of the total field can be obtained by expanding the curve shown in Fig. 3 in a Fourier integral. It may be noticed at once that the spectral composition of the signal (especially at frequencies whose period is commensurate with the pulse duration of the intrinsic field of the particle) will depend strongly on the transverse coordinate of the receiver.

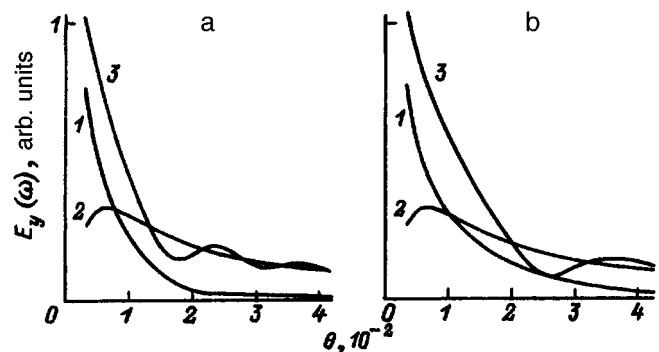Figure 4 shows the angular distribution of the E_y com-

FIG. 4. Angular distribution of the spectral components $E_y(\omega)$. $\gamma = 150$, $x_p = 3$ m; (1) $E_y^q(\omega)$, (2) $E_y^r(\omega)$, (3) $E_y^{r+q}(\omega)$.

ponent of the field for two wavelengths: $\lambda = 1$ (a) and 2 mm (b). Curve 1 shows the distribution of the self-field of the uniformly moving charge, curve 2, that of the radiation field, and curve 3, the combined field. It is clear that even at fairly large distances from the metal surface (the measurement plane is 3 m away), the distribution of the total field differs greatly from that of the radiation field. First, these curves have several peaks, whose amplitude and position depend on wavelength, while the radiation field has a single peak at an angle of $\theta = \gamma^{-1}$ for all wavelengths. Second, in the space near the particle trajectory at angles $\theta < \gamma^{-1}$, the magnitude of the field is determined mainly by the self-field of the particle.

We now select a criterion for defining the boundary of the region within which the intrinsic field of the charge has an effect. We shall compare the radiation and intrinsic fields at those points in space which lie in the direction of the maximum intensity of the radiation field, i.e., at an angle $\theta = \gamma^{-1}$ to the particle velocity at the point where it leaves the metal. If $E^q(\omega)$ and $E^r(\omega)$ are of the same order of magnitude in this direction, then we can say that for angles $\theta < \gamma^{-1}$ the self-field of the particle has a significant effect on the spectrum of the electromagnetic field.

We shall compare only the $E_y(\omega)$ components of the Fourier harmonics. For relativistic particles ($\gamma \gg 1$) Eq. (9) yields an amplitude for the Fourier harmonic of the radiation field in the direction $\theta = \gamma^{-1}$ of

$$E_y^r(\omega) = \frac{q}{2\pi c r_p} \frac{\beta}{\gamma\left(1 - \beta + \frac{1}{2\gamma^2}\right)} \approx \frac{q\beta\gamma}{2\pi c L}. \tag{12}$$

In this derivation it was assumed that $\sin\theta \approx \theta$, $\cos\theta \approx 1 - (\theta^2/2)$, and $r_p \approx L$. In order to obtain a expression for the Fourier component of the intrinsic field of the charge, it is necessary to calculate the integral

$$E_y^q(\omega) = \frac{q y_p}{2\pi\gamma^2 v^3} \int_0^\infty \frac{\exp(i\omega t_1)}{\left(\frac{y_p^2}{\gamma^2 v^2} + t_1^2\right)^{3/2}} dt_1. \tag{13}$$

In this integral we have used only the first term on the right of Eq. (7) and made the substitution $t_1 = t - (x_p/v)$. The second term in Eq. (7) is much smaller than the first and we neglect it. The lower limit of integration is the time at which the radiation field reaches the radiation receiver. Since the receiver lies at an angle $\theta = \gamma^{-1}$, the charge lies at the point x_p at this time, i.e., at the least distance from the receiver.

In fact, the field of the uniformly moving charge will be picked up by the receiver immediately after the radiation front passes, at time $t = r_p/c$. At this time the particle will be at the point $x = vt = vr_p/c$. In our case, the longitudinal coordinate x_p of the receiver and the distance r_p from it to the escape point are related by the formula $r_p = x_p/\cos(1/\gamma)$. Given that $\gamma = (1 - \beta^2)^{1/2}$ and taking $\cos(1/\gamma) \approx 1 - (1/2\gamma^2)$, we obtain the coordinate of the particle at time $t = r_p/c$. It is $x = x_p$. After integrating, we have
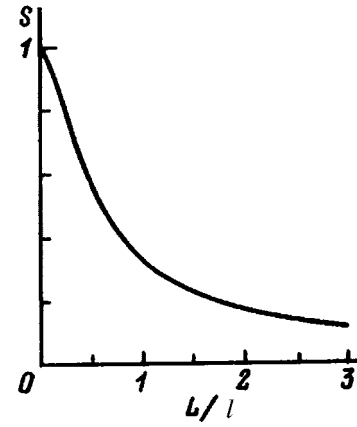

FIG. 5. The ratio of the harmonics of the self-field $E_y^q(\omega)$ and radiation field $E_y^r(\omega)$ as a function of the reduced receiver distance.

$$E_y^q(\omega) = \frac{q\omega}{2\pi\gamma v^2}\left\{ K_1\left(\frac{\omega y_p}{\gamma v}\right) \right.$$
$$\left. + i\frac{\pi}{2}\left[-I_1\left(\frac{\omega y_p}{\gamma v}\right) + L_1\left(\frac{\omega y_p}{\gamma v}\right) + \frac{2}{\pi} \right] \right\}, \tag{14}$$

where K_1 is the Macdonald function and I_1 and L_1 are the modified Bessel and Struve functions.

The argument of the special functions is $\omega y_p/\gamma v$. Since in this case $y_p = L/\gamma$, while the formation length for the radiation at wavelength λ is $l = \lambda\gamma^2/2$, the argument of the functions can be written in the form $\omega y_p/\gamma v = \pi L/l$, i.e., as the ratio of the metal-to-receiver distance to the formation length for the radiation.

The ratio of the harmonics of the intrinsic field $E_y^q(\omega)$ and radiation field $E_y^r(\omega)$ is given by

$$S = \left| \frac{E_y^q(\omega)}{E_y^r(\omega)} \right| = \left| \pi\frac{L}{l}\left\{ K_1\left(\pi\frac{L}{l}\right) \right. \right.$$
$$\left. \left. + i\frac{\pi}{2}\left[-I_1\left(\pi\frac{L}{l}\right) + L_1\left(\pi\frac{L}{l}\right) + \frac{2}{\pi} \right] \right\} \right|, \tag{15}$$

which implies that it depends only on the ratio of the receiver distance L to the radiation formation length l. the function $S(L/l)$ is plotted in Fig. 5, from which we see that even when the distance L exceeds the formation length l by a factor of 3, the amplitude of the harmonic of the self-field of the particle is still 10% of that of the radiation field.

CONCLUSION

The classical expressions for transition radiation fields[4] have been obtained under the assumption that the radiation field does not interfere with the self-field which is dragged along with the charge. The above remarks imply that, depending on the location of the detection device, interference between these two fields can have a significant effect on the result and this circumstance must be taken into account in real measurements.

[1] V. L. Ginzburg and V. N. Tsytovich, *Transition Radiation and Transition Scattering* [in Russian], Nauka, Moscow (1984).

[2] *Bibliography of Papers on Transition Radiation from Charged Particles (1945-1982)* [in Russian], Erevan Physical Institute, Erevan (1983).

[3] B. M. Bolotovskiĭ, Tr. Fiz. Inst. Akad. Nauk SSSR **140**, 95 (1982).

[4] V. L. Ginzburg and I. M. Frank, Zh. Éksp. Teor. Fiz. **16**, 15 (1946).

[5] J. D. Jackson, *Classical Electrodynamics*, Wiley, New York (1962) [Russian trans., Mir, Moscow, 1965].

Translated by D. H. McNeill

Modeling the wave parameters of a narrow slotted transmission line based on a superconducting film

O. G. Vendik, I. S. Danilov, and S. P. Zubko

St. Petersburg State University of Electrical Engineering, 197376 St. Petersburg, Russia

An analytical expression for calculating the wave parameters of a narrow-slot transmission line is found in the quasistatic approximation. A two-fluid model is used for analyzing processes in a superconducting film of thickness comparable to the London penetration depth. The wave parameters of the slotted line are calculated under assumptions about the current distribution near the edges of the slot which are analogous to those used previously for analyzing microstrip and coplanar lines. © *1997 American Institute of Physics.* [S1063-7842(97)02009-6]

INTRODUCTION

High-temperature superconducting (HTSC) slotted transmission lines have recently attracted considerable attention. This has happened, in particular, because of the combined use of HTSC slotted line with a Josephson junction in microwave squids.[1,2] In order to impedance match this type of system, it is necessary to have a minimum wave impedance in the slotted line. This is achieved by reducing the slot width to micron sizes. The existing analytical descriptions of the parameters of slotted lines[3,4] are restricted to slot widths exceeding 2% of the substrate thickness. For slots that are narrow enough, it is possible to use a quasistatic approximation for calculating their linear parameters and to obtain simple analytical expressions. Because of the small, but finite resistance of the conductor, HTSC transmission lines have Ohmic losses, and these can be significant if the slotted line is sufficiently narrow. The method employed in this article has already been used to calculate the wave parameters in HTSC microstrip[5] and coplanar[6] transmission lines. The computational results are in good agreement with experimental data, which indicates that this approach can be applied to other types of flat transmission lines. In this paper we present analytical expressions which make it possible to calculate the wave parameters of HTSC slotted lines with given parameters.

QUASISTATIC CALCULATION OF THE LINEAR PARAMETERS OF A NARROW SLOTTED LINE

The transverse cross section of the line is shown in Fig. 1a. The substrate has a thickness h and width D. A superconducting film of thickness d is deposited on the substrate and the slot width is w. The effective permittivity of the line, which determines the phase velocity of a wave in a line with an ideally conducting coating, is taken to be

$$\varepsilon_{\text{eff}} = (\varepsilon_r + 1)/2, \tag{1}$$

where ε_r is the relative permittivity (dielectric constant) of the substrate material.

A comparison with electrodynamic calculations of ε_{eff} for slotted lines[3,4] shows that as w/h approaches zero, the actual value of $\varepsilon_{r,\text{eff}}$ approaches that given by Eq. (1) (Fig. 2a).

Let L_1 and C_1 denote the linear and capacitance per unit length of the line (in the equivalent circuit, the inductances are in series, while the capacitances are in parallel). We use the fact that[3]

$$L_1 = (Z_0\sqrt{\varepsilon_{\text{eff}}})/c, \quad C_1 = \sqrt{\varepsilon_{\text{eff}}}(Z_0 c)^{-1}, \tag{2}$$

where Z_0 is the wave impedance of the line and c is the speed of light in vacuum.

In order to calculate the capacitance per unit length of a slot in a shield of width D, we have used the approximation of partial capacitances,[7,8]

$$C_1 = \varepsilon_0 [(\varepsilon_r - 1) - F(k_1) + F(k_2)], \tag{3}$$

where ε_0 is the vacuum permittivity. $F(k_1)$ gives the contribution of the substrate to the slot capacitance, and $F(k_2)$ gives the contribution of the surrounding space.
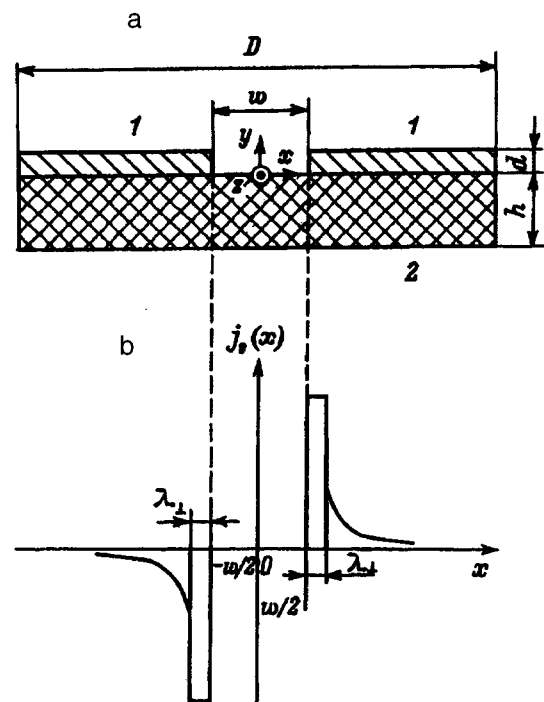

FIG. 1. Transverse cross section of a slotted line. (a) A sketch of the transverse cross section of the slotted line: (*1*) conducting film, (*2*) dielectric substrate; (b) the distribution of the surface current density in the transverse cross section of a superconducting slotted line.
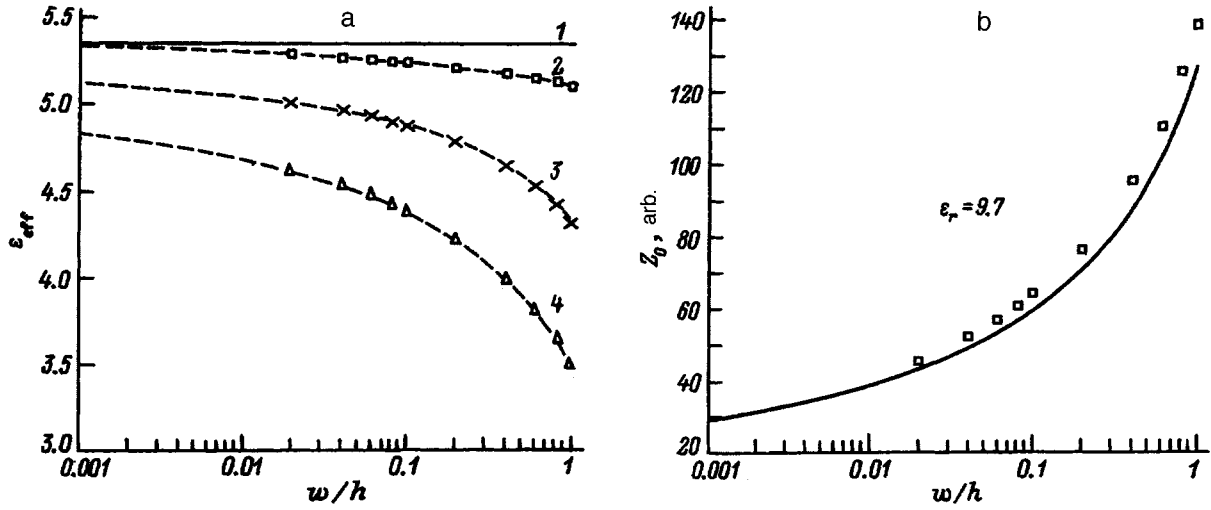
FIG. 2. Wave parameters of a slotted line as a function of the ratio w/h. (a) Effective dielectric permittivity ε_{eff} as a function of the ratio w/h for $\varepsilon_r = 9.7$: (1) $\varepsilon_{\text{eff}} = (\varepsilon_r + 1)/2$; (2–4) data from Ref. 4; (2) $h/\lambda_0 = 0.075$, (3) $h/\lambda_0 = 0.05$, (4) $h/\lambda_0 = 0.025$. The dashed curve is an "intuitive" extrapolation of the data of Ref. 4 for $w/h \leqslant 0.02$. (b) The wave impedance as a function of w/h; the smooth curve is the impedance calculated using the quasistatic approximation and the points are data from Ref. 4.

To find $F(k_1)$ and $F(k_2)$ we used a conformal mapping[7,8] which gives the following expressions for the $F(k_i)$:

$$F(k_i) = \frac{K(k_i^*)}{K(k_i)}, \tag{4}$$

where $K(k)$ is the complete elliptic integral of the first kind, and $k_i^* = \sqrt{1 - k_i^2}$.

For $w \ll h \ll D$ we have $k_1 = \pi w/h$ and $k_2 = w/D$. For the subsequent transformation, we used the approximation[3]

$$F(k_i) = \pi^{-1} \ln(2 \cdot (1 + \sqrt{k_i^*})/(1 - \sqrt{k_i^*}))$$

for $0 \leqslant k \leqslant 0.707$. (5)

Expanding the argument of the logarithm in Eq. (5) in a series in the small parameter k_i, we obtain a simple analytical expression for the capacitance per unit length of the slotted line and then, using Eqs. (1) and (2), an expression for Z_0:

$$Z_0 \sqrt{\mu_0/\varepsilon_0}((\varepsilon_r + 1)/2)^{0.5} \pi((\varepsilon_r - 1)$$
$$\times \ln(16h/(\pi w)) + 2 \ln(4D/w))^{-1}. \tag{6}$$

It is important to note that this approximation is valid for $D < \lambda_0$, where λ_0 is the free-space wavelength. For $D \geqslant \lambda_0$, it is necessary to substitute $D^* = \lambda_0/2$ in Eq. (6). The wave impedances Z_0 calculated using Eq. (6) are in good agreement with the data of Ref. 4 for $w/h > 0.02$ (Fig. 2b).

THE CONTRIBUTION OF THE SUPERCONDUCTING LINE

When deriving the true values of the linear parameters for an HTSC slotted line, it is necessary to add the kinetic inductance $L_1^{(k)}$ per unit length to L_1 and to account for the resistance R_1 owing to the real component of the complex conductivity of the superconducting film.

The sequence for finding the losses in the transmission line is as follows: initially the finite conductivity of the film is neglected and the boundary value problem is solved, in

this case, in a quasistatic approximation. Then it is assumed that including $L_1^{(k)}$ and R_1 does not change the distribution of the fields or the magnitude of the currents. Using the current distribution obtained under the assumption that the film does not have finite conductivity, we then find values of $L_1^{(k)}$ and R_1 which are, therefore, determined by the surface current distribution and by the surface impedance of the superconducting film.

The complex conductivity of the superconductor can be written in the approximation of a two-fluid model as follows:

$$\sigma = \sigma_1 - j\sigma_2. \tag{7}$$

The active part (σ_1) originates in the normal electronic conductivity and is the cause of the losses in the line. The reactive part ($j\sigma_2$) originates in the inertia of the dissipationless motion of the superconducting carriers. This part of the conductivity is related to the London penetration depth λ_L by

$$\sigma = (\omega\mu_0\lambda_L^2)^{-1}, \tag{8}$$

where ω is the angular frequency and μ_0 is the vacuum magnetic permeability.

An expression for $j_s(x)$ was found from a quasistatic analysis of current flow in the transmission line with the use of conformal mapping.[5,6] In order to avoid a divergence in the integral of the square of the surface current density near the edges of the film, the transformation described in Ref. 5 was carried out. The resulting distribution has the form

$$j_s(x) = \begin{cases} \pm\dfrac{4}{\pi}\dfrac{l}{\sqrt{w\lambda_\perp}}, & w/2 \leqslant |x| \leqslant w/2 + \lambda_\perp, \\[3mm] \pm\dfrac{1}{x}\dfrac{2}{\pi}\dfrac{1}{\sqrt{(2x/w)^2 - 1}}, & |x| \geqslant w/2 + \lambda_\perp, \end{cases} \tag{9}$$

where I is the total current of the microwaves propagating in the slotted line, and $\lambda_\perp = 2\lambda_L^2/d$.

The current distribution give by Eq. (9) is plotted in Fig. 1b. The conducting halfplanes of the slotted line have almost equal tangential components of the magnetic field on the two surfaces but in different directions. This meas that the magnetic field parallel to the surfaces is zero in the middle of the film ($y = d/2$, Fig. 1a), which, in turn, is equivalent to a magnetic wall parallel to the x axis. The equivalent surface impedance of the film in this case can be written as a parallel combination of two ''halves,'' each of which has a surface impedance formed by a film of half the thickness; this yields the following result:

$$Z_{s,\text{eff}} = 0.5 \cdot Z_s \coth(j \cdot 0.5d \sqrt{-j\omega\mu_0(\sigma_1 - j\sigma_2) + j\omega\varepsilon_0}), \tag{10}$$

where Z_s is the wave impedance for a wave propagating in the material of the superconducting film and

$$Z_s = \sqrt{\frac{j\omega\mu_0}{\sigma_1 - j\sigma_2}}. \tag{11}$$

The values of $L_1^{(k)}$ and R_1 can now be found from

$$Z_1 = R_1 + j\omega L_1^{(k)}$$

$$= Z_{s,\text{eff}} \cdot \left(\int_{-\infty}^{+\infty} |j_s(x)|^2 ds \right) \left(\int_0^{+\infty} j_s(x) dx \right)^{-2}, \tag{12}$$

where we have used the distributions of the surface current density given by Eq. (9).

A calculation of the integrals yields

$$Z_1 = Z_{s,\text{eff}} \, 32 \cdot (\pi^2 w)^{-1}$$

$$\times (0.5 + \lambda_\perp/w + 0.25 \ln(1 + w/\lambda_\perp)). \tag{13}$$

Assuming that L_1 and C_1 are known (2), the complex propagation parameter can be written as

$$k_z = \beta - j\alpha = -j \sqrt{j\omega C_1 (R_1 + j\omega(L_1 + L_1^{(k)}))}, \tag{14}$$

where α is the damping constant and β is the phase constant including losses.

RESULTS OF THE MODEL

The wave impedance and propagation constant of a slotted line have been calculated on the basis of Eqs. (6) and (14). The effective dielectric constant $\varepsilon_{r,\text{eff}} = \varepsilon_{\text{eff}}(\beta \cdot c/\omega)^2$ and α are plotted in Fig. 3 as functions of the thickness d of a $YBa_2Cu_3O_7$ film at $T = 78$ K on a substrate with $\varepsilon_r = 9.7$.

It is clear from Fig. 3 that $\varepsilon_{r,\text{eff}}$ and α rise rapidly for small d. For small w/h, this variation can be related to an increase in the kinetic inductance $L_1^{(k)}$ and resistance R_1 and, accordingly, in their contribution to these parameters.

CONCLUSION

Analytical expressions have been found for calculating the wave parameters of HTSC slotted lines. The kinetic inductance and Ohmic resistance per unit length have been
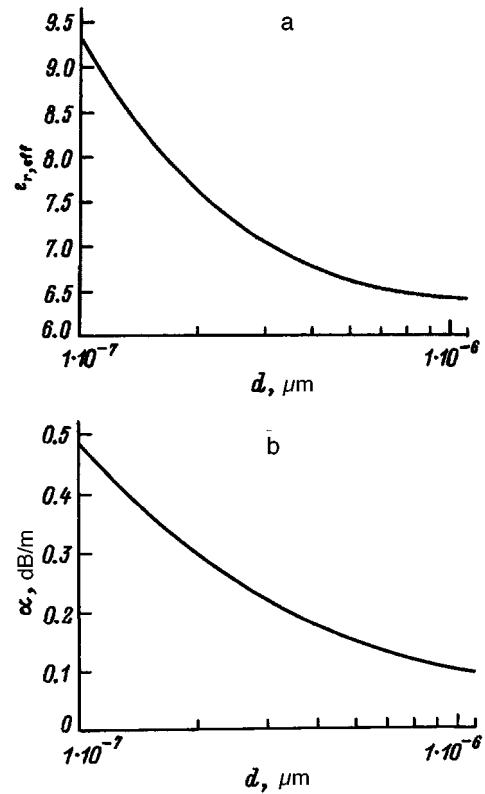

FIG. 3. The wave parameters of a narrow slotted line as a function of the film thickness d for $w = 15$ μm, $h = 0.6$ mm, $f = 3$ GHz, $D = 20$ mm, and $\varepsilon_r = 9.7$. (a) The effective dielectric permittivity $\varepsilon_{r,\text{eff}}$ given by the formula $\varepsilon_{r,\text{eff}} = \varepsilon_r(\beta/\beta_0)$ as a function of film thickness d; (b) the damping coefficient α given by Eq. (14) as a function of film thickness d.

calculated using a quasistatic distribution for the current in the slotted line. The formulas given here are convenient for use in computer-aided design.

One of the authors (O. G. Vendik) thanks Dr. Alex Braginski for discussions in which he called attention to the suitability of narrow slot lines for the construction of microwave squids.

[1] A. H. Miklich *et al.*, IEEE Trans. on Appl. Supercond. **AS-3**, 2434-2438 (1993).

[2] Y. Zhang, H. Soltner, and M. Gottschlich, in *Jahresberichte*, Institut für Schicht- und Ionentechnik. Forschungszentrum Jülich KFA (1995), pp. 89–92.

[3] K. C. Gupta, R. Garg, and R. Chadha, *Computer-Aided Design of Microwave Circuits*, Artech House, Dedham, Mass. [Russian translation, Radio Svyaz', Moscow, 1987].

[4] K. C. Gupta, R. Garg, I. Bahl, and P. Bhartia, *Microstrip Lines and Slot-lines*, 2nd. ed. Artech House, Inc., London (1996).

[5] O. G. Vendik and A. Yu. Popov, Zh. Tekh. Fiz. **63** (7), 1 (1993) [Tech. Phys. **38**, 535 (1993)].

[6] L. M. Löfgren and O. G. Vendik, in *Proceedings of the 23rd EuMC*, Madrid (1993), pp. 644–645.

[7] V. I. Lavrik *et al.*, *Conformal Mapping of Physical and Topological Models* [in Russian], Naukova Dumka, Kiev (1990).

[8] R. K. Hoffmann, *Handbook of Microwave Integrated Circuits*, Artech House, Norwood, Mass. (1987).

Translated by D. H. McNeill

Numerical simulation of space-charge dynamics in a gyrotron trap

D. V. Borzenkov and O. I. Luksha

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

The particle-in-cell (PIC) method is used to simulate the self-consistent accumulation and bunching of space charge in the trap of a gyrotron electron-optical system. It is shown that it is possible to generate charge bunches that oscillate along the direction of the magnetic field. The dependence of the characteristics of these oscillations on the magnitude of the electron current into the trap is determined, along with the effect of the accumulated charge on the velocity distribution of electrons in the current passing through the magnetic mirror. Satisfactory agreement with the experimental data is obtained. © *1997 American Institute of Physics.*
[S1063-7842(97)02109-0]

1. The space charge that accumulates in the trap between the cathode and the magnetic mirror of a gyrotron electron-optical system reduces the quality of the helical electron beam produced.[1-3] Experimental data[1,3-5] indicate that oscillations appear in this trapped charge at frequencies $f = 10$ to 100 MHz, which are associated with longitudinal oscillations of charge bunches. Theoretical analysis of the collective electron processes in the trap can unify the results of experiments and point to effective ways to control these oscillations. In this paper we present the results of such an analysis, derived from numerical simulations using the particle-in-cell (PIC) method (also known as the method of macroparticles); for examples, see Refs. 6 and 7.

2. Our analysis of the electron motion is based on the approximation of adiabatic paraxial drift for a magnetized azimuthally symmetric beam. Using this approach, we need only treat one-dimensional motion of the guiding centers of electron orbits along the z coordinate parallel to the magnetic field. Changes in the longitudinal velocity v of an electron arise from the electric field E_{ex} created by external sources, the self-field E_{sc} of the space charge, and changes in its transverse velocity as it moves through a spatially nonuniform magnetic field.

The equations of motion were integrated using time steps $\Delta t = 0.1$ to 0.5 ns. At a given step N, the charge emitted from the cathode $Q = I \cdot \Delta t$ (where I is the beam current) was divided up into $M = 40$ macroparticles. At the cathode, these particles had zero longitudinal velocity v_k and various transverse velocities $v_{\perp k}$. A Gaussian distribution was used for the electron distribution function with respect to transverse velocity

$$F(v_{\perp k}) = \exp\left[-\frac{4(v_{\perp k} - \bar{v}_{\perp k})^2}{\Delta v_{\perp k}^2} \right],$$ (1)

where $\bar{v}_{\perp k}$ is the transverse velocity averaged over the electron ensemble, and $\Delta v_{\perp k}$ is the velocity spread.

The initial transverse velocity of the ith macroparticle $v_{\perp k}(i)$ for $i = 1 \dots M$ and for given I, Δt, $\bar{v}_{\perp k}$, $\Delta v_{\perp k}$ is determined from the following relations:

$$v_{\perp k}(i) = \frac{v_{\perp k}^{\max}(i) + v_{\perp k}^{\min}(i)}{2},$$ (2)

$$I\Delta t \frac{\int_{v_{\perp k}^{\min}(i)}^{v_{\perp k}^{\max}(i)} F(v_{\perp k}) dv_{\perp k}}{\int_0^\infty F(v_{\perp k}) dv_{\perp k}} = q(i),$$ (3)

$$v_{\perp k}^{\max}(i) = v_{\perp k}^{\min}(i-1).$$ (4)

The charge $q(i)$ is the same for all the macroparticles except for particles with low indices $i < 6$, for which the charge is decreased so as to more uniformly distribute these macroparticles with respect to the transverse velocity. This method of charge partitioning presupposes the presence of several (2 to 5) particles with the largest values of initial velocity $v_{\perp k}$, which undergo reflection from the magnetic mirror even in the absence of a space-charge field.

In calculating the intrinsic space-charge field, we assume that the azimuthally uniform hollow electron beam is thin enough that the change in potential over its transverse cross section can be neglected. Furthermore, in keeping with the typical geometry of the gyrotron electron-optical system, in which characteristic longitudinal sizes are much larger than the distance between the beam and the drift tube, we assume that

$$\Delta_{(z)} U \ll \Delta_{(r)} U,$$ (5)

where $\Delta_{(z)} U$ and $\Delta_{(r)} U$ are respectively the longitudinal and radial components of the Laplacian of the beam potential.

Fulfillment of Eq. (5) was monitored during the calculation. Under these assumptions, the longitudinal component of the space-charge field is given by the expression

$$E_{sc} = -\frac{\partial U(z)}{\partial z},$$ (6)

where

$$U(z) = \rho(z) \frac{S(z)}{2\pi\varepsilon_0} \ln\left(\frac{R_{mp}(z)}{R_0(z)} \right).$$ (7)

Here $S(z)$ and $R_0(z)$ are the area of the transverse cross section and the average beam radius respectively, and $R_{mp}(z)$ is the radius of the drift tube. The space-charge den-
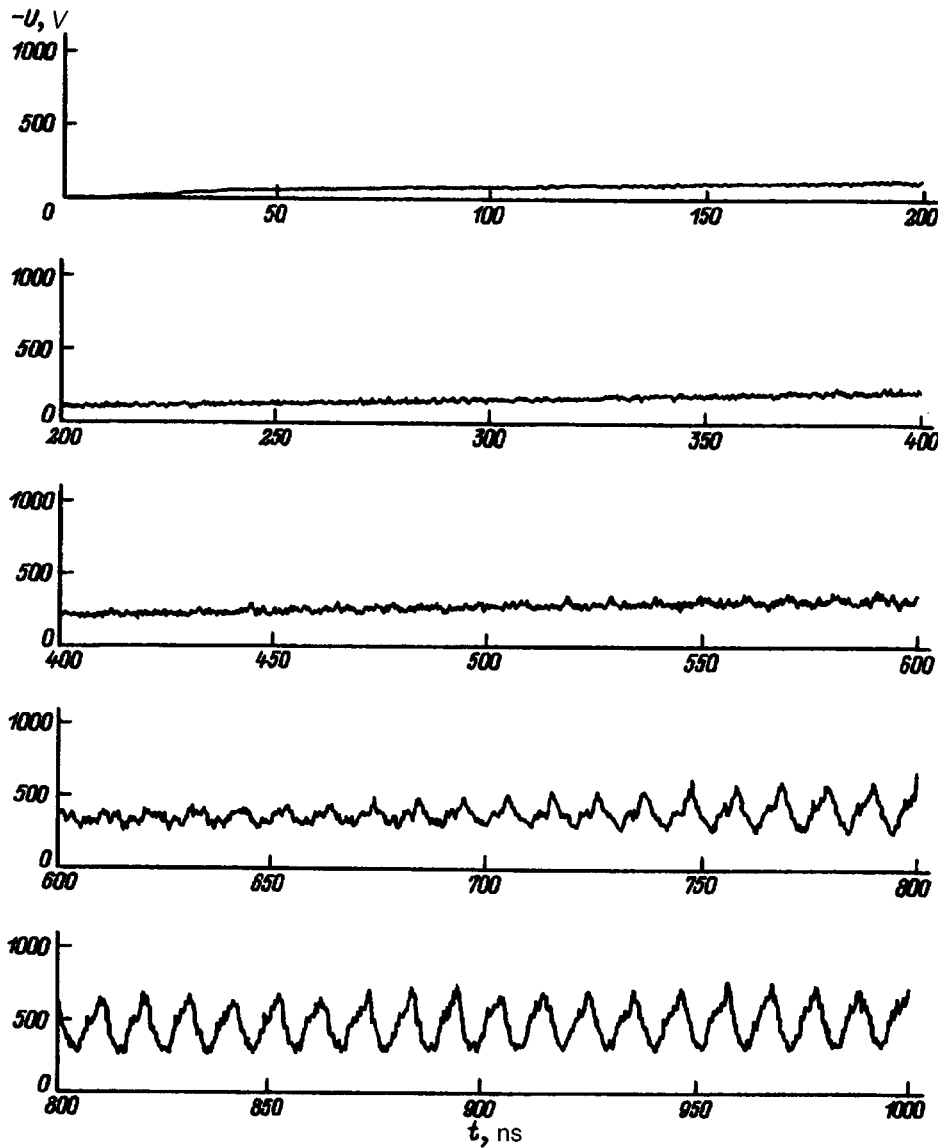
FIG. 1. Dependence of the beam potential $-U$ on the time t (beam current $I=$ 0.2 A, pitch factor $g_0=1.4$, transverse velocity spread $\delta v_{\perp k}$ =0.25.

sity $\rho(z)$ was determined by dividing space along the z coordinate into $L=$ 100 cells and calculating the macroparticle charge in all the cells at each time step, taking into account the finiteness of the particle dimensions.

3. Our calculations were designed to match the gyrotron electron-optical system configuration in which the space-charge oscillations were studied experimentally.[3] The basic structural dimensions and parameters of the operating regime used in these calculation were the following: average cathode radius $R_k=$ 4 mm, width of the cathode emission ring $l_k=$ 2 mm, width of the anode–cathode gap $d_{ak}=5$ mm, slope angle of the cathode surface relative to the device axis $\psi_k=15°$, distance from the cathode to the magnetic field plateau $z_{max}=24$ cm, drift tube radius in the neighborhood of the magnetic field plateau $R_{mp}(z_{max})=3$ mm, external accelerating potential difference $U_{ex}=16$ kV, overmagnetization coefficient $\alpha=B_0/B_k=15$ to 17 (where B_0 and B_k are the magnetic inductions at the plateau and at the cathode), and magnetic induction $B_0=$ 2.5 T. The longitudinal distributions of the external electric field $E_{ex}(z)$ and magnetic induc-

tion $B(z)$ were specified approximately using a piecewise-linear approximation.

We analyzed several cases for values of the current I, relative transverse velocity spread $\delta v_{\perp k}=\bar{v}_{\perp k}\cdot\Delta v_{\perp k}$, and pitch factor of the ''cold'' beam g_0 within the ranges $0.1 \leqslant I\leqslant 0.5$ A, $0.25\leqslant \delta v_{\perp k}\leqslant 0.3$, and $1.2\leqslant g_0\leqslant 1.6$. The quantity g_0 can be regulated by varying the magnetic reversal coefficient α in accordance with the relation

$$g_0=\frac{v_{\perp 0}}{v_0},\qquad(8)$$

where $v_{\perp 0}=\bar{v}_{\perp k}\cdot\sqrt{\alpha}$ is the average transverse electron velocity, and $v_0=\sqrt{2\,\eta U_{ex}-v_{\perp 0}^2}$ is the average longitudinal velocity at the magnetic field plateau.

4. When the parameters I, $\delta v_{\perp k}$, g_0 lie within these ranges, accumulation of space charge in the trap is accompanied by excitation of low-frequency oscillations ($f\sim 100$ MHz). Figure 1 shows a typical time dependence of the beam potential $-U$ in a cell at a distance of 16.8 cm from

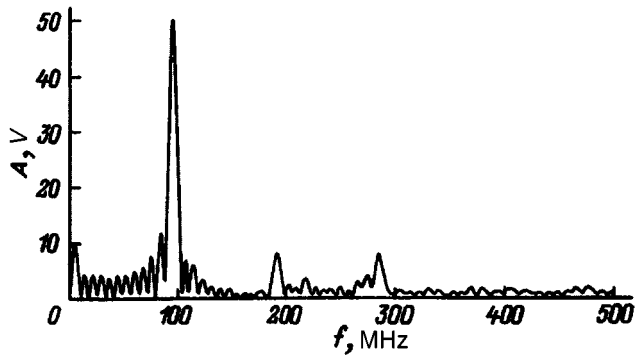FIG. 2. Spectrum of the time series $-U(t)$ shown in Fig. 1 over the time interval $835 \leqslant t \leqslant 960$ ns ($I = 0.2$ A, $g_0 = 1.4$, $\delta v_{\perp k} = 0.25$).

FIG. 3. Spectrum of the time series $-U(t)$ at a low particle flux into the trap over the interval $935 \leqslant t \leqslant 1060$ ns ($I = 0.2$ A, $g_0 = 1.2$, $\delta v_{\perp k} = 0.25$).

the cathode calculated for $I = 0.2$ A, $\delta v_{\perp k} = 0.25$, $g_0 = 1.4$, and $\Delta t = 0.2$ ns. For the initial time segment ($t < 600$ ns) we observe an approximately linear increase in the potential $-U$ and an increase in the number of macroparticle caught in the trap. At time $t \cong 600$ ns the charge accumulated in the trap equals roughly half of the total charge of particles located in the interaction region between the cathode and the magnetic field plateau. The total number of particles in this region reaches ~ 9500. At later times the amplitude of the beam potential modulation increases, with a modulation period of ~ 10 ns. Figure 2 shows the Fourier spectrum of a fragment of a time series of $-U(t)$ with duration 125 ns (835 ns $\leqslant t \leqslant 960$ ns). Against a background of wideband noise, discrete peaks can be seen at frequency multiples, with a fundamental frequency component at $f \cong 94$ MHz. After the period in which the oscillations grow, the system reaches a quasisteady state in which neither the amount of charge in the trap nor the oscillation amplitude change very much ($t > 900$ ns in Fig. 1).

The oscillation frequencies obtained in these calculations agree approximately with those of a single electron moving between reflection planes. We also observed a phase shift between values of the function $-U(t)$ recorded at points that were separated along the axis. These data confirm the prior experimental conclusion that the oscillations have a spatial structure, indicating that they are associated with longitudinal oscillations of charge bunches in the trap.

The charge bunching is clearly based on an instability of negative-mass type in the ensemble of nonisochronous electron oscillators. The computed value of the oscillation frequency is approximately 1.5 to 2 times larger than the experimental values, which may be due to the approximate nature of our computational model and the external field approximations we used.

An increased particle flux into the trap, which can be implemented by increasing I, $\delta v_{\perp k}$ or g_0, is accompanied by a decrease in the startup time of the oscillations and the time for the system to reach its quasisteady state. For example, increasing the pitch factor g_0 from 1.2 to 1.6 causes the oscillation startup time to decrease from ~ 700 ns to ~ 450 ns. For values of the pitch factor $g_0 \leqslant 1.1$ no oscillations are observed at all within an analysis time of $t \sim 1300$ ns, which at this stage is the limit imposed by the capabilities of the
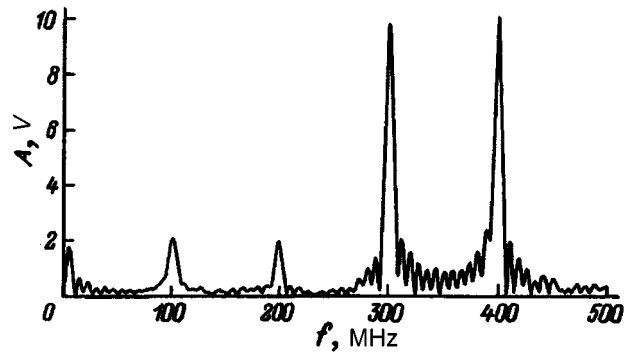
computer. The threshold pitch factor for oscillations to appear in the trap is experimentally found to be ~ 1.0 in the operating regimes corresponding to these computations.[3]

The spectral content of the oscillations also changes as the particle flux into the trap is varied. For the small pitch factor $g_0 = 1.2$, several bunches can exist simultaneously during the initial stage of charge bunching in the trap. In the oscillation spectrum this is reflected by increased harmonic amplitudes relative to the fundamental component (Fig. 3). As the "amplitude" of the bunches increases, their interaction increases as well, and eventually one bunch "absorbs" the others.

For larger particle fluxes into the trap ($g_0 = 1.6$), the oscillations are characterized by increased amplitude of the wideband noise and a decrease in the power at the fundamental peak of the spectrum (Fig. 4). This randomization of the collective processes is probably associated with the increased rate of renewal of the trapped space charge, which decreases the fraction of "long-lived" electrons in the trap. In this regime, the oscillator dynamics are also characterized by time-periodic amplification and suppression of the fundamental spectral component.

Bulk charge that accumulates in the trap acts on the velocity distribution of electrons passing through the magnetic mirror in the region of uniform magnetic field. In Fig. 5 we plot the total charge of particles passing through the plane $z = z_{max}$ at time $\Delta \tau = 100$ ns versus their longitudinal velocity. These data were obtained for a regime with $I = 0.2$ A,
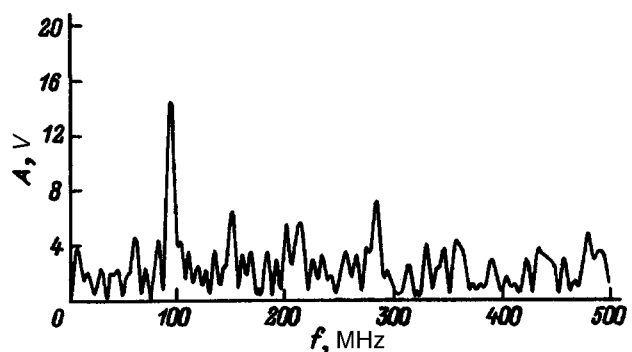

FIG. 4. Spectrum of the time series $-U(t)$ at a high particle flux into the trap over the interval $1175 \leqslant t \leqslant 1300$ ns ($I = 0.2$ A, $g_0 = 1.6$, $\delta v_{\perp k} = 0.275$).
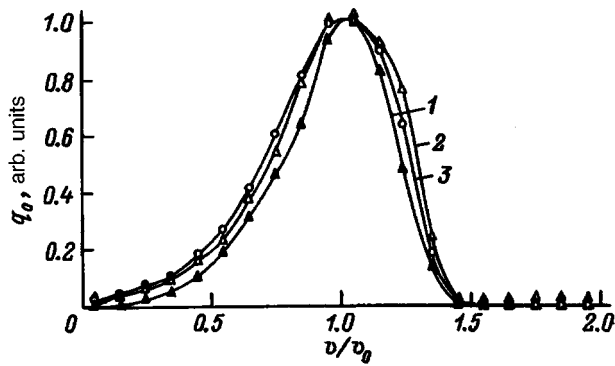
FIG. 5. Charge distribution q_0 of particles entering the magnetic field plateau at time $\Delta\tau = 100$ ns versus their longitudinal velocity v. The value of v is normalized by the average longitudinal velocity for a ''cold'' beam v_0.

$\delta v_{\perp k} = 0.25$, and $g_0 = 1.4$. Curve *1* corresponds to times $\Delta\tau$ in the range 0–100 ns, when the charge in the trap is small; curve *2* is for 500–600 ns, when the static charge reaches its highest level without significant oscillations; curve *3* is for 800–900 ns, when the oscillations have maximum amplitude. As the space charge in the trap increases, the velocity distribution function of the particles becomes washed out (compare curves *1* and *2*). The increase in the oscillations is accompanied by a change in the form of the velocity distribution, as the fraction of macroparticles with velocity above average decreases and the number of slow particles increases. The value of the velocity spread in this case is practically unchanged.

In conclusion, it is noteworthy that the assumptions we have used in this model significantly simplify the real processes involved in electron current oscillations in a gyrotron electron-optical system. Nevertheless, our data are in satisfactory agreement with the experimental results, in particular the threshold conditions for excitation of space-charge oscillations in the trap, their spectra, and their spatial characteristics. As we pursue our approach to the simulation of collective processes in the gyrotron electron-optical system, we will address the possibility of controlling the oscillation characteristics via the nonuniform fields in the drift-tube region.

[1] Sh. E. Tsimring, *Lectures on Microwave Electronics and Radio Physics*, Vol. 4 [in Russian], Saratov State University Publishers, Saratov (1974), p. 3.
[2] V. E. Zapevalov and Sh. E. Tsimring, *Lectures on Microwave Electronics and Radio Physics* [in Russian], GosUNTs ''Kolledzh,'' Saratov (1996), Vol. 2, p. 129.
[3] O. I. Luksha and O. Yu. Tsybin, *Lectures on Microwave Electronics and Radio Physics*, [in Russian], Izd. SGU, Saratov (1993), p. 20.
[4] O. I. Luksha and G. G. Sominskiĭ, Zh. Tekh. Fiz. **64**(11), 160 (1994) [Tech. Phys. **39**, 1173 (1994)].
[5] O. I. Luksha and G. G. Sominskiĭ, Zh. Tekh. Fiz. **65**(2), 198 (1995) [Tech. Phys. **40**, 221 (1995)].
[6] C. K. Birdsall and A. B. Langdon, *Plasma Physics via Computer Simulation* (McGraw-Hill, New York, 1985) [Russian trans., Énergoatomizdat, Moscow, 1989].
[7] A. S. Roshal', *Simulation of Charged Beams* [in Russian], Atomizdat, Moscow (1979).

Translated by F. Crowne

Structure and composition of tungsten silicide thermal–field microprotrusions

M. V. Loginov and V. N. Shrednik

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

The thermal–field microprotrusions that grow on the surface of a tungsten tip coated with silicon when the tip is heated in an electric field are investigated by a suite of field emission methods, including electron field emission, ion desorption microscopy, and the atomic-probe method. For Si coatings more than a few monolayers thick, microprotrusions are observed to grow in the field desorption regime when the tip is heated to temperatures $T = 1100–1200$ K in an electric field with initial intensity $F = 5.7–8.6 \times 10^7$ V/cm. The field at which they evaporate is $1.2–1.8 \times 10^8$ V/cm. The set of moving spots (i.e., microprotrusions) forms rings whose collapse signals the dissolution of the thermal–field growths on the developed faces. The most interesting structures are the sharp microprotrusions that grow on the central facet of a {110} tungsten tip under certain conditions. Atomic-probe analysis of their composition reveals that they consist of tungsten trisilicide WSi_3 with a monolayer surface skin whose composition is close to WSi_2. The intense growth of these formations on an initially smooth close-packed {110} face of tungsten is evidence that reconstruction of the latter takes place under the influence of the strong field and the interaction with silicon. © *1997 American Institute of Physics.* [S1063-7842(97)02209-5]

INTRODUCTION

Thermal–field microprotrusions form at the surface of a conducting crystal heated in the presence of an electric field F when this field is strong enough that the field-induced ponderomotive forces, which stretch the crystal, overcompensate the compressive force of surface tension (which smooths out the surface). For this to happen, the surface temperature should be high enough that the active surface diffusion of atoms responsible for the crystal growth kinetics can occur. Data of this kind on microprotrusions grown on metal tips was analyzed and summarized in Refs. 1 and 2. Usually microprotrusions grow readily in the neighborhood of a crystal vertex (more rarely an edge); this growth is limited by so-called field -induced reconstruction, which causes the growth of microprotrusions to cease. For pure metals it is not possible to grow thermal–field microprotrusions on the most closely packed faces, probably because it is difficult to nucleate elements of a new layer on these faces, which are as a rule ideally smooth.

However, when certain impurities are present on the surface of a smooth face, thermal–field microprotrusions can grow even at the center of the facet. Thus, in Refs. 3 and 4 the authors describe cases where typical thermal–field microprotrusions grow in a reproducible fashion on {110} faces of tungsten coated with silicon (in quantities of more than a monolayer). However, their experiments yielded only indirect data (e.g., the magnitude of the desorbing field, the smoothing temperature, etc.) about the chemical composition of the microprotrusions. It is important to identify more precisely where these microprotrusions are on this particular close-packed {110} face, since this information is relevant to the reconstruction pattern generated by the interaction with the adsorbate and the strong electric field.[4] An isolated microprotrusion on the central facet of a tip point is an interesting object in itself from the point of view of creating field-emission sources of electrons and ions that are highly ''pointlike'' in their properties. These facts make investigations of the reproducible growth of microprotrusions, the stability of their emission properties, and peculiarities of their field-induced evaporation (which in particular determines ion emission) even more valuable.

In this paper we will investigate the chemical composition of thermal–field microprotrusions grown under preset conditions on the central facet of a (110)W tip coated with silicon, refine the conditions for the growth of these microprotrusions, and identify distinctive features and characteristics of their field-induced evaporation (both low-temperature and high-temperature). Summarizing our work in advance, we have discovered that such microprotrusions consist of tungsten silicides of various compositions. Whereas the authors of Refs. 3 and 4 investigated low-temperature field-induced evaporation of microprotrusions, our focus of attention is the high-temperature field evaporation of fluctuating and self-maintaining microprotrusions. Moreover, ours is the first study of these phenomena, which up to now have been described only for certain pure metals,[5–7] in a chemical compound. We are also the first to undertake a precise atomic-probe analysis of the composition of a microprotrusion, which constitutes a special methodological problem.

EXPERIMENTAL TECHNIQUE

In this work we used the time-of-flight atomic probe described in Ref. 8. As we pointed out in that paper, it is possible to observe emission images of the tip by using an image brightness amplifier consisting of two microchannel plates and a luminescent screen. Our source of silicon vapor was a rod heated by a current passing through it, mounted in the atomic probe chamber to the side of the tip. The residual

gas pressure at the time of the experiments was $(3–4)\times10^{-9}$ Torr. The tungsten tips were oriented along the $\langle 110 \rangle$ axis.

The experiments, which were designed to obtain mass spectra of the field-evaporated material in the atomic probe under conditions of thermal–field microprotrusion growth, were made difficult by contradictory requirements. On the one hand, the tips had to be strongly heated for a long time prior to the experiments in order to purify them of impurities. This procedure blunts the tips very rapidly if there is any appreciable concentration of Si in the tungsten bulk. On the other hand, the samples must be maximally sharp in order to avoid unacceptably high evaporating voltages, both dc, i.e., base-level, and pulsed. The situation is only partially alleviated by the fact that the microprotrusions that are grown enhance the local field intensity by a factor of 1.5 to 3.0 compared to the original surface. Field evaporation, especially at low temperatures, requires extremely high fields both for pure tungsten and for silicides or silicon.

In order to keep the tips sufficiently fine, we avoided strong heating (we did not go above a temperature of 1800 K) and heated the samples in the presence of a moderately strong electric field (in feeding a voltage of positive sign to the tip). The silicon deposited on the lateral surface of the tip was difficult to remove completely under these circumstances, and it deposited in the neighborhood of the end of the tip. This hindered our *a priori* calibration of the amount of Si. However, after this it became easy to evaluate this quantity using the mass spectra.

Despite our good vacuum, tips heated in the chamber often contained a certain quantity of carbon at its surface. This we inferred from the characteristic image of a ribbed crystal of W, which is typical of solid solutions of C in W at subcarbide concentrations.[9,10] Attempts to remove the carbon by annealing in oxygen at oxygen pressures of order 1×10^{-6} Torr and temperatures of 1700–1800 K led to temporary cleaning; however, the tips were blunted, and after subsequent heating once more showed signs of carbon contamination. However, it is known from the literature[11] that after Si deposition even a monolayer of C on a W surface is displaced by the Si into the substrate bulk, where it is unable to appreciably influence surface processes. Therefore, we used a surface with traces of carbon as our starting surface. The field evaporation spectra from such surfaces sometimes exhibit signs of carbon contamination, but our preliminary atomic-probe experiments on the surface of a ribbed crystal of W revealed principally tungsten ions in the field evaporation spectra.

The temperature of the tip T was determined using a Piro optical pyrometer. The tip could be heated at the same time that high voltages V of either polarity were fed to it by passing the current through a small bow to which it was welded. In this work, ''low temperatures'' should be taken to mean room temperature (300 K). We used room temperature rather than cryogenic temperatures in order to simplify the experiments, since at 300 K migration of W and Si atoms is quite reliably frozen out.

HIGH-TEMPERATURE FIELD EVAPORATION FROM TUNGSTEN COATED WITH DEPOSITED SILICON

For a tungsten tip coated with low concentrations of silicon at room temperature, corresponding to a fraction of a monolayer, the conditions (on T,V) required for microprotrusion growth are close to those for pure tungsten,[5] and a significant decrease in both T and V observed only when larger amounts of silicon are deposited (several monolayers). Under these conditions, after exposing the tip to temperatures of 1100–1200 K and voltages 2.5 to 4 times larger in absolute value than the V needed to observe the initial electron field-emission image (i.e., for $F=(5.7–8.6)\times10^7$ V/cm at the original surface) and with opposite sign, we were able to grow microprotrusions and study them (*in situ*) by using ion desorption images. The field at the peaks required for effective evaporation (rather than growth) was estimated to be $(1.2–1.8)\times10^8$ V/cm, which is considerably smaller than the values required for tungsten (at least the lower limit is).

Images of these closely spaced and fluctuating microprotrusions were quite dim, even for extremely large (of order 10^6) enhancement coefficients of the microchannel plates. They appeared over the entire visible field. Sets of spots — microprotrusions — arranged themselves in rings around the planar facets, which then collapse inward (at $T=1100–1200$ K). This behavior is similar to that of Ir reported in Ref. 6 and Pt in Ref. 7. In this regime, and for these surface concentrations of Si, we did not observe any special localization of microprotrusions on the {110} face. For pure tungsten, inward collapse of rings under conditions of hot field evaporation has not been reported. It could be that higher temperatures are required for tungsten, due to the so-called ''micro-outgrowths'' associated with its growth.[1,6] We were unable to determine the composition of the microprotrusions present on a tungsten surface coated with a film of Si under these conditions, because the contrast between the evaporation rate for the base-level voltage and the rate at the peak of the evaporating pulse turned out to be too small; moreover, the recorded spectra consists of chaotic pulses, indicating a noise process.

The use of *in situ* recording of desorbed ion images allowed us to control microprotrusion growth in the neighborhood of the {110} face on the tungsten tip. However, we were able to do this only after depositing a large amount of Si (probably more than 10 monolayers) on a tungsten surface that was practically free of previous Si deposits. When the aforementioned microprotrusions were grown on the tip, we had to apply voltages 2.8–3.3 times larger than the voltage required to observe electron emission images (and naturally of the opposite sign) while increasing the temperature to 1000 K or higher before we observed the expected ion desorption pattern. In the temperature range $T=1100–1200$ K, against a background of dim spots that flickered at various positions (ion currents) an extremely bright spot would appear unexpectedly and abruptly at the center of the screen on a {110} face. This spot, which had a tendency to increase in brightness and observed size, was unstable and could disappear as rapidly as it appeared. Increasing the voltage could generate new bright spots along with the previous one, while decreasing it often led to an abrupt disappearance of the spot.
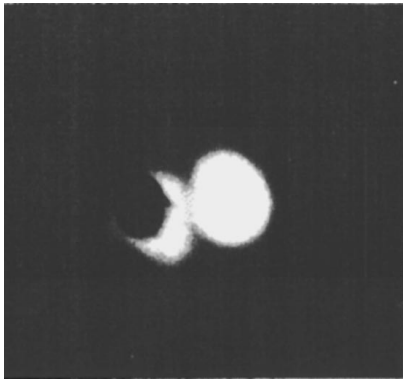
FIG. 1. Electron field emission image of two tungsten silicide thermal–field microprotrusions. $V_e = 2$ kV. One of the microprotrusions is pointing at the atomic-probe iris.

FIG. 2. Field evaporation spectra from thermal–field microprotrusions grown on 110 W. a — spectrum obtained immediately after growth of the microprotrusion, b — spectrum obtained after 15 hours exposure of the microprotrusion surface to residual gases at a pressure of 10^{-5} Torr.

We were able to preserve these strongly emitting microprotrusions (sometimes there were two or three) only by rapidly lowering the temperature while maintaining the applied voltage, which was removed only after the tip had cooled considerably (almost to room temperature). In these cases, by feeding a negative voltage to the tip we could image one or two, more rarely three, microprotrusions in the region of the tip axis (i.e., on the {110} face). They appeared as circular bright spots of electron field emission at voltages in the range 1.4 to 2.0 kV (most often at 1.8 kV). Often they would occur within the probe aperture without additional adjustment, which in these experiments coincided with the tip axis.

By evaporating these microprotrusions, which were grown repeatedly and reproducibly, we obtained mass spectra, which we discuss in the next section. Figure 1 shows an example of an electron field-emission image of such microprotrusions. Note that sharp microprotrusions appear spontaneously only when the {110} face is at the peak, and do not occur on lateral facets of this kind where the field is considerably lower. With regard to their external shape and manner of appearance in the electron regime, and in the relative emission parameters of the latter, these protrusions are analogous to those described in Refs. 3 and 4. In order to grow them on a {110} face it is necessary to generate a new layer on the planar portion of the facet. Consequently, the facet cannot be kept smooth, and must reconstruct. The sharp threshold character of microprotrusion growth with increasing field attests to its influence on this reconstruction.

FIELD-EVAPORATION MASS SPECTRA AND MICROPROTRUSION COMPOSITION

Figures 2a and 2b show field-evaporation mass spectra of microprotrusions obtained at low (room) temperature and the same base-level (18 kV) and pulsed voltages (4.8 kV) (a voltage of 1.8 to 2.0 kV is required in order to observe an electron field-emission image). The spectrum shown in Fig. 2a was obtained immediately after the growth of the microprotrusion, while that shown in Fig. 2b was obtained the next day some 15 hours after the growth (during this time the tip was kept in an atmosphere of residual gases at a pressure of order 10^{-5} Torr at room temperature, after which the spectrum was recorded in a vacuum of 5×10^{-9} Torr). Clearly
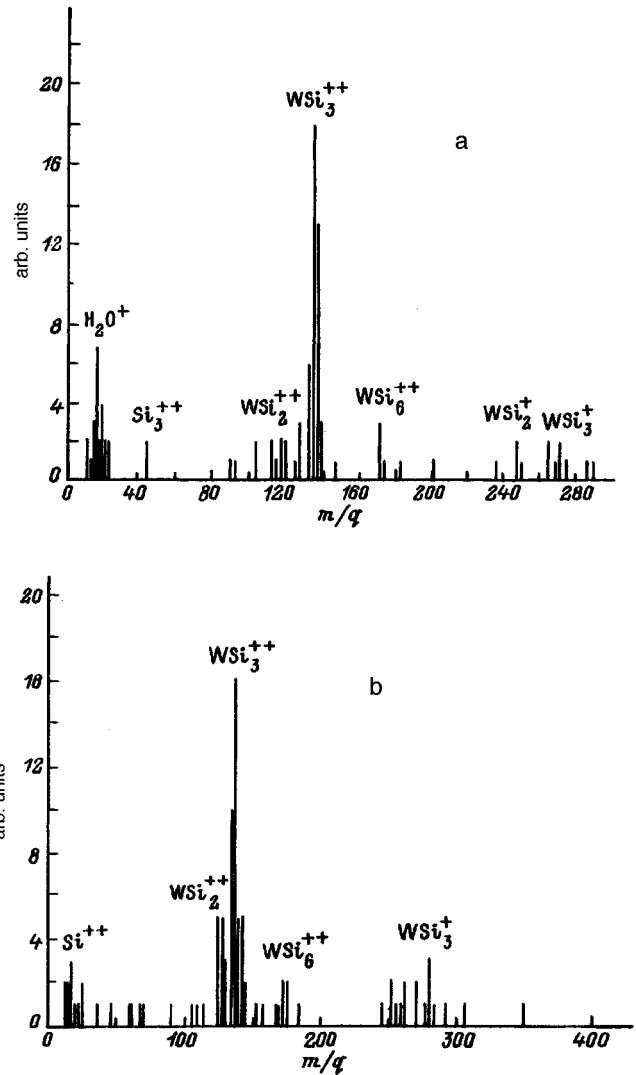
the composition spectra are practically the same. It is astonishing that the field-emission characteristics of the microprotrusion remain unchanged after 15 hours exposure to a poor vacuum, indicating an extraordinary (and atypical, e.g., for tungsten) stability against corrosion and passivity against adsorption.

The dominant peak in the spectra of Fig. 2 is at mass number 134, which corresponds to a silicide with the composition WSi_3 (in the form of WSi_3^{++} ions). This peak also dominates in other spectra of other microprotrusions grown on {110} tungsten faces and recorded for other evaporating voltage parameters. Furthermore, the spectra of Fig. 2 (and in other analogous spectra) always exhibit a peak with low but still significant amplitude at mass number 120, which corresponds to the ion WSi_2^{++}. As a rule, we find a group of low-amplitude peaks (i.e., near the noise level) in the neighborhood of mass number 268, corresponding to the ion WSi_3^+. Peaks in the low-mass region (30 to 45), which are sharply expressed in Fig. 2a, may contain traces of carbon
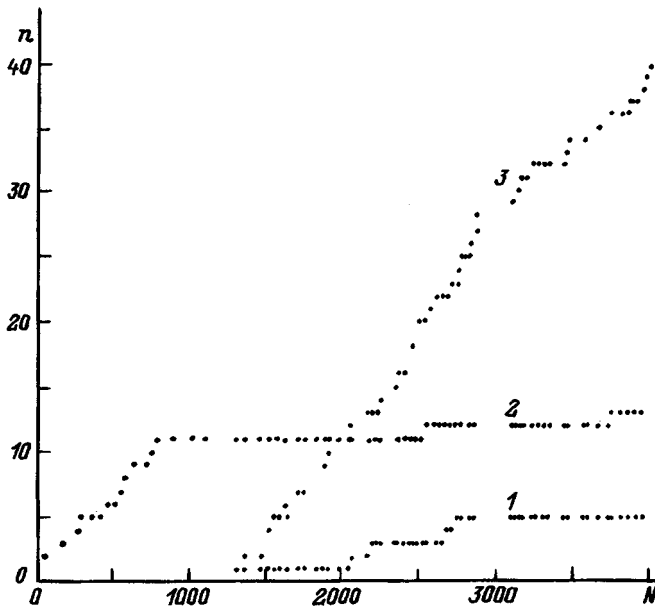
FIG. 3. Curves for the cumulative number of recorded ions n versus the number of prior voltage pulses N for three peaks of the spectrum of Fig. 2a. $1 — WSi_3^+$, $2 — WSi_2^{++}$, $3 — WSi_3^{++}$.

impurities C^+, C_2^+, C_3^+, and SiC^+ along with the ions Si^+, Si_3^{++}, and H_2O^+. At the noise level pulses were encountered in the spectra corresponding to still heavier ions, e.g., $(WSi_3)_2^+$ and other cluster ions up to $(WSi_2)_5^+$ and $(WSi_3)_5^+$, as well as $(WSi_n)_m^+$, where m varied from 1 to 5 and $m > n$ by a factor of 2 to 3. Thus, the field-evaporation spectra of thermal–field microprotrusions consisted mostly of ions of WSi_3—56 ions out of the 117 recorded in the spectrum of Fig. 2a—followed by ions of WSi_2—12 ions (in the spectrum of Fig. 2a the ions were primarily doubly charged, and to a lesser extent singly and triply charged).

Analysis of the accumulation of ions recorded by the spectra of Fig. 2a (Fig. 3) shows that tungsten disilicide and trisilicide ions evaporate nonuniformly with time. Initially we record evaporation of WSi_2 ions, and then WSi_3 ions. This dynamic behavior clearly indicates spatial localization of the corresponding compounds in the microprotrusions: layers of tungsten disilicide located at the peak of the microprotrusion, and beneath them layers of tungsten trisilicide. A plot of the total accumulation of all recorded ions (Fig. 4) reveals a wavelike structure, with appreciable pauses between groups of arriving ions, often with an increase in the rate of accumulation before the pause. This structure of the accumulation curve can be explained by postulating a layer-by-layer evaporation of the microprotrusion. The monolayers evaporating at the periphery draw inward in rings toward the axis of the microprotrusion. As such a ring ''collapses'' inward, the rate of evaporation might increase, after which some slowing of the process would begin. Comparison of the plots of $n(N)$ (Figs. 3 and 4a) show that the WSi_2^{++} ions were concentrated principally in the first wave, i.e., in the first monolayer to be analyzed. However, the detection of WSi_3^{++} and WSi_3^+ ions indicates evaporation of subsequent layers.

At sufficiently high base-level voltages, evaporation of a microprotrusion can take place in the absence of a voltage pulse as well. In this case, the evaporating pulses select ions from the same outermost steps of monolayers moving within the zone of the probe aperture. From this it follows that field evaporation spectra obtained at high base-level voltages need not indicate the evaporation of the very first portion of ions from the microprotrusion surface, but may correspond to evaporation of deeper lying monolayers.

In order to identify the conditions necessary for recording ions evaporated from the very top of the tip, we monitored the microprotrusions experimentally to determine their stability in the evaporation field. We did this by subjecting a freshly grown microprotrusion to the action of a constant evaporating field for a certain fixed time, and then monitoring the voltage V_e at which electron field-emission images were observed. Table I lists the results of these experiments. From the table it follows that when a high base-level voltage (greater than 3–3.5 V) is applied to the tip, the very top may evaporate almost immediately; this top probably consists of only a few atoms.

In order to estimate the actual magnitude of the electric field intensity F, we will assume that the brightness of the spot observed in the electron field-emission regime for sharp microprotrusions (with radii of curvature $r = 30–50$ Å) corresponds to electron field-emission current densities of around 1 Å/cm. If we use the tables of Dolan[12] to choose a field intensity F corresponding to this current density and a work function $\varphi = 4.65$ eV close to the expected work functions of silicides,[13] we obtain $F = 2.7 \times 10^7$ V/cm, and for $V = 2.2$ kV we calculate a field multiplier $\alpha = F/V$ for the microprotrusion ($\alpha = 1.23 \times 10^4 cm^{-1}$). The first sign of blunting of the microprotrusion is observed for F in the range 7.4×10^7 V/cm $> F > 6 \times 10^7$ V/cm (V between 5 and 6 kV), while fields $F = 0.8–1.0 \times 10^8$ V/cm systematically and continuously blunt the microprotrusion.

We used a rather careful method to analyze the microprotrusion in order to select ions from its very top. We fed a pulsed voltage with amplitude 4 kV to tips with freshly grown microprotrusions and an initial $V_e = 2–2.2$ kV (for zero base-level voltage), which was clearly insufficient to evaporate even a single atom from the top. We then increased the base-level voltage slowly until the first recorded ions appeared at $V_b = 4$ kV (i.e., at 8 kV total voltage or in a field $F = 1.04 \times 10^8$ V/cm). At this point, the first ninety pulses generated 10 recorded ions, without a single ion more even after a rather long wait. The recorded ions arrived in the

TABLE I. Blunting of a tungsten silicide microprotrusion due to field evaporation of its top when a positive voltage V_+ is fed to it. The blunting is revealed by the increased value of the voltage V_e.

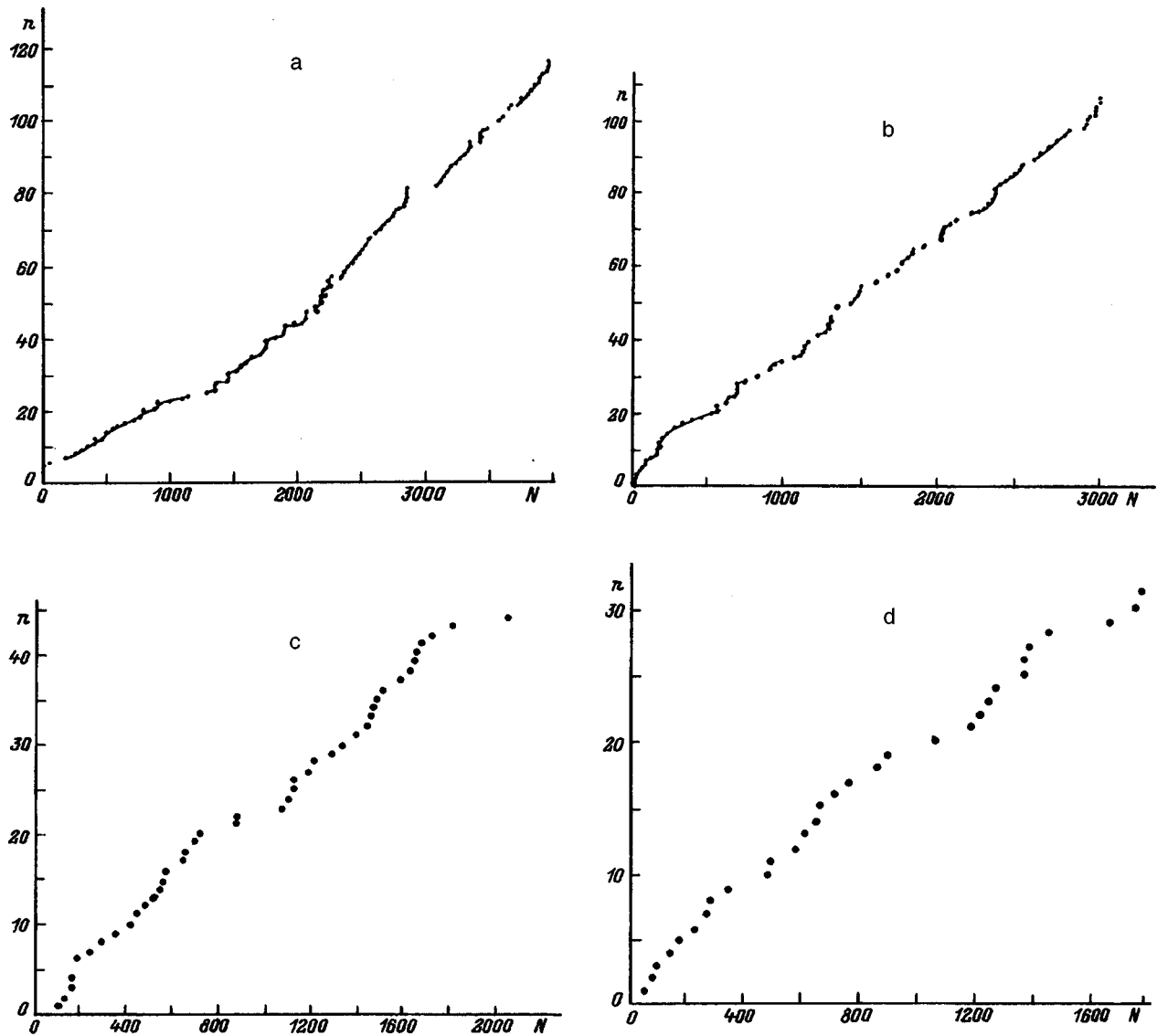| V_+, kV	V_e, kV
4	2.2
5	2.2
6	2.3
7	2.6
8	2.9
10	3.4
11	3.6

FIG. 4. Curves for the cumulative number of recorded ions n versus the number of prior voltage pulses N for four field evaporation spectra of tungsten silicides at room temperature. a — from spectrum 2a; b — from spectrum 2b. $V_b + V_i$, kV: a–c — 18+4.8, d — 14+4.8; n_0/N_0: a — 106/4000, b — 105/3060, c — 44/2108, d — 33/1788 (the label 0 denotes total numbers of ions and pulses).

order shown in Table II. We have reason to believe that in this experiment we recorded all the ions that evaporated from the tip of the microprotrusion. The total recorded ion content corresponding to the disilicide WSi_2 (12 atoms of W and 24 atoms of Si) included a large variety of ions. This implies atomic disorder of the top of the microprotrusion, which

could be a consequence of the continuous field evaporation at high temperature in the process of ''quenching.'' Further increases in V_b did not affect the rate of recording of ions appreciably (even when V_p was increased to 4.8 kV) up to a value of $V_b = 14$ kV (more precisely, in two cases when V_b was increased smoothly by 1 kV followed by a pause, four

TABLE II. Ions field-evaporated from the top of a silicide microprotrusion when $V_b + V_i = 4 + 4$ kV.

Pulse number	Mass numbers m/q of recorded ions	Presumed ions	Computed m/q
12	145	WSi_4^{++}	148
21	307	WSi_4^{+}	296
23	217	WSi^{+}	212
40	424 and 478	$(WSi)_2^{+}$, $(WSi_2)_2^{+}$	424, 480
50	170 and 199	$(WSi_6)^{++}$, $(W_2Si)^{++}$	176, 198
66	65 and 76	W^{+++}, $(WSi_2)^{+++}$	61.3, 80
84	6	C^{++}	6

more ions were recorded: 2 WSi^{++} and 1 WSi^{++} ion at $7+4.8$ kV, and one WSi_3^{++} ion at $10+4.8$ kV).

For $V_b + V_p = 14+4.8$ kV we generated a spectrum out of of 33 ions (the ratio of the total number of evaporating pulses to the number of resulting ions $= 54$), consisting of 25 ions corresponding to the composition WSi_3 and only five WSi_2 ions. However, the spectrum for this experimental run was incomplete and selective since the three waves in the function $n(N)$ (Fig. 4d) record the evaporation of a monolayer at the applied base-level voltage for every 10 ions recorded. Upper bounds on F for V_b and $V_b + V_p$ correspond to 1.82×10^8 V/cm and 2.44×10^8 V/cm (assuming that α is unchanged after the evaporation of 14 ions from the top). Of course the real fields were smaller; nevertheless, they provided an evaporation rate of 0.2 ions/s for the base-level voltage (estimate) and 2×10^6 ions/s for the pulse. These results confirm that only the top layer of the microprotrusion has the composition WSi_2, and that subsequent layers consist of tungsten trisilicide.

DISCUSSION OF RESULTS

In summary, our analysis shows that when high concentrations of silicon are deposited on the {110} face of a tungsten tip heated in an electric field up to temperatures T too low for diffusion and thermal–field reconstruction of pure W(1000–1200 K), sharp silicide microprotrusions grow with compositions close to WSi_2 in the topmost layer and to WSi_3 in the bulk of the microprotrusion. At relatively low temperatures the {110} face of tungsten is capable of nucleating and growing microprotrusions in whose construction from 1/3 to 1/4 of the tungsten atoms participate. Under these circumstances the original tungsten face cannot remain unperturbed. In addition to reconstruction, which favors crystal growth, the face also provides the conditions necessary for the formation of tungsten silicides. The processes of crystal growth and chemical interaction is strictly controlled by the electric field intensity.

Under conditions of high-temperature field evaporation, the silicide skin of silicon-impregnated tungsten exhibits the phenomenon of inwardly collapsing rings, which is atypical of pure tungsten at these (relatively low) temperatures. Using this effect we can make pointlike and controllable sources of silicide ions, which should make it possible to controllably deposit tungsten silicide on a substrate on a scale of tens or hundreds of angstroms.

In this paper we describe the first investigations of the composition of thermal–field microprotrusions for a chemically multicomponent conducting material. It is characteristic that at low temperatures ratios of evaporation rates for the

steady base-level voltage and the sum of pulsed and base-level voltages are automatically established with a contrast of greater than 10^6. This allows us to obtain correct, easily calibrated spectra that reflect the composition of the microprotrusion. If active field evaporation takes place when the voltage V_b is applied, then the microprotrusions are blunted until they reach a state in which the evaporation is unacceptably slow. Elucidation of the composition of a microprotrusion starting at its very tip requires the use of special methods, several of which were described in the previous section.

The thermal–field growth of silicide microprotrusions deserves a special analysis and discussion of its own. Our preliminary observations show that the treatment of equilibrium and steady-state microprotrusions developed in Refs. 2 and 14 for metals is entirely applicable to multicomponent conducting materials.

[1] V. N. Shrednik, *Crystal Growth* [in Russian], Nauka, Moscow (1980), Vol. 13, pp. 68–79.

[2] Yu. A. Vlasov, O. L. Golubev, and V. N. Shrednik, *Crystal Growth* [in Russian], Nauka, Moscow (1991), Vol. 19, pp. 5–21.

[3] V. G. Butenko, Yu. A. Vlasov, O. L. Golubev, and V. N. Shrednik, Surf. Sci. **266**, 165 (1992).

[4] O. L. Golubev, E. L. Kontorovich, and V. N. Shrednik, Zh. Tekh. Fiz. **66**(3), 88 (1996) [Tech. Phys. **41**, 279 (1996)].

[5] Yu. A. Vlasov, V. G. Pavlov, and V. N. Shrednik, Pis'ma Zh. Tekh. Fiz. **12**, 548 (1986) [Sov. Tech. Phys. Lett. **12**, 224 (1986)].

[6] V. G. Butenko, O. L. Golubev, E. L. Kontorovich, and V. N. Shrednik, Pis'ma Zh. Tekh. Fiz. **18**(8), 86 (1992) [Sov. Tech. Phys. Lett. **18**, 275 (1992)].

[7] O. L. Golubev, E. L. Kontorovich, and V. N. Shrednik, Zh. Tekh. Fiz. **66**(3), 97 (1996) [Tech. Phys. **41**, 284 (1996)].

[8] M. V. Loginov, O. G. Savel'ev, and V. N. Shrednik, Zh. Tekh. Fiz. **64**(8), 123 (1994) [Tech. Phys. **39**, 811 (1994)].

[9] A. P. Komar and Yu. N. Talanin, Izv. Akad. Nauk Ser. Fiz. **22**(5), 580 (1958).

[10] O. L. Golubev, B. M. Shaĭkin, and V. N. Shrednik, Pis'ma Zh. Tekh. Fiz. **1**, 714 (1975) [Sov. Tech. Phys. Lett. **1**, 313 (1975)].

[11] V. N. Ageev, E. Yu. Afanas'eva, N. R. Gall' et al., Pis'ma Zh. Tekh. Fiz. **12**, 565 (1986) [Sov. Tech. Phys. Lett. **12**, 231 (1986)].

[12] W. W. Dolan, Phys. Rev. **91**, 510 (1953).

[13] V. O. Fomenko, O. L. Golubev, and V. N. Shrednik, *Handbook of Emission and Absorption Properties of Materials* [in Russian], Atomizdat, Moscow (1975).

[14] Yu. A. Vlasov, O. L. Golubev, and V. N. Shrednik, Izv. Akad. Nauk SSSR, Ser. Fiz. **52**, 1538 (1988) [in Russian].

Translated by F. Crowne

A method for determining errors of a static fiber-optic gyrocompass

I. A. Matisov, V. E. Strigalev, V. A. Nikolaev, and Yu. V. Ivanov

M. A. Bonch-Bruevich St. Petersburg State Telecommunications University, 191186 St. Petersburg, Russia

The influence of intrinsic noise and signal drift on the accuracy of a static fiber-optic gyrocompass (FOG) is analyzed theoretically, and a method is proposed to determine the errors of the gyrocompass experimentally. From this it is possible to choose an optimum algorithm for measuring the FOG signal and calculating the direction angle with respect to north. Real measurements of the accuracy of an operating gyrocompass yield errors of order 1° (rms). © *1997 American Institute of Physics.* [S1063-7842(97)02309-X]

INTRODUCTION

The so-called ''static'' gyrocompass design[1,2] is based on a high-sensitivity fiber-optic gyroscope (FOG) which detects the projection of the Earth's rotation vector onto its own sensitivity vector. It then uses the resulting signal to determine a direction angle with respect to north. The FOG is placed on a rotating platform such that its sensitivity vector lies in the plane of the horizon for any rotation of the platform. The accuracy with which the angle of rotation of the platform is monitored exceeds the accuracy required to determine the direction angle relative to north. The platform and the FOG placed on it can turn relative to one another through fixed angles chosen according to a special algorithm. The FOG signal is measured after each of these turns, and an analog-digital converter is used to load the measured value into a computer where the direction angle relative to north is calculated.

In our paper Ref. 3 we conducted a comparative analysis of the algorithms used to compute direction angles relative to north. The parameter we used to compare these algorithms theoretically was the mean-square error in computing the angle. In this paper, the only source of inaccuracy we considered was the intrinsic noise of the FOG. As for drift of the output FOG signal, we assumed that it was either absent or could be approximated by linear or quadratic functions. In the latter case, the effect of signal drift can be calculated separately from additional measurements of the FOG signal and then taken into account in computing the angle. In general, the drift of the output FOG signal is an unpredictable function of time which does not yield to a unique approximation, and therefore its exact calculation is impossible.

The goal of this work is to study the combined influence of intrinsic noise and FOG signal drift on the accuracy of the static gyrocompass scheme, and to propose a method for determining the errors of the operating scheme.

INFLUENCE OF FOG CHARACTERISTICS ON THE ACCURACY OF THE STATIC GYROCOMPASS SCHEME

Several algorithms are known for measuring the FOG signal and calculating the direction angle with respect to north[1,4] that can deal with a constant shift in the output voltage of the high-sensitivity FOG. They do this by measuring the FOG signal three times in a row, rotating the FOG between measurements. In the first method, the rotation step is 90° (see Ref. 1) while in the second it is 120° (see Ref. 4). In our Ref. 3 we showed that if only intrinsic noise in the FOG is taken into account the first algorithm allows the angle to be computed with high accuracy. It is obvious that the use of this method is preferable from the point of view of decreasing the total measurement time, which also implies a decrease in the influence of signal drift of the FOG on measurement accuracy. In this method, the direction angle relative to north Θ_p is computed from the following expression:

$$\Theta_p = \tan^{-1}\frac{S_2 - S_3}{S_2 - S_1} - 45°, \tag{1}$$

where S_1, S_2, S_3 are the values of the FOG signal measured after successive rotations of the FOG by 90°, i.e.,

$$S_1 = U_0 \sin \Theta, \quad S_2 = U_0 \cos \Theta, \quad S_3 = -U_0 \sin \Theta. \tag{2}$$

Here Θ is the angle between the sensitivity vector of the FOG and the projection of the Earth's rotation vector onto the horizontal plane, and U_0 is the maximum amplitude of the FOG signal, which records the rotation of the Earth at a given point on the Earth's surface.

According to our theoretical calculations, which we described in Ref. 3, the mean-square error in calculating the angle Θ_p depends on the value of the intrinsic FOG noise as follows:

$$\sigma\Theta_p = \frac{\sigma U}{U_0}\sqrt{1 - 0.5\cos 2\Theta}\,(\text{rad}),$$

where σU is the mean-square measurement error for the FOG signal at a single position determined by the intrinsic noise of the FOG.

It is clear from this expression that the error in calculating Θ_p depends on the true value of Θ, and is a minimum when $\Theta = 0$.

We cannot obtain an exact analytic expression for the way the error in calculating Θ_p depends on the value of the FOG signal drift. However, if we postulate some behavior for the drift that is close to reality, we can obtain sufficiently precise dependences for the error in calculating Θ_p as a function of Θ and the rate of change of the FOG signal drift. We did this for two typical cases of drift behavior. In the first

case, the drift was a linearly varying function over the entire measurement time; in the second case, we assumed that the time derivative of the signal drift changes sign within the total measurement time so that the absolute value of the drift in the signal has the same value at the end of the measurement as it does at the beginning of the measurement. The absolute error in computing Θ_p in the first and second cases are respectively

$$\Theta - \Theta_p \approx \frac{DT}{2U_0} \cos \Theta,$$

$$\Theta - \Theta_p \approx \frac{DT}{2U_0} \sin \Theta \, (\text{rad}),$$

where D is the time derivative of the FOG signal drift at the beginning of the measurement cycle and T is the total measurement time of the FOG signal.

The relative error of these expressions is smaller than the ratio $DT/2U_0$. In general, the optimum value of Θ relative to which it is better to measure the FOG signal and then calculate Θ_p is determined by the real FOG characteristics, i.e., its intrinsic noise and the drift of its output signal.

EXPERIMENTAL STUDY OF THE ACCURACY OF A STATIC FIBER- OPTIC GYROCOMPASS

The sample FOG we used was based on a single-mode polarization-preserving fiber with the following characteristics: scaling coefficient 270 mV/(deg/h), intrinsic noise 0.2 $(\text{deg/h})/\sqrt{\text{Hz}}$, and output signal drift 0.2 deg/h. The first stage of development of the static design scheme for a fiber-optic gyrocompass should include a computer simulation of the calculation of directional angles relative to north. The simulation method we used was described in Ref. 5. Following this method, we created long-period files (up to two hours) recording the behavior of the output signal from our FOG under vibration-isolated conditions. Then this sample was split up into intervals of equal duration t corresponding to the time for a single measurement of the FOG signal, and took averages over them. The measurement intervals alternated strictly with intervals long enough to rotate the FOG from one position to another and for transient processes arising from vibrations at the instant of rotation to die away (total time 5 seconds). To the signal obtained in this way we added a theoretically calculated signal based on Eq. (2) for $\Theta = -45°$. The calculation of Θ_p was based on Eq. (1). Figure 1 shows the behavior of the mean-square error in calculating the angle Θ_p as a function of time t. It is clear that for the FOG sample we used the error in calculating Θ_p reaches a minimum for a measurement time (averaged) of 10 seconds; this error is a quantity of order 0.5°. It is obvious that for this sample FOG we cannot obtain higher accuracy for the gyrocompass for a single calculation of Θ_p.

At the next stage we made measurements on an actual gyrocompass. As we showed in Ref. 3, we can use an algorithm for calculating Θ_p that allows us to first estimate the value of the drift and then include it in the calculations of Θ_p. In order to check this possibility, a single-type cycle of measurements of the FOG signal was repeated 30 times. Each cycle consisted of four measurements of the FOG sig-
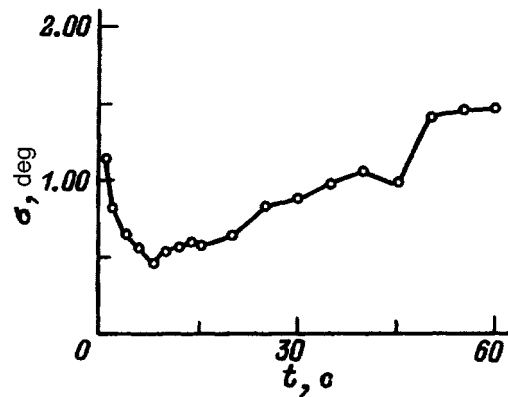

FIG. 1. Computational error Q_p in simulating the gyrocompass.

nal lasting 15 seconds. Between measurements the FOG was rotated in one direction by an angle of 90° relative to the previous measurement (the time between measurements of the FOG signal was 5 seconds). In all the measurements, the FOG sensitivity vector remained roughly at an angle of 45° to the north-south axis. The results of the measurements were stored in the computer. Then the angle Θ_p was calculated according to Eq. (1) by three methods (Fig. 2). In the first method only the first three measured values of the FOG signal for each cycle were used to calculate Θ_p, while in the second method only the second, third, and fourth measured values of the FOG signal were used. In the third method, based on four measurements of the FOG signal in a single cycle, we first estimated the drift in the FOG signal under the assumption that its variation was linear in nature: $d = 0.5(S_2 + S_4 - S_1 - S_3)$. Then we computed the angle Θ_p based on the value of the FOG signal corrected by this calculated value of d. The mean-square error in calculating the angle Θ_p was roughly 1.5° for the first and second methods, while in the third method it was 2° This implies that for our FOG the best algorithm for calculating Θ_p was the simple algorithm that uses three measurements of the FOG signal with rotations of the FOG by 90°. The FOG rotations have an additional instability in the temperature regime of operation of this sample FOG, which increases its intrinsic noise and drift. We can use this to explain the increase in error in
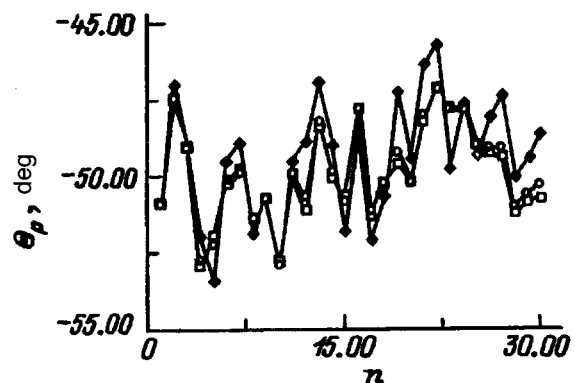

FIG. 2. Comparison of algorithms for calculating Q_p. Methods: $\bigcirc$ — 1, $\square$ — 2, $\blacklozenge$ — 3.
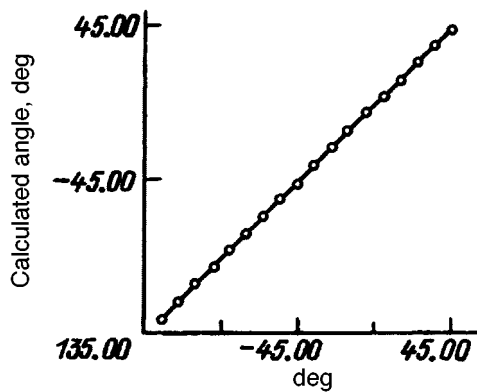
FIG. 3. Calculated Q_p versus the angle of rotation of the first position of the FOG.

FIG. 4. Computational error Q_p versus the angle of rotation of the first position of the FOG.

calculating Θ_p compared to the error obtained previously in the simulation process.

The third stage of our study of the gyrocompass accuracy was to measure its error as a function of the value of the calculated angle Θ_p. For this we measured the FOG signal as we did previously, except that each new measurement cycle of four values of the FOG signal began with a shift of the first measurement position by $10°$ with respect to the first position of the previous measurement cycle. Thus, we made nine measurement cycles in succession, after which the first measurement position coincided with the original position, and then made nine more measurement cycles, doing so 30 times. The results of the measurements were also accumulated in the computer.

Then, based on the results of the measurements, we calculated the angle Θ_p using Eq. (1) for eighteen orientations of the FOG sensitivity vector in the first position for measurements relative to the north-south axis. The first half of this range of values was calculated from the first three measured values of the FOG signal of each cycle, while the second half was based on the second, third and fourth values of the FOG signal. Figures 3 and 4 show plots based on 30 calculations of the average value of the angle Θ_p and the mean-square deviation of the calculated angle respectively. We then passed a straight line through the calculated values of the angle Θ_p, as shown in Fig. 3. The deviation of the computed values from a straight line was a quantity smaller than the mean-square deviation shown in Fig. 4. The plot shown there has several obvious extrema. The minimum value of the error in computing the angle Θ_p was a quantity of order $0.6°$ for a total measurement time of order 1 minute. The analogous result given in Ref. 1 was $0.2°$ for the same measurement time. The character of the function shown in Fig. 4 confirms the conclusions stated in the first part of this paper. The lower error in computing the angle Θ_p within the range $-30°$ to $0°$ is explained by the smaller influence of the intrinsic FOG noise on the accuracy of computing the angle Θ_p over this range of angles. Likewise, the lower computational error within the range $-120°$ to $-90°$ is explained by the smaller influence of the FOG signal drift on the accuracy of computing Θ_p within this range. From this plot we may conclude that it is preferable to orient the FOG sample with respect to the north-south axis in such a way that the com-
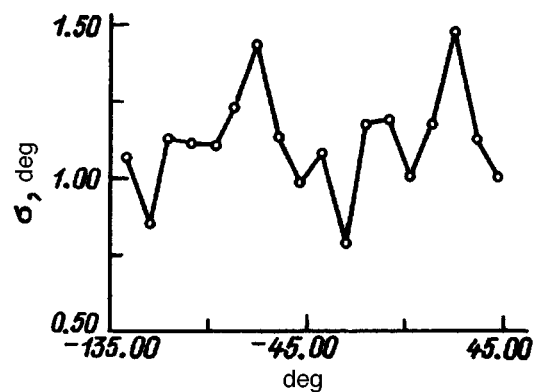
puted angle Θ_p falls on a value close to $-30°$. This orientation of the FOG during the operation of the gyrocompass can be implemented after the first preliminary measurement of the FOG signal and calculation of the directional angle relative to north, possibly based on a simplified algorithm or within a shorter measurement time. The smaller error in calculating Θ_p for the angle $-45°$ compared to the results obtained in the previous series of measurements is explained by the fact that this test series had a considerably larger measurement time (of order six hours), which led to a decrease in the intrinsic noise and drift of the signal during the measurements, since the sample FOG was warmed up.

CONCLUSION

Based on this work, we propose a method for determining the measurement error of a static fiber-optic gyrocompass.

The first stage of this method is to determine the ultimate accuracy of the gyrocompass for the sample FOG used and the optimum time for a single measurement of the FOG signal. For our sample FOG we obtained an accuracy of order $0.5°$ (rms) and a measurement time of 10 seconds respectively. The second stage is to determine the optimum algorithm for measuring the FOG signal and compute the directional angle relative to north. For our sample FOG we concluded that it was necessary to use an algorithm without compensation of the drift, with a minimum number of measurements of the FOG signal equal to three. The third stage is to determine that value of the angle relative to the direction of north for which the computational error is a minimum. For our sample FOG this optimum angle was $-30°$.

We determined how accurately the angle relative to the direction of north was calculated in the course of making the test measurements and calculations. For our sample fiber-optic gyrocompass the error was a quantity of order $1°$ (rms) or smaller.

[1] R. B. Dyott and D. E. Allen, in *Proceedings of the Tenth Optical Fibre Sensors Conference*, Glasgow, Scotland (1994), p. 442.
[2] T. Tanaka and Y. Igarashi, Appl. Opt. **33**, 120 (1994).
[3] I. A. Matisov, V. A. Nikolaev, and V. E. Strigalev, in *Proceedings of the Educational Institutes of Communications* [in Russian], St. Petersburg (1996), Vol. 162, p. 92.

[4] V. N. Logozinski and V. A. Solomatin, in *Proceedings of the Second International Conference on Gyroscopic Engineering and Navigation* [in Russian], St. Petersburg, 1995.

[5] Yu. V. Ivanov, V. E. Strigalev, V. A. Nikolaev, and I. A. Matisov, in *Proceedings of the Educational Institutes of Communications* [in Russian], St. Petersburg (1996), Vol. 162, p. 87.

Translated by F. Crowne

Thermal stability of the magnetic parameters of epitaxial iron garnet films subjected to planar radial stresses

V. T. Dovgiĭ, T. G. Astaf'eva, F. G. Bar'yakhtar, and G. I. Yampol'skaya

Donetsk Physico-technical Institute, Ukrainian Academy of Sciences, 340114 Donetsk, Ukraine

The effect of external planar radial pressure on the thermal stability of the magnetic parameters of epitaxial iron garnet films is investigated in the temperature range 200–500 K for external mechanical stresses in the range 0–40 kgf/mm^2. It is shown that external planar radial pressure can be used to improve the thermal stability of these magnetic parameters by a factor of 1.5–2, and also to alter significantly the temperature interval of single-domain behavior for orientational phase transitions near a compensation point. © *1997 American Institute of Physics.* [S1063-7842(97)02409-4]

One of the important requirements imposed on magnetic-bubble materials is temperature stability of their magnetic parameters.[1–11] In this paper we investigate the effect of an external planar radial pressure on the thermal stability of these parameters for iron garnet films with the compositions $(YSmLuCa)_3(FeGe)_5O_{12}$ (1), $(YSmLu)_3(FeGa)_5O_{12}$ (2), and $(YGdTm)_3(FeGa)_5O_{12}$ (3), grown on (111)-oriented $Gd_3Ga_5O_{12}$ substrates. Using magnetooptic devices based on the Faraday effect, we studied these films in the temperature range 200–500 K under external mechanical stresses in the range 0–40 kgf/mm^2. In this paper we show that external planar radial stresses can significantly improve the thermal stability of the magnetic parameters of epitaxial iron garnet films.

Figure 1 shows the temperature dependence of the characteristic length l, the wall energy σ_w, the saturation magnetization $4\pi M_s$, the anisotropy field H_A, the exchange constant A, the quality factor Q, the half period of the stripe structure $P_0/2$, the magnetic- bubble collapse field H_k, and

the period a and magnetic-bubble diameter d of a lattice of magnetic bubbles for films with composition $(YSmLuCa)_3(FeGe)_5O_{12}$ in an external stress $\sigma_{ex} = 0$ (open circles) and $\sigma_{ex} \neq 0$ (filled circles).

Table I lists various temperature coefficients, defined as $P = [(dP/dT)/P] \cdot (100\%)$ where P is any of the (generalized) static magnetic parameters mentioned above, for epitaxial iron garnet films with $\sigma_{ex} = 0$ kgf/mm^2 and $\sigma_{ex} = 18$ kgf/mm^2. The temperature dependence of the characteristic length l is a minimum when the variations in $4\pi M_s$ and σ_w compensate one another, i.e., when $\Delta M_s/M_s = \Delta\sigma_w/2\sigma_w$.[4] Figure 2 and Table II list analogous temperature dependences of the magnetic parameters and their temperature coefficients for films with the composition $(YSmLu)_3(FeGa)_5O_{12}$.

Epitaxial iron garnet films with composition $(YGdTm)_3(FeGa)_5O_{12}$ having a compensation point close to room temperature (~ 215 K). Fig. 3 illustrates the change in magnetic parameters of these films with temperature, while
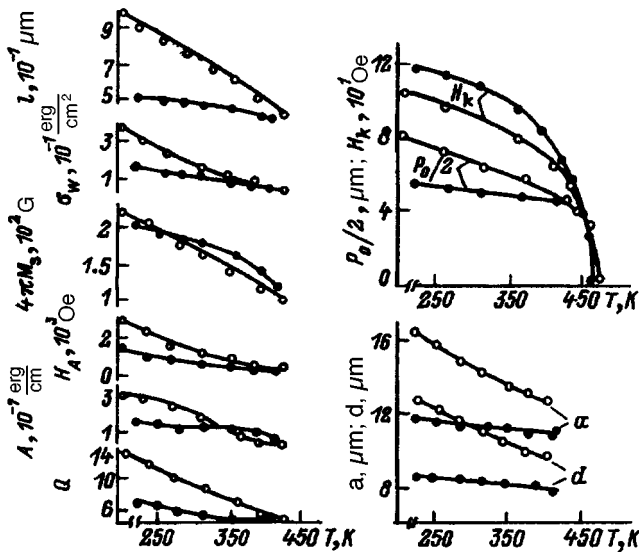

FIG. 1. Temperature coefficients of the magnetic parameters of epitaxial iron garnet films with the composition $(YSmLuCa)_3(FeGe)_5O_{12}$ for $\sigma_{ex}=0$ (open circles) and $\sigma_{ex}\neq 0$ (filled circles).
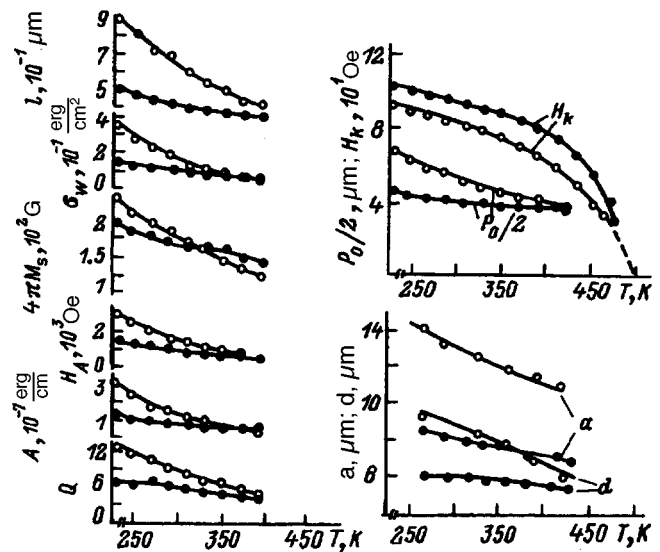
FIG. 2. Temperature coefficients of the magnetic parameters of epitaxial iron garnet films with the composition $(YSmLu)_3(FeGa)_5O_{12}$ for $\sigma_{ex}=0$ (open circles) and $\sigma_{ex}\neq 0$ (filled circles).

TABLE I. Temperature Coefficients of the Magnetic Parameters of Films with the Composition $(YSmLuCa)_3(FeGe)_5O_{12}$

σ_{ex}, kgf/mm²	T,K	l_T,%/K	σ_T,%/K	M_T,%/K	H_A,%/K	A_T,%/K	H_k,%/K	$P_0/2$,%/K	a_T,%/K	d_T,%/K	K_u,%/K	$\frac{\Delta\sigma_w}{\sigma_w}/\frac{\Delta M}{M}$
0	293	−0.32	−0.82	−0.28	−0.73	−0.4	−0.17	−0.22	−0.14	−0.13	−0.92	2.7
18	293	−0.16	−0.48	−0.16	−0.29	−0.35	−0.11	−0.07	−0.07	−0.05	−0.67	2.2

TABLE II. Temperature Coefficients of the Magnetic Parameters of Films with the Composition $(YSmLu)_3(FeGa)_5O_{12}$

σ_{ex}, kgf/mm²	T,K	l_T,%/K	σ_T,%/K	M_T,%/K	H_A,%/K	A_T,%/K	H_k,%/K	$P_0/2$,%/K	a_T,%/K	d_T,%/K	K_u,%/K	$\frac{\Delta\sigma_w}{\sigma_w}/\frac{\Delta M}{M}$
0	293	−0.43	−1.6	−0.34	−0.87	−0.94	−0.15	−0.30	−0.26	−0.22	−1.19	2.9
15	293	−0.19	−0.5	−0.16	−0.44	−0.4	−0.09	−0.12	−0.1	−0.13	−0.60	2.9

Table III lists their temperature coefficients at room temperature (293 K) and 353 K respectively for initial films ($\sigma_{ex}= 0$ kgf/mm²) and for $\sigma_{ex}= 16$ kgf/mm².

Analysis of the plots for films with the first two compositions show that the magnetic parameters of the initial films ($\sigma_{ex}= 0$) and those subjected to external stress ($\sigma_{ex}\neq 0$) decrease monotonically with temperature with a dependence that is close to linear (except for a region around the Néel point). Application of a radial stress improves the thermal stability of the magnetic parameters of these films. A criterion of temperature stability of the magnetic parameters (characteristic length l, saturation magnetization $4\pi M_s$, boundary energy density σ_w, bubble collapse field H_k, and equilibrium period of the stripe domain structure P_0) was formulated in Refs. 1,2,8 for $|P_T|\leqslant 0.2$–0.3 %/K. From these tables it is clear that the temperature coefficients of the magnetic parameters for films subjected to stress decrease by a factor of 1.5–2 and satisfy better the criterion of temperature stability.

The criteria of small changes in the characteristic length l (see Ref. 4) is $\Delta M/M = \Delta\sigma/2\sigma_w$, where $\Delta M/M$ and $\Delta\sigma/\sigma_w$ are relative changes in the saturation magnetization and domain wall energy respectively. For films with the first and second composition, at $\sigma_{ex}= 0$ and $\sigma_{ex}\neq 0$ we find that $(\Delta\sigma/\sigma)/(\Delta M/M)=2.62$–$3.1$.

The thermal stability of the magnetic parameters under the action of an external stress is improved because the change of the magnetic parameters with respect to temperature is compensated by their change with respect to stress.

In Ref. 12 it was shown that the values of the magnetic parameters decrease with increasing external stress (the derivative is negative), analogous to changes in these parameters with temperature. If a film is under a certain external stress at room temperature which decreases as the temperature increases and vice versa, then there is a certain temperature range in which the change in magnetic parameters with
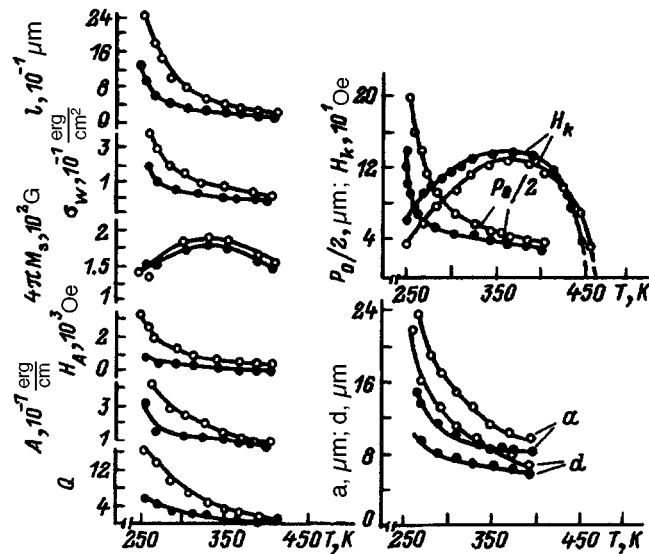

FIG. 3. Temperature coefficients of the magnetic parameters of epitaxial iron garnet films with the composition $(YGdTm)_3(FeGa)_5O_{12}$ for $\sigma_{ex}= 0$ (open circles) and $\sigma_{ex}\neq 0$ (filled circles).

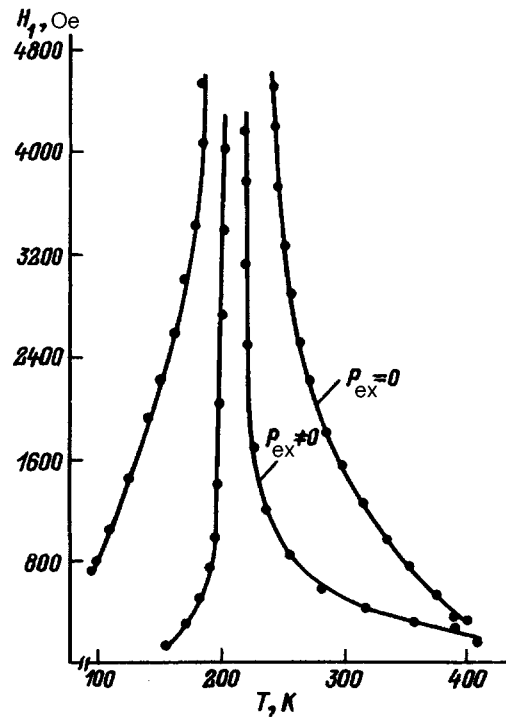FIG. 4. Dependence of the saturation magnetization of a film in the planar field $H_\perp$ on temperature near the compensation point for $\sigma_{ex}= 0$ and $\sigma_{ex}\neq 0$.

TABLE III. Temperature Coefficients of the Magnetic Parameters of Films with the Composition $(YGdTm)_3(FeGa)_5O_{12}$

σ_{ex}, kgf/mm^2	T,K	l_T,%/K	σ_T,%/K	M_T,%/K	H_A,%/K	A_T,%/K	H_k,%/K	$P_0/2$,%/K	a_T,%/K	d_T,%/K	K_u,%/K
0	293	-1.91	-0.98	$+0.51$	-1.53	-1.08	$+1.37$	-1.3	-0.93	-0.93	-0.86
16		-1.38	-0.55	$+0.38$	-1.06	-0.62	$+0.67$	-0.66	-0.65	-0.68	-0.48
0	353	-1.37	-1.47	-0.11	-1.52	-1.36	$+0.31$	-0.75	-0.53	-0.54	-1.57
16		-0.82	-0.95	-0.08	-0.97	-0.88	$+0.05$	-0.41	-0.22	-0.35	-1.04

respect to temperature and stress will compensate each other $(-dP/dT = +dP/d\sigma)$. The improvement in thermal stability of the parameters of the domain structure is primarily due to the change in uniaxial anisotropy under the action of the external stress.

Because epitaxial iron garnet films with composition $(YGdTm)_3(FeGa)_5O_{12}$ have a point of compensation near room temperature, the temperature coefficients of their magnetic parameters at 293 and 353 K are worse than for films with compositions 1 and 2. The thermal stability of magnetic parameters for films with composition 3 is improved by the presence of an external planar radial stress (Fig. 3 and Table III). In the range of temperatures close to the Néel point and the compensation point, the changes in magnetic parameters with respect to temperature are large (especially M_s and H_A), and are not compensated by their changes with respect to pressure.

Figure 4 shows the temperature dependence of the saturation magnetization of a film with composition $(YGdTm)_3(FeGa)_5O_{12}$ in a planar field $H_\perp$ near the compensation point for $\sigma_{ex} = 0$ and $\sigma_{ex} \neq 0$. The single-domain interval in the neighborhood of the compensation point for $\sigma_{ex} = 0$ has a width $\Delta T = 60$ K, while for $\sigma \approx 25$ kgf/mm^2 it is $\Delta T = 10$ K ($\sigma = 25$ kgf/mm^2 corresponds to a temperature of 210 K). Thus, the temperature interval for single-domain behavior of this epitaxial iron garnet film (the state with uniform magnetization) decreases by a factor of 6 under stress; this narrowing of the single-domain region is prima-rily due to changes in the uniaxial anisotropy under planar radial stress. The shift in the compensation point (defined as the midpoint of the region of single-domain behavior with respect to the saturation field) is insignificant, ≈ 5 K.

Thus, we have shown that external planar radial stress can improve the thermal stability of magnetic parameters of iron garnets, and also significantly alter the temperature interval of single-domain behavior near a compensation point for orientational phase transitions.

[1] R. V. Telesnin, V. N. Dudorov, and V. V. Randoshkin, Fiz. Tverd. Tela (Leningrad) **17**, 3015 (1975) [Sov. Phys. Solid State **17**, 2000 (1975)].
[2] R. V. Telesnin, V. N. Dudorov, A. T. Marcheko, and V. V. Randoshkin, Mikroelektronika **8**(1), 84 (1979).
[3] N. Ohta, F. Ishida, T. Ikeda, and Y. Sugita, J. Appl. Phys. **51**, 589 (1980).
[4] D. M. Smith and A. W. Anderson, AIP Conf. Proc. **5**, 120 (1972).
[5] M. J. Dimayan and E. Della Torre, J. Appl. Phys. **43**, 1285 (1972).
[6] M. J. Dimayan and E. Della Torre, J. Appl. Phys. **43**, 5207 (1972).
[7] K. Yamaguchi, M. Uchishiba, and T. Suzuki, IEEE Trans. Magn. **MAG-16**, 616 (1980).
[8] J. W. Nielsnin, IEEE Trans. Magn. **MAG-12**, 327 (1976).
[9] K. P. Belov, M. A. Belyanchikova, R. Z. Levitin, and S. A. Nikitin, *Rare-Earth Ferro- and Antiferromagnets* [in Russian], Nauka, Moscow (1965).
[10] D. M. Gualtieri, P. F. Tumelty, and M. A. Gilleo, J. Appl. Phys. **50**, 7824 (1979).
[11] M. Kestigian, A. B. Smith, and W. R. Bekebrede, J. Appl. Phys. **49**, 1873 (1978).
[12] V. G. Baryakhtar, V. T. Dovgiĭ, B. Ya. Sukharevskiĭ *et al.*, Fiz. Tverd. Tela (Leningrad) **25**, 1415 (1983) [Sov. Phys. Solid State **25**, 813 (1983)].

Translated by F. Crowne

Effect of external noise on the optical thermal breakdown of semiconductor wafers

Yu. V. Gudyma

Yu. Fed'kovich Chernovtsy State University, Chernovtsy, Ukraine

The optical thermal breakdown of a semiconductor wafer with a small Biot number under conditions of fluctuating incident light intensity is discussed. It is shown that external noise shifts the breakdown region toward somewhat higher values of the control parameter, and as the fluctuations grow this process leads to suppression of the critical point. Light intensities below the threshold noise can induce a nonequilibrium phase transition. Numerical estimates are given for germanium illuminated by a CO_2 laser. © *1997 American Institute of Physics.*
[S1063-7842(97)02509-9]

Optical thermal breakdown is itself an example of a cooperative optical effect. It occurs in systems excited by light whose distinctive peculiarity is the appearance of multiple stable states when the system parameters exceed certain critical values, even in the absence of external feedback.[1] Physically, this phenomenon is associated with the rapid (exponential) growth of the coefficient of free-carrier absorption of light (when the energy of the optical quantum is smaller than the width of the band gap) for a semiconductor heated by previously absorbed light. It manifests itself in an avalanche-like rise in the semiconductor temperature when the light intensity (the external parameter) exceeds the breakdown threshold. As an example of a dissipative structure that appears in an open system far from a state of thermodynamic equilibrium,[2,3] optical thermal breakdown is valuable in that it has both been observed experimentally[4] and explained theoretically.[5] Modeling of this phenomenon demonstrates distinctive features of hysteresis in distributed systems, in particular the existence of a switching wave between states corresponding to different branches of the hysteresis function.[6] Thus the description of optical thermal breakdown encompasses almost all nonequilibrium effects.

Let us discuss how the state of the system changes under the influence of fluctuations in the intensity of the incident light. Note that this system is especially convenient for experimental studies of these effects. The intensity of the incident light is an external parameter that can be varied easily in the experiment over a considerable range. Also important is the fact that the light intensity is a multiplicative parameter in the Lambert–Beer law. From this we infer that fluctuations will have a strong effect on the system, because, as is well known, we can completely control the dominant nonequilibrium properties of the system by changing the energy flux of the incident light. In order to neglect internal fluctuations, we will start with a semiconductor wafer that is spatially uniform with a small Biot number $Bi = H\lambda/l \ll 1$ (where l is the wafer thickness, λ is the thermal conductivity, and H is the heat transfer coefficient, equal to the ratio of the specific thermal conductivity to the specific material density of the material per unit length). Assume also that the transverse thermal resistance of this wafer is small compared to the thermal resistance of the contact with the heat bath, which means a rapid equalization of the temperature over the wafer thickness. Then it is sufficient to describe the nonstationary thermal regime of the system dynamics by a single variable — the temperature T, whose value is determined from the heat balance equation:[7]

$$c\frac{dT}{dt} = G(x,t)\{1-\exp(-\alpha(T)l)\}/l - H(T-T_0), \quad (1)$$

where $G(x,t)$ is the intensity of a broad beam of incident light, c is the specific heat of the semiconductor material, T_0 is the temperature of the heat bath,

$$\alpha(T) = \alpha_0 \exp(-E_g/2kT) \quad (2)$$

is the absorption coefficient of light by free carriers, and E_g is the width of the band gap.

The temperature dependence of the magnitude of the absorption coefficient α_0 is usually neglected at high temperatures. It is not difficult to transform from (1) to dimensionless variables

$$\frac{d\theta}{d\tau} = \beta\{1-\exp(-\eta\exp(\gamma-\gamma/\theta))\} - (\theta-1), \quad (3)$$

where $\theta = T/T_0$, $\tau = tH/c$, $\gamma = E_g/2kT_0$, and $\eta = \alpha(T_0)l$.

In what follows we will also assume that the correlation time for external fluctuations is negligibly small, so that we may use the white-noise approximation to describe them. Let us replace the parameter β by a stationary random process,[8] i.e., $\beta(t) = \beta + \sigma\xi(t)$, where the external noise $\xi(t)$ has zero average value and a correlation function $\langle\xi(t)\xi(t')\rangle = \sigma^2\delta(t-t')$, while β equals the average value of $\beta(t)$. Taking into account that Gaussian white noise is the derivative of a Wiener process in the sense of generalized functions, we can transform Eq. (3) into Stratonovich's stochastic differential equation, which determines the evolution of the transition probability density $p(\theta,\tau|\theta',\tau')$. This in turn leads to the corresponding Fokker–Planck equation:

$$\frac{\partial p(\theta,\tau|\theta')}{\partial\tau}$$

$$= -\frac{\partial}{\partial\theta}\left[\beta\{1-\exp(-\eta\exp(\gamma-\gamma/\theta))\} - (\theta-1)\right.$$

$$-\frac{\eta\gamma\sigma^2}{2\theta^2}\{1-\exp(-\eta\exp\times(\gamma-\gamma/\theta))\}$$

$$\times\exp(-\eta\exp(\gamma-\gamma/\theta))\bigg\}$$

$$\times\exp(\gamma-\gamma/\theta)]p(\theta,\tau|\theta')+\frac{\sigma^2}{2}\frac{\partial}{\partial\theta^2}$$

$$\times\{1-\exp(-\eta\exp(\gamma-\gamma/\theta))\}^2 p(\theta,\tau|\theta'). \qquad (4)$$

Stationary regimes that describe macroscopic behavior of a nonlinear system driven by external noise are adequately and simply described by the stationary probability density $p(\theta)$. If the boundary of the state space is natural (or regular with instantaneous reflection), then the stationary solution to the uniform Fokker–Planck equation is

$$p_s(\theta)=N\{1-\exp(-\eta\exp(\gamma-\gamma/\theta))\}^{-1}$$

$$\times\exp\left[\frac{2}{\sigma^2}\int_0^\theta[\beta\{1-\exp(-\eta\exp(\gamma-\gamma/u))\}\right.$$

$$\left.-(u-1)]\{1-\exp(-\eta\exp(\gamma-\gamma/u))\}^{-2}du\right], \qquad (5)$$

where N is a constant obtained from the normalization condition

$$\int_0^b p_s(\theta)d\theta=1. \qquad (6)$$

The upper limit of integration is determined by limitations imposed by physical conditions of the problem (for example, the melting point of the crystal). The solution to Eq. (5) is referred to as a potential solution, since its maximum corresponds to stable stationary states while its minima correspond to unstable stationary states.[8] Thus extrema of the stationary probability density can be identified with macroscopic stationary states of the system. Equation (5) was analyzed numerically for a germanium wafer with thickness
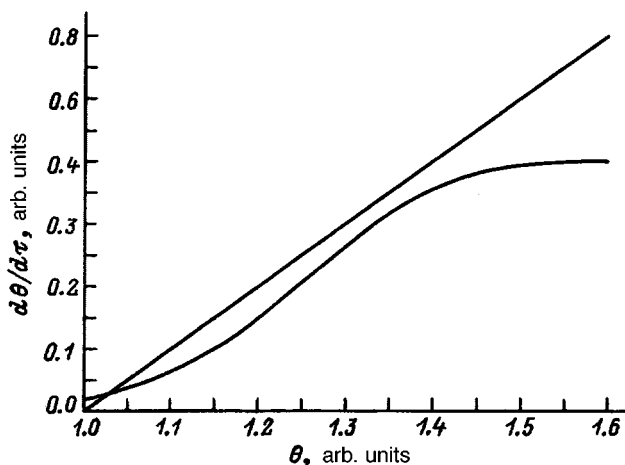

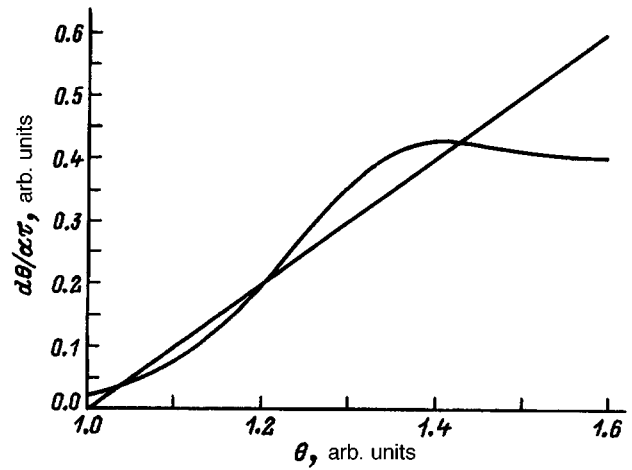FIG. 1. Graphical solution to the evolution equation (3) when $\beta=0.4$.

FIG. 2. Graphical solution to the evolution equation, taking into account fluctuations in the incident light with $\sigma^2=0.1$, for the same value $\beta=0.4$. A transition takes place when $\sigma^2=0.04$.

6 mm illuminated by a CO_2 laser with wavelength 10.6 μm at room temperature $T_0=300$ K. The material constants $E_g=0.67$ eV, $\alpha(T_0)=0.09$ cm^{-1} were taken from data given in Ref. 4. Recall that the melting point of germanium is 1210 K.

It is well known[6] that the thermal balance Eq. (3) has two stable deterministic stationary solutions for values of β above a certain critical value. The first of these corresponds to a phase in which the temperature of the system ceases to depend on the temperature of the external medium and is completely determined by the pump, while the second is the intrinsic phase of optical thermal breakdown. However, the action of external noise (both additive and multiplicative) can only shift the region of optical thermal breakdown towards somewhat larger values of the control parameter. Further increases in fluctuations in the external light intensity lead to suppression of the critical point. However, the primary question regarding qualitative readjustments of the regimes of this dynamic system, i.e., transitions induced by noise, remains unclear.[9] For this we consider the behavior of the system for values of the control (cooperative) parameter below the critical point. In the absence of external multiplicative noise this regime is characterized by a single stationary state (Fig. 1); however, increasing the fluctuations in the incident light leads to the appearance of bistability (Fig. 2). Thus we obtain a situation where a phase transition is induced exclusively by noise. Also important is the fact that predictions based on an analysis in which the noise is idealized as white noise are also justified for any colored noise in the limit of small correlation times.[9] It is worth emphasizing that the effect of external noise on optical thermal breakdown in semiconductor wafers is quite amenable to experimental simulation.

[1] H. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic Press, Orlando, 1985) [Russian trans., Mir, Moscow, 1988].
[2] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems* (Wiley, New York, 1977) [Russian trans., Mir, Moscow, 1979].
[3] H. Haken, *Synergetics: An Introduction. Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry, and Biology* (Springer-

Verlag, Berlin, 1977) [Russian trans., Mir, Moscow, 1980].

[4] P. A. Young, Appl. Opt. **10**, 638 (1971).

[5] E. M. Epshtein, Izv. Vyshch. Ucheb. Zaved. Radiofiz. **15**, 33 (1972).

[6] N. N. Rozanov, Zh. Éksp. Teor. Fiz. **80**, 96 (1981) [Sov. Phys. JETP **53**, 47 (1981)].

[7] E. M. Epshtein, Zh. Tekh. Fiz. **48**, 1733 (1978) [Sov. Phys. Tech. Phys. **23**, 983 (1978)].

[8] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer-Verlag, Berlin, New York, 1985) [Russian trans., Mir, Moscow, 1986].

[9] W. Horsthemke and R. Lefever, *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology* (Springer-Verlag, Berlin, New York, 1984) [Russian trans., Mir, Moscow, 1987].

Translated by F. Crowne

Salient features of the analysis of layered structures by means of ion sputtering

A. A. Dorozhkin, A. V. Filimonov, A. P. Kovarskiĭ, Yu. A. Kudryavtsev, and N. N. Petrov

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

Results are presented from an experimental investigation of layered structures made up of elements with strongly different atomic masses, using ion sputtering with post-ionization of secondary neutrals in a gas discharge plasma. An anomalous increase of the signal at boundaries between layers is noted. It is proposed that the phenomenon of selectivity in etching affects the measurement of the profile of secondary particle yield. © *1997 American Institute of Physics.* [S1063-7842(97)02609-3]

One of the most widely used methods of layer-by-layer analysis of materials is ion sputtering with monitoring of the composition of the bottom of the etch pit is monitored by one of several well-known methods of element surface analysis: secondary-ion mass spectroscopy, mass spectroscopy of ionized neutral particles, Auger electron spectroscopy, x-ray photoelectron spectroscopy, etc.[1] The first two methods are the ones most often used, since the composition of the sputtered particles in these cases is unaffected by the possibility of selective sputtering and reflects the bulk concentrations of the corresponding elements in the subsurface layer of the solid. In the presence of selective etching, the composition of the surface changes in such a way that the rate of etching of various components turns out to be proportional to their bulk concentration.[2]

We have made a systematic analysis of layered structures using the method of post-ionization of sputtered neutrals. In this case, a number of atypical effects are observed at a boundary between layers consisting of elements with significantly different masses. In this paper we present the results of these measurements, interpret them, and estimate possible manifestations of the effects described in the analysis.

The measurements were based on the method of post-ionization of sputtered neutrals in the plasma of a high-frequency gas discharge using an INA-3 apparatus made by Leibold-Hearues. The samples we used were irradiated by Ar^+ ions with energies 0.2–2 keV (the size of the beam was $\simeq 1$ mm at the target). After they leave the target, the sputtered neutral particles arrive in a region where a high-frequency gas discharge is created in argon, and there they are partially ionized; then they are accelerated and sent to a quadrupole mass analyzer ($M/\Delta M \simeq 500$). Thus, we are measure the mass spectra of secondary neutrals. In order to measure the profile of the distribution of some specified element with respect to depth, we measured the dependence of the flux of corresponding ions on the time in which the sample is irradiated by the ion beam (the sputtering time). Usually the sputtering rate was $\simeq 3$ Å/s under our conditions. Six elements could be measured simultaneously. In our measurements we recorded the yield of monatomic ions of the elements that make up the layered structure. For samples we used layered structures suitable for making electrical contacts to semiconductor electronics (Si/Pt, Au/Mo/Au/GaAs,

etc.). It is well known that the quality of such contacts is improved by preparing them in the form of layered structures. The samples were prepared by thermal deposition of metallic films on semiconductor surfaces and were subjected to subsequent processing. The layer-by-layer analysis of these structures with the goal of establishing distinctive features of the etching is the primary focus of this work.

Figure 1 shows the results of measurements on a sample of Si/Pt (the primary Ar^+ ions had an energy of 400 eV). For Pt^+ ions we observe a time dependence that is characteristic of layered systems: initially the Pt^+ signal is practically independent of t, while at a film boundary it abruptly drops down to the background level. The falloff time is determined by the possibility of interdiffusion of Pt and Si and the depth resolution of the monitoring method. For Si^+ and Si_2^+ ions the time dependence turned out to be somewhat different: in the Pt layer the signals were not large (with respect to the background), and in the Si layer they had a constant value characteristic of this element. However, at the boundary we observed a maximum in which the magnitude of the signal is considerably higher than the Si signal. Furthermore, the po-
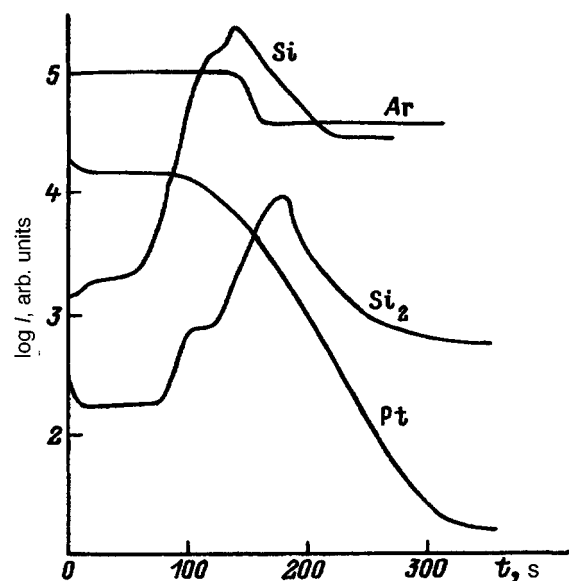

FIG. 1. Profile of the distribution of the elements in a layered Si/Pt structure. Primary particles - Ar^+, $E_p = 400$ eV.
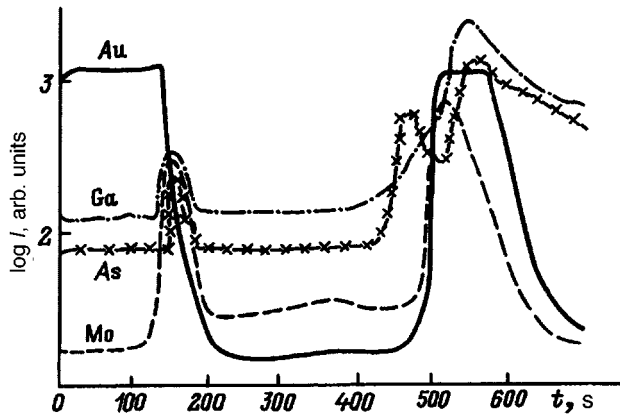
FIG. 2. Profile of the distribution of the elements in a layered Au/Mo/Au/GaAs structure. Primary particles - Ar^+, $E_p = 400$ eV.

sitions of the Si^+ and Si_2^+ maxima do not coincide, and attentive examination of the spectra reveal fine structure.

Analogous measurements were made for more complex systems: we deposited films of Mo and Au on the substrate of GaAs. The results of these measurements are shown in Fig. 2. The shape of the time dependence of the Au^+ signal corresponds to the profile of the distribution of this element in the structure. For Mo we trace two very significant maxima at the boundary with gold. Similar anomalies are observed for Ga^+ and As^+. For these elements we also observe maxima at the boundary with the Au film, on either side of the latter.

Based on this data and results obtained for a number of other samples, we can formulate the following general rule. Whereas the measured profiles for the heavier elements (Au, Pt) rather regularly reflect the distribution of these elements with respect to depth into the sample, maxima are observed for the lighter elements (Si, Ga, As, Mo) at boundaries with heavy elements. In a number of cases, the magnitude of the analysis signal is 5–8 times larger than for the corresponding pure element (Si, Ga, As, Mo). Increasing the ion energy to 2 keV attenuates this effect.

It is obvious that the effect we are observing is impossible to explain by anomalously high concentrations of one of the elements at the boundary between films. In contrast to secondary-ion mass spectroscopy, we cannot associate it with changes in the degree of ionization of secondary particles, since the ionization takes place in a plasma which is in no way connected with the surface. One natural possible explanation remains: an increase in the rate of sputtering of light elements at a boundary with heavy elements. However, we still have to find possible reasons for this phenomena of selective sputtering.

It is known, and also demonstrated by us in Ref. 2, that during sputtering of a mixture or a compound of light and heavy elements, a preliminary sputtering of atoms of the lighter element usually occurs. However, as the result of changes in the composition of the surface the rates of sputtering of different elements are found to be proportional to their bulk concentrations.[3] This process, in principle, requires a certain finite establishment time, and is described by an equation of type

$$\frac{dn}{dt} = \frac{j_0 S}{e N_0} n_{\text{bulk}} - \frac{j_0}{e} s_1 n,$$

whose solution is

$$n = \frac{S}{N_0} \frac{n_{\text{bulk}}}{s_1} \left[1 - \exp\left(-\frac{j_0}{e} s_1 t \right) \right],$$

where

$$\tau = \left(\frac{j_0}{e} s_1 \right)^{-1}$$

is a characteristic time for establishing the equilibrium state: j_0 is the current density of the primary beam, and s_1 is the probability for sputtering of lighter atoms located directly at the surface.

We can approximately estimate this quantity by taking $\sigma \simeq 10^{-16} \text{cm}^2$. Then $\tau \simeq 0.2 - 0.5$ s. For normal analysis of the elemental composition of the material, the time it takes to begin the measurement after switching on the ion beam greatly exceeds τ; therefore, selective etching is irrelevant and makes practically no contribution to these measurements of. However, a different situation arises when we analyze layered structures. Usually mixing of the components occurs at a boundary between layers within a comparatively thin transition layer ($\simeq 10$ Å) due to interdiffusion that takes place in the course of preparing the structure and its subsequent processing. Such interlayers of mixed composition are rapidly sputtered through, and equilibrium states cannot be established. Therefore, selectivity is exhibited in the etching of the lighter component at layer boundaries. As experiment shows, this selectivity can be quite significant.

To a certain extent, this assumption is confirmed by the dependence of the selectivity on the ion energy E_0: with increasing E_0 the selectivity of the sputtering should decrease.[3] We also made measurements at decreased sputtering rates. The results for one of the samples of Au/Mo/Au/
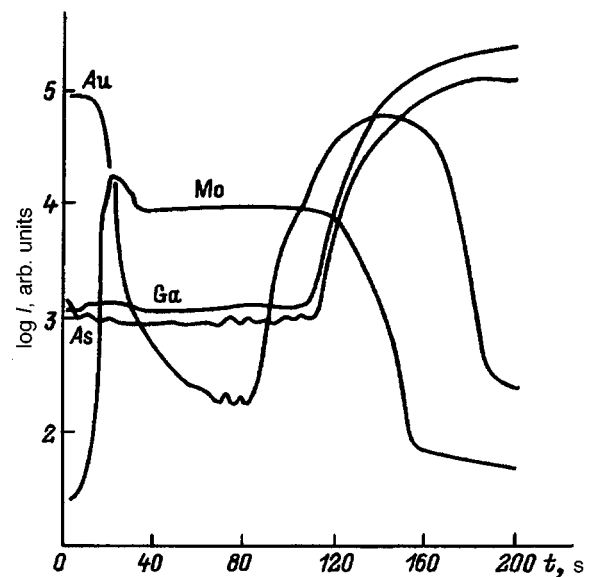

FIG. 3. Profile of the distribution of the elements in a layered Au/Mo/Au/GaAs structure. Primary particles - Ar^+, $E_p = 400$ eV. The sputtering rate has been lowered by a factor of 7.

GaAs are shown in Fig. 3 (the sputtering rate was decreased by a factor of 7). In this case the effect under study almost disappears.

The effects we have observed should also appear when secondary-ion mass spectroscopy is used. However, in this case they are superimposed on changes in the analysis signal connected with changes in the coefficient of ionization of the sputtered particles (which can be considerably larger), but it is possible to take our effect into account in this case as well.

In order to observe artifacts of a similar kind we can use measurements at various values of j_0.

[1] A. W. Czanderna (ed.), *Methods of Surface Analysis* (Elsevier, New York, 1975) [Russian trans., Mir, Moscow, 1979].
[2] A. A. Dorozhkin, A. P. Kovarsky, Y. Kudriavtsev, and M. Yagodkina, *Abstracts of SIMS IX*, Yokohama, 1993.
[3] A. Benninghoven, F. G. Rudenauer, and W. Werner, *Secondary Ion Mass Spectrometry* (Wiley, New York, 1987).

Translated by F. Crowne

Ultrasonic monitoring of the accumulation of aging damage and recovery of the useful lifetime of industrial parts

L. B. Zuev, V. Ya. Tsellermaer, V. E. Gromov, and V. V. Murav'ev

Institute of Strength Physics and Materials Science, Siberian Departmant of the the Russian Academy of Sciences, 634021 Tomsk, Russia

A method is described for monitoring changes in the strength properties of steels during fatigue testing based on measuring the velocity of propagation of ultrasound. It is shown to be possible to detect the dangerous stage of development of fatigue damage for individual industrial parts. A method is also proposed for restoring the function of manufactured items by treating them with high-power electric current pulses. It is shown that such treatment significantly increases the useful lifetime and can prevent fatigue failure. © *1997 American Institute of Physics.* [S1063-7842(97)02709-8]

INTRODUCTION

It is well known that the prediction of useful lifetimes of industrial parts subjected to fatigue-inducing loading is a very complicated problem. Data on fatigue limits obtained by plotting the so-called Weller curve[1] can only be used to estimate average characteristics of a material, and never give information about such important reliability indicators of the individual components as the useful lifetime.[2] This methodology significantly limits progress in increasing the reliability of machines and components. Moreover, fatigue-induced failure is usually sudden in character, and its approach is not accompanied by any noticeable external signs. Microscopic studies[1,3] show that during fatigue a gradual accumulation of microdamage takes place, followed by a slow hidden rise in fatigue-induced cracks, leading to catastrophic growth of a primary destructive macrocrack. The existence of this long-term preparatory stage suggests that suitable intervention can help to prevent the transition to the next stage of the process.

Thus there are two problems to address in application studies of the process of metal fatigue and fatigue in alloys: the choice and justification of provable and sufficiently informative signs of arrival at a dangerous stage of structural change, leading to failure; and development of methods that can be directly used on machine parts or components to delay the development of defects whose growth within a short time can lead to failure. Several approaches to this problem will be discussed in this paper.

ON THE POSSIBILITY OF DIAGNOSING FATIGUE-INDUCED FAILURE

A very convenient quantity that carries reliable information on the structure of a material and its changes is the velocity of propagation of ultrasound in metals and alloys.[4] The fact that this quantity is determined by elastic modulus (G for transverse waves) and the material density ρ, i.e., $V_t = (G/\rho)^{1/2}$, while correct, does not fully reflect the complexity of the problem.[5] It has been discovered[4] that practically all structural changes caused by heat treatment or strain lead to small but measurable changes in the ultrasonic velocity. Therefore one promising approach to structural monitor-

ing is to use the measured ultrasonic velocity to diagnose a material under fatigue-induced loading. Figure 1 illustrates changes in the ultrasound velocity during fatigue testing of samples made of 45-steel; these changes are inferred from measurements of bending oscillations. Analogous dependences were also obtained for samples made of M76 rail steel by the method of autocirculation of sound pulses at a carrier frequency 2.5 MHz and the use of an ISP-12 apparatus.[4] These data indicated that the form of the dependence of the velocity of transverse ultrasonic waves on the number of loading cycles n was qualitatively the same for all the samples tested. In all cases, $V_t(n)$ exhibited three successive drops in the ultrasonic wave velocity, although the size and rate of the quantitative change differed from sample to sample (Fig. 1). Other measurement methods have also revealed a characteristic three-stage kinetics for the change of several metal properties during fatigue testing. For example, in Ref. 6 it was shown that the Young's modulus of steel changes analogously (see, also, Ref. 7); in Ref. 8 similar behavior was observed in the electrical resistance; and in Refs. 9,10 the authors noted similar patterns in the buckling amplitudes of samples during bending oscillations. Under-
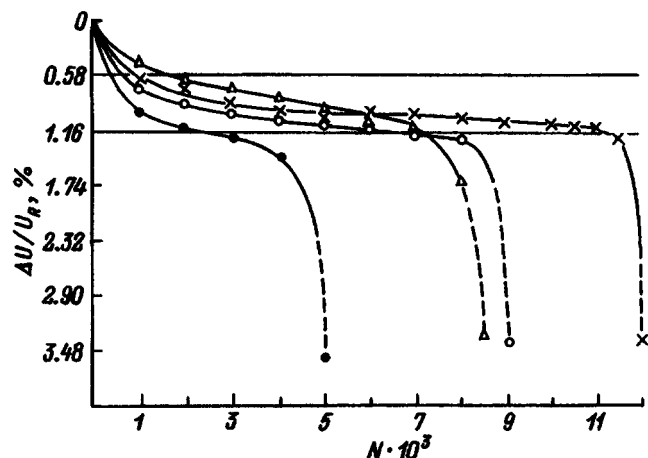

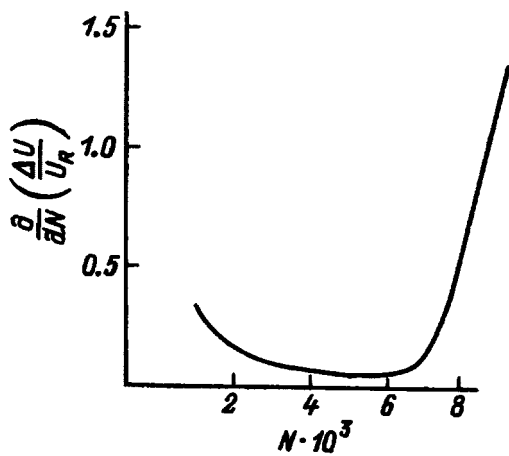FIG. 1. Relative change of the ultrasound velocity during fatigue testing of samples of 45-steel.

FIG. 2. Change in ultrasound velocity from its value at test cycle No. 1 versus the number of cycles.

standing of the essence of the processes responsible for such step-like curves is achieved through microstructural analysis of the metal undergoing fatigue testing. It turns out that while no particular change is detectable in the microstructure at stages 1 and 2, as stage 3 is approached traces of plastic deformation appear. Immediately after the function $V_t(n)$ begins its steep decrease one observes signs of destruction in the form of microcracks with dimensions $\geqslant 0.01$ mm.

Thus there is reliable evidence that the transition to a stage-3 dependence of the ultrasound velocity on the number of loading cycles signals the approach of the catastrophic stage of fatigue and the exhaustion of the working life of the component. Since measurements of the ultrasonic velocity using the ISP-12 apparatus or something similar to it[4] are quite simple, they can be implemented in practice without hindering the operation of structures and machines that are in use and subjected to sign-changing loading. This makes it possible to detect signs that the parts material has reached a critical state beforehand while the parts are in use.

We also argue that this type of dependence of $V_t(n)$ is useful in explaining the U-shaped character of the dependence of the intensity of outages (i.e., their number per unit time) on the time of use t, which is well known in the theory of reliability.[11] Clearly, the plot of dV_t/dn versus n shown in Fig. 2 (which is based on the data of Fig. 1) is also U-shaped (obviously $n \sim t$). In terms of reliability theory[2,11] its initial stage corresponds to outages during operation, the stage of slow falloff of the ultrasound velocity is the period of sudden outages (normal use), and the rapid falloff in the ultrasonic velocity signals a transition to the most dangerous period of wear-induced outages. In this stage, serious changes have already taken place in the structure of the material, in particular, as microscopic investigations show, the appearance of cracks of dangerous size.

RESTORING THE LIFETIME OF COMPONENTS AFTER FATIGUE TESTING

If timely observation of the onset of a critical stage of fatigue is feasible, then it also should be possible to restore the working lifetime of an industrial part by healing the mi-

crocracks by exposing the material to some specific external factors,[12] since there is still time to use them. The most promising approach in this context is to use high-power electric current pulses. For example, the authors of Refs. 13–15 are of the opinion that the plastification of alloys via the electroplastic effect takes place because microcracks are healed by the passage of a train of electric current pulses through the alloy with current density $j \geqslant 10^3$ MA/m^2. In the course of this work, we attempted to apply this effect to steel samples whose state after fatigue testing corresponded to the beginning of stage 3 for the function $V_t(n)$, which was identified based on measuring the ultrasound velocity.

The scheme of the experiment was as follows. Samples made of M76 rail steel were subjected to fatigue testing according to a scheme from a zero loading cycle with amplitude 1.6 MPa, with simultaneous measurement of the ultrasonic velocity using the ISP-12 apparatus. For the material in its initial state the value of the ultrasonic velocity was 2910 m/s. Samples under these conditions were destroyed after ~ 3150 loading cycles (averaged over eight samples), whereas the beginning of a significant drop in the ultrasound velocity (to a value of 2890 m/s) was noted after only 300 loading cycles. Samples that had reached this state were subjected to treatment by current pulses with frequency 20 Hz and amplitude ~ 20 kA for 30 minutes. Longitudinal fatigue testing showed that this treatment not only practically restored the original value of the ultrasound velocity, which indicates the return to the initial material structure, but also allowed the samples to undergo an additional 1000–1500 loading cycles before failure, i.e., the effect was quite significant.

Thus, electrical treatment of materials after they have reached the critical stage of accumulation of fatigue damage can appreciably increase the working lifetime of samples, When real components are used, this procedure can prolong their useful lifetime. We will assume that the most probably reason for this increase in working lifetime is healing of nucleating microcracks that form during fatigue loading, due to local heating of the material in the region of current line concentrations at the ends of the cracks, the blunting of the latter due to relaxation of stresses and the corresponding decrease in the level of stress concentration in these zones. Effects of this kind are widely discussed in the literature[12–14] for unrelated problems in fatigue.

Note that a large role is played in this case by the regime of electrical treatment. Thus, the authors of Ref. 16 made analogous attempts to increase the fatigue strength by preliminary treatment of steel with electrical pulses. However, this did not lead to an increase in the working capability of the fatigue samples, which contained prior cracks that had grown to various lengths (1.5–2.0 mm); rather, it only decreased the scatter in the data from one sample to another. Possibly this is connected with the shape of the electrical pulses used, which were rapidly attenuating sinusoids, a shape that is characteristic of the discharge from a capacitor into a low-inductance circuit. Contrast this with the well known fact[15,17] that the largest electroplastic effect is observed in alloys subjected to unipolar pulses from a special oscillator.

Thus when *in situ* monitoring of the state of a metal during fatigue testing by measuring its ultrasound velocity is combined with treatment by electrical pulses, it becomes feasible in principle to restore the working life of industrial parts operating under conditions of fatigue loading, thereby contributing to the increased reliability of machines and mechanisms.

[1] A. J. Kennedy, *Processes of Creep and Fatigue in Metals* (Oliver and Boyd, Edinburg, 1963; Wiley, New York, 1963) [Russian trans., Metallurgiya, Moscow, 1965].

[2] D. N. Reshetov, A. S. Ivanov, and V. Z. Fadeev, *Reliability of Machines* [in Russian], Vysshaya Shkola, Moscow (1988).

[3] V. M. Goritskiĭ and V. F. Terent'ev, *Structure and Fatigue-Induced Failure of Metals* (Metallurgiya, Moscow, 1980).

[4] V. V. Murav'ev, L. B. Zuev, and K. L. Komarev, *Velocity of Sound and Structure of Steels and Alloys* [in Russian], Nauka, Novosibirsk (1996).

[5] T. Tokuoka and Yu. Iwashimizu, Int. J. Solids Structures **4**, 383 (1968).

[6] A. Karius, E. Gerold, and E. H. Shulz, Archiv für Eisenhüttenwesen, **5/6**, 113 (1944).

[7] V. S. Ivanov, *Metal Fatigue* [in Russian], Metallurgiya, Moscow (1963).

[8] V. Ya. Bash, *Study of Stress and Strain by the Thermoelectric Method* [in Russian], Naukova Dumka, Kiev (1984).

[9] L. B. Zuev, N. K. Korotkikh, and V. M. Muratov, Izv. Vyshch. Uchebn. Zaved. Chernaya Metallurgiya, No.10, p. 81 (1980).

[10] L. B. Zuev and N. K. Korotkikh, Izv. Vyshch. Uchebn. Zaved Chernaya Metallurgiya, No.10, p. 80 (1984).

[11] B. Dilon and Ch. Singh, *Engineering Reliability: New Techniques and Applications* (Wiley, New York, 1981) [Russian trans., Mir, Moscow, 1984].

[12] V. M. Finkel', *Physical Priciples of Forestalling Damage* [in Russian], Metallurgiya, Moscow (1977).

[13] V. M. Finkel', Yu. I. Golovin, and A. A. Sletkov, Dok. Akad. Nauk SSSR **227**, 848 (1976) [Sov. Phys. Dokl. **21**, 216 (1976)].

[14] K. M. Klimov and I. I. Novikov, Dok. Akad. Nauk SSSR **260**, 1360 (1981) [Sov. Phys. Dokl. **26**, 722 (1981)].

[15] A. F. Sprecher, S. L. Mannan, and H. Conrad, Acta Metall. **34**, 1145 (1986).

[16] G. V. Stepanov and A. I. Babushkin, Problemy Prochnosti, No. 5/6, pp. 74–78 (1995).

[17] V. E. Gromov, V. I. Stashenko, and O. A. Troitskiĭ, Izv. Akad. Nauk. SSSR Metally, No. 2, 154 (1991).

Translated by F. Crowne

Positive temperature coefficient of resistance in the lead-containing ceramic barium titanate

A. V. Myasoedov and S. R. Syrtsov

Institute of Acoustical Engineering, Belorussian Academy of Sciences, 219923 Vitebsk, Belarus

Recent years have seen a continuing interest in the synthesis and development of ceramic materials with a high temperature jump in the resistance [positive temperature coefficient of resistance (PTCR)] due to the wide use of such materials in various electron devices.[1,2] One of the most promising materials of this group is the solid solution $(Ba,Pb)TiO_3$ doped by a small quantity of rare-earth elements. The function of such dopants is to endow the ceramic with semiconductor properties. Among such dopants, yttrium has gained widest use thanks to the fact that its presence results in a wider minimum (0.2–0.5 at. %) in the resistivity ρ_0 in comparison with the other dopant elements. At the same time, the effect of the preparation process parameters (in particular, annealing temperature and time) on the characteristics of the PTCR of such systems has received insufficient study.

The present paper examines the effect of the annealing temperature and time on the characteristics of the PTCR of 0.5 at. % yttrium-doped $Ba_{0.9}Pb_{0.1}TiO_3$.

As the initial reagents we used OSCh-grade barium titanate, lead oxide, and titanium oxide and ChCh-grade yttrium oxide Y_2O_3. The time of wet grinding and mixing of components was 24 h. Thermal synthesis of the ceramic took place in two steps. The first step consisted of preliminary annealing of the material at 950 °C for 2 h followed by grinding for 4 h in water. Pellets with the dimensions $20\times20\times2$ mm were then pressed and subjected to high-temperature annealing. The anneal regime was the following. The samples were kept at 600 °C for 2 h and then heated at the rate of 6 °C/min to the required annealing temperature and kept at that temperature for a prescribed time. They were then cooled at the rate of 50 °C/min to 1135 °C, and kept at that temperature for 30 min, with subsequent cooling to room temperature at a rate of 5 °C/min.

The temperature dependence of the resistivity was determined by the two-probe method using electrodes of indium–gallium in the ratio 40:60. The voltage on the samples did not exceed 1 V.

Results of our study of the effect of the annealing temperature T_0 and time t_0 on the characteristics of the PTCR are shown in Figs. 1 and 2. It can be seen that as t_0 is increased from 3 to 30 min a steep decrease is observed in the room-temperature resistance of the ceramic and the resistance jump $\gamma = R_{max}/R_{min}$ also decreases (especially strongly in the region of small annealing times). The magnitude of the resistance jump depends substantially on the annealing temperature, takes its maximum value at $T_0 = 1350$ °C while ρ_0 depends weakly on T_0.

Note that the resistance temperature coefficient[2] $\alpha = (2.3\Delta \lg R)/(T_2 - T_1)$ depends weakly on t_0 and T_0 (and is approximately equal to 20% 1/K) due to an increase in the temperature interval of the transition $T_2 - T_1$.

The shift of the transition temperature (for pure $BaTiO_3$ the Curie temperature $T_C \simeq 120$ °C) was somewhat less than the values given in the literature. This is due, apparently, to evaporation of part of the volatile PbO during sintering of the ceramic in air.

On the basis of the Heywang model,[2] which links the anomalous behavior of the resistance with the presence of acceptor states on the grain boundaries, it may be supposed that the above-established regularities in the PTCR parameters are due to a difference in the rate of formation of acceptor and donor states respectively on the boundaries and inside the grains of the ceramic.
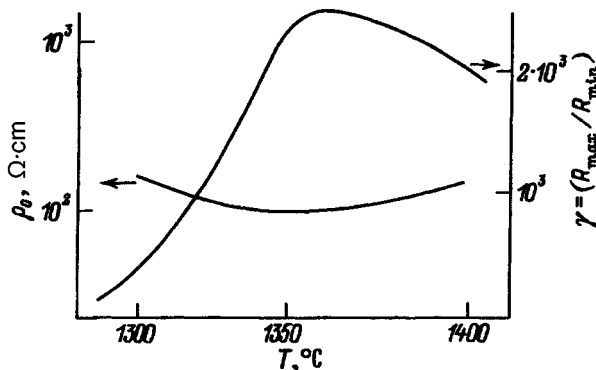

FIG. 1. Dependence of the room-temperature resistivity ρ_0 and resistance jump $\gamma = R_{max}/R_{min}$ on the annealing temperature ($t_0 = 30$ min).
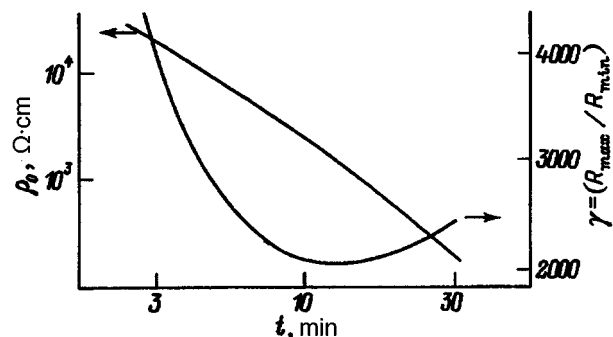

FIG. 2. Dependence of the room-temperature resistivity ρ_0 and resistance jump $\gamma = R_{max}/R_{min}$ on the annealing time ($T_0 = 1350$ °C).

[1] E. D. Politova and V. G. Elenskiĭ, Zarubezh. Radioélektron., No. 10, 66–75 (1990).

[2] *Barium-Titanate Based Semiconductors* [Russian translation from Japanese] (Énergoizdat, Moscow, 1982).

Use of hysteresis in bifurcation systems to measure noise

O. Ya. Butkovskiĭ, Yu. A. Kravtsov, and E. D. Surovyatkina

Space Research Institute, Russian Academy of Sciences, 117810 Moscow, Russia

It is proposed to make use of the influence of noise on the delay in dynamic bifurcations to measure the level of weak intrinsic noise in nonlinear chaotic systems. The dependence of the size of the hysteresis loop on the rate of change of the control parameter and on the noise level in a system described by a logistic map is found. A calibration curve for determining the noise level from measured values of the size of the hysteresis loop is obtained. © *1997 American Institute of Physics.* [S1063-7842(97)02909-7]

INTRODUCTION

The phenomenon of delayed loss of stability in bifurcation systems has been known for quite a long time.[1] It has been analyzed in detail in many papers (see Refs.2–8 and the references therein). As is well known, when the system passes through a bifurcation point the number of stable periodic points of the map is doubled while the previously stable point becomes unstable. The phenomenon of delay is manifested in the circumstance that in dynamical bifurcations, after passing through a bifurcation point the system remains for quite a long time on an unstable branch and after some time makes a comparatively rapid transition to one of two stable points (for simplicity we consider the case of the first period-doubling bifurcation). A similar process takes place for the reverse transition, i.e., when the control parameter is decreased. The smaller the control parameter, the more pronounced is the phenomenon of delay.

When the control parameter is swept cyclically, the system is delayed in the vicinity of the preceding stable points both in the forward and in the reverse direction. In this situation the phenomenon of delay leads to the appearance of a hysteresis loop.

It is clear that in the presence of noise the time the system spends on an unstable branch is shortened, since the action of noise hastens the deviation from the unstable position. This phenomenon is reminiscent of the acceleration of phase transitions in the presence of noise. As is well known, at low noise levels matter has a tendency to be hung up around unstable (metastable) states, a phenomenon which is similar to the delay in bifurcation systems. At the same time, additional noise shortens the residence time on an unstable branch in a way similar to how noise shrinks the dimensions of the hysteresis loop in bifurcation systems.

Attempts to describe the phenomenon of delay under the action of noise were made in Refs. 9–13, which considered not only qualitative but also some quantitative characteristics of the phenomenon. In the present paper we propose to use the dependence of the delay time on the external noise level to measure weak internal noise in the system.

PECULIARITIES OF HYSTERESIS IN DYNAMIC BIFURCATIONS IN THE PRESENCE OF NOISE

We will consider characteristics of ''noise'' hysteresis, i.e., hysteresis in the presence of noise, for a model dynamical system described by a nonstationary noisy logistic map:

$$x_{n+1} = r_n x_n (1 - x_n) + f_n, \quad r_{n+1} = r_n + s, \tag{1}$$

where r_n is the time-varying control parameter, s is the rate at of change of the control parameter, and f_n is a random disturbance.

For simplicity we consider only the situation in which the random variable f_n is uniformly distributed over the interval $(-\gamma, +\gamma)$, so that the variance σ_f^2 is equal to $\gamma^2/3$. The logistic model, complicated by noise and nonstationarity, was already used to analyze bifurcations in Refs. 7 and 10. In the quasistationary regime the first period-doubling bifurcation for the logistic map takes place at the critical value $r_{c1} = 3$. Let us investigate the vicinity of this point by varying the control parameter from $r_0 = 2.8$ to $r_1 = 3.2$ (the forward direction) and from 3.2 to 2.8 (the reverse direction; in the reverse direction the quantity s in Eq. (1) changes sign). Results of a numerical calculation of the dependence of the dynamical variable x_n on the control parameter r_n are plotted in Fig. 1 for the case of very small noise $\sigma_f^2 = 10^{-14}$. For the rate of change s equal to 0.0004, the forward bifurcation takes place at $r_+ = 3.08$, which exceeds the critical value $r_{c1} = 3$ whereas in the reverse direction the return to the initial stable state takes place at $r_- = 2.95$, which is less than the critical value.

After passing through the critical value $r = r_{c1} = 3$ the stable branch *A* in Fig. 1 splits into two stable branches *1* and *2* and an unstable branch *3*. As a consequence of delay of the loss of stability the system can be found for a long time in the vicinity of the unstable branch, after which it transitions quite rapidly to one of the two stable branches *1* and *2* (Fig. 1). The residence time of the system on the unstable branch *3* is limited by the action of noise. With growth of the noise level σ_f^2 the system moves away from the vicinity of the unstable branch *3* and transitions to one of the two stable branches *1* and *2*. Note that phenomena to those described above also take place at the second, third, etc. period-doubling bifurcations.[13,14]

Results of a numerical analysis of the effect of noise on the bifurcation process are shown in Fig. 2. According to Fig. 2, the larger the noise level, the faster the system transitions to one of the stable branches. With growth of the noise level the time of the actual bifurcation approaches its quasistationary value $r_{c1} = 3$. Thus, noise accelerates the bifurcation process and shrinks the size of the hysteresis loop.

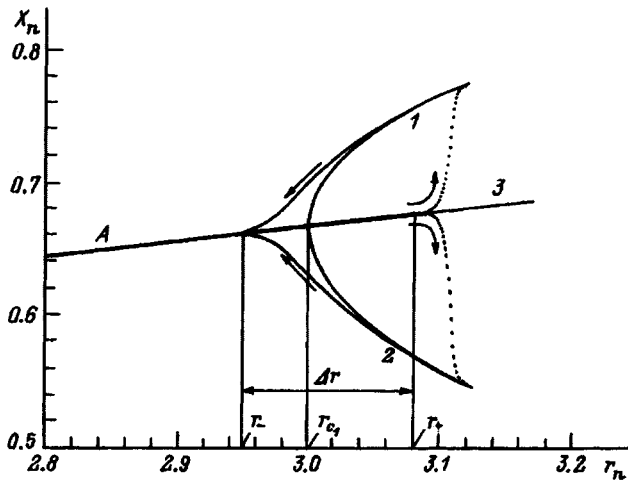The size of the hysteresis loop Δr may be characterized

FIG. 1. ''Noise'' hysteresis loop ($s = 0.0004$, $\sigma_f^2 = 10^{-14}$) in a dynamical system described by a logistic map.

PROBABILITY DISTRIBUTION OF THE DYNAMICAL VARIABLE FOR A FAST BIFURCATIONAL TRANSITION

To analyze the evolution of the probability distribution of the dynamical variable, we calculated the distributions of this quantity for several values of the control parameter $r = 2.9$, 2.95, 3.00, 3.05, 3.10, and 3.20. Each histogram was calculated from 500 time series. Figure 3 presents a set of histograms corresponding to rate of variation of the control parameter $s = 0.0004$ and noise level $\sigma_f^2 = 10^{-8}$. Bifurcation in the histograms is manifested first in the form of a split peak which then rapidly separates into two independent peaks corresponding to the two stable bifurcation branches 1 and 2. Splitting occurs at $r_{c1} \approx 3.04$. Thus, splitting takes place at a value of the control parameter that exceeds the stationary bifurcation value $r_{c1} = 3.00$. As the system is made to pass through the bifurcation value $r = r_{c1} = 3.00$ we do not detect any splitting of the histograms corresponding to ''pre-bifurcation noise''[17] (the analog of ''pre-oscillation'' fluctuations). An insignificant amplification of the fluctuations was observed only in the vicinity of the actual splitting points of the trajectories.

by the position of the break point r_+ of the trajectory along branch 3 for the forward direction and the position of the return point r_- for the reverse direction (Fig. 1). It is natural to define the points r_+ and r_- by the condition that the trajectory distance itself from the unstable branch 3 (or approach the stable branch A) by a certain distance, say on the order of 1% from its stationary value. With growth of the noise level σ_f^2 the upper limit of the loop r_+ approaches the critical value r_{c1} while the lower limit r_- is practically independent of the noise. This circumstance was not noted in previous publications.

The regularities of ''noise'' hysteresis revealed by the above analysis of a noisy nonstationary logistic map are also manifested in other bifurcation systems, e.g., in polarization bifurcations in nonlinear optics.[15,16]
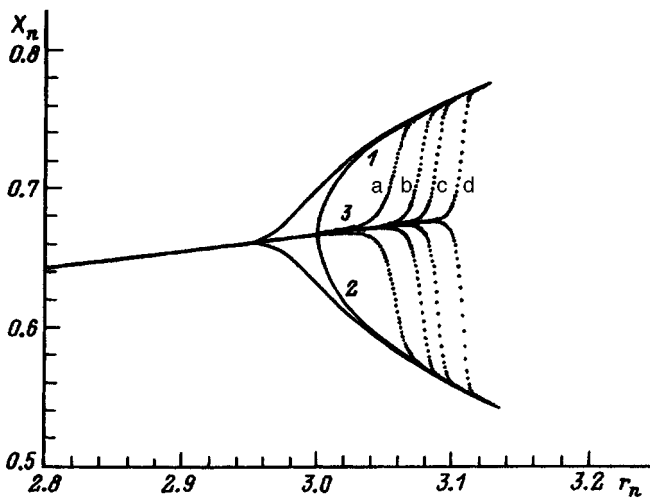
POSSIBILITY OF MEASURING WEAK INTRINSIC NOISE IN BIFURCATIONAL DYNAMICAL SYSTEMS

The phenomenon of hysteresis in bifurcational transitions may be useful in measuring the level of weak internal noise, which is usually a complicated experimental task.

The point is that in the case of weak noise the fluctuations $\xi = x - \bar{x}$ about the equilibrium position $\bar{x}$ are quite small. Proceeding from the linearized equations for ξ (Refs. 7 and 9), we note that the variance of the fluctuations $\sigma_\xi^2 = \langle \xi^2 \rangle$ is equal to

$$\sigma_\xi^2 = \frac{\sigma_f^2}{1 - A^2}, \tag{2}$$

where $A = \partial F / \partial x$ is the ''slope'' of the map at the stationary point.

FIG. 2. ''Noise'' hysteresis in a logistic map with rate of change of the control parameter $s = 0.0004$ for various values of the noise level $\sigma_f^2 = 10^{-6}$ (curve a), 10^{-8} (curve b), 10^{-10} (curve c), 10^{-14} (curve d).
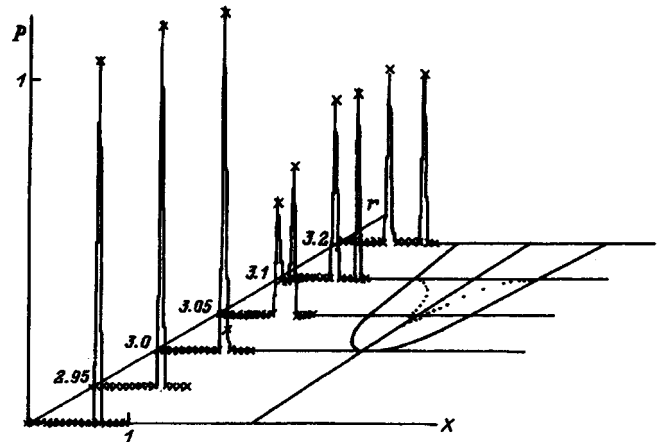

FIG. 3. Histograms characterizing the probability distribution of the dynamical variable x with growth of the control parameter with rate of change $s = 0.0004$ at the noise level $\sigma_f^2 = 10^{-6}$.
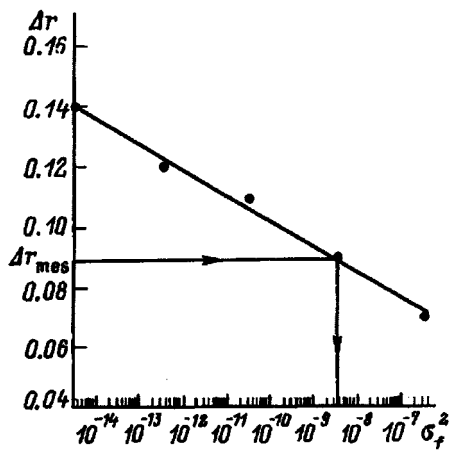
FIG. 4. Dependence of the size of the noise hysteresis loop $\Delta r = r_+ - r_-$ on the noise level. This dependence allows one to estimate the noise level from measurements of the size of the hysteresis loop.

For the logistic map, $A = 2 - r$. Far from the bifurcation point the variance σ_ξ^2 is comparable with σ_f^2 by virtue of formula (2), which complicates the measurements.

As the bifurcation point is approached, the quantity $|A|$ tends to unity. At $r = r_{c1}$ the quantity $|A| = |2 - r|$ is equal to unity and the variance σ_ξ^2, calculated according to the linearized theory [formula (2)], diverges. Indeed, as $r \to r_{c1}$ the linearized theory loses force, but a rough account of the nonlinear terms shows that the estimate of σ_ξ^2 in the limit $r \to r_{c1}$ saturates at the level $\sigma_\xi^2 = \sigma_f^2/G$, which "characterizes the growth of fluctuations"[17] (here $G \approx 3$). For the noise variance less than 10^{-8} the magnitude of the pre-bifurcation fluctuations σ_ξ^2 turns out to be quite small: $\sigma_\xi^2 < 0.4 \times 10^{-4}$.

Under these conditions the dynamic method of determining the noise level from characteristics of the hysteresis loop provides much greater sensitivity since the size of the loop is very sensitive to weak noise.

By way of an example, Fig. 4 plots the dependence of the size of the hysteresis loop $\Delta r = r_+ - r_-$ on the noise level σ_f^2 for the logistic map. As can be seen from this figure, the size of the loop Δr decreases with growth of the noise σ_f^2.

In order to estimate the noise level σ_f^2, the idea is to measure the length of the hysteresis loop $\Delta r = r_+ - r_-$ and determine the variance of the internal noise of the system from the calibration graph $\Delta r = F(\sigma_f^2)$. This method allows one to estimate very low noise levels σ_f^2 all the way down to $10^{-12} - 10^{-14}$. In our numerical experiments we were not able to realize lower noise levels due to the effect of round-off error, but this limitation is not present in real physical systems so that it should be possible to measure very low noise levels.

It may be expected that in other dynamical systems the dependence of the size of the loop Δr on the noise level will have a behavior qualitatively similar to that just described. Of course, the calibration curve $\Delta r = F(\sigma_f^2)$ will be different for different systems.

It is possible to recommend another method for determining σ_f^2, based on measuring the residence time τ of the system in a narrow vicinity of the unstable branch 3. This time depends substantially on the noise level so that by measuring τ it should be possible to estimate the noise variance σ_f^2 with the help of calibration curves similar to the one in Fig. 4.

[1] M. A. Shishkova, Dokl. Akad. Nauk SSSR **209**, 576 (1973).
[2] A. U. Neĭshtadt, Usp. Mat. Nauk **41**, 300 (1985).
[3] A. U. Neĭshtadt and V. V. Sidorenko, Preprint of the Institute of Applied Mathematics, Russian Acad. Sci., No. 56, p. 28 (1995).
[4] B. Morris and F. Moss, Phys. Lett. A **118**, 117 (1986).
[5] P. Mandel and T. Erneux, Phys. Rev. Lett. **53**, 1818 (1984).
[6] M. Lefebre, D. Dangoisse, and P. Glorieux, Phys. Rev. A **29**, 1486 (1984).
[7] C. Baesens, Physica **53**, 319 (1991).
[8] E. Benoit (ed.), *Dynamical Bifurcations. Lecture Notes in Mathematics* (Springer-Verlag, Berlin, 1993).
[9] P. Pieranski and J. Malecki, Nuovo Cimento D **9**, 757 (1987).
[10] C. Van den Broeck and P. Mandel, Phys. Lett. A **122**, 36 (1987).
[11] H. Zeghlache, P. Mandel, and C. Van den Broeck, Phys. Rev. A **40**, 286 (1989).
[12] O. L. Anosov, O. Ya. Butkovskii, Yu. A. Kravtsov, and E. D. Surovyatkina, in *Proceedings of the Third Technical Conference on Nonlinear Dynamics (CHAOS) and Full Spectrum Processing*, Connecticut, USA, 1995.
[13] O. Ya. Butkovskii, J. S. Brush, and Yu. A. Kravtsov, *Predictability of Complex Dynamical Systems* (Springer-Verlag, Berlin, 1995).
[14] R. Kapral and P. Mandel, Phys. Rev. A **32**, 1076 (1985).
[15] E. Arimondo, D. Dangoisse, C. Grabbanini, E. Menchi, and F. Papoff, J. Opt. Soc. Am. B **4**, 892 (1987).
[16] I. N. Zheludev, Usp. Fiz. Nauk **157**, No. 4, 683 (1989) [Sov. Phys. Usp. **32**, 357 (1989)].
[17] W. Horsthemke and R. Lefever, *Noise-Induced Transitions*, (Vol. 15 of Springer Series in Synergetics, Springer, New York, 1983).

Translated by Paul F. Schippnick

Application of the velocity spectrum to a spatiotemporal study of high-speed processes

A. A. Aliverdiev

Dagestan State University, 367025 Makhachkala, Dagestan

The dynamical problem of reconstructing an internal structure from a set of integrals depending on time and the velocity of the signals being integrated is solved. The results can be used to study dynamic objects of various kinds. © *1997 American Institute of Physics.*
[S1063-7842(97)03009-2]

The use of computer tomography in velocity space has been the subject of much discussion in the literature. For example, the authors of Ref. 1 proposed to use it to enhance the quality of photographic recording of fast processes while the authors of Refs. 2 and 3 proposed using it to determine the three-dimensional velocity distribution of particles in a gas or plasma. However, the problem of using the velocity spectrum of recorded signals to examine the internal structure of a dynamic object has still not been solved. The aim of the present article is to develop the mathematical part of this problem.

Thus let a set of n-tuple integrals over the region Ω be known (these integrals are recorded by sensors located on an n-dimensional sphere of radius r enclosing this region) for a set of times t and a set of velocities v. Then the mathematical problem can be formulated as follows. Knowing

$$G(\alpha,t,v) = \int_{\Omega} f(x,t+a/v)dx, \tag{1}$$

it is possible to find $f(x,t)$. Here x is the n-dimensional position vector of the object of interest (in practice n can take the values 1, 2, 3); α is the n-dimensional unit vector characterizing the orientation of the sensor (obviously, for $n=1$ α is a scalar quantity which can take the values 1 and -1); and a is the distance from the point x to the sensor α.

To solve the problem, we set $g(\vartheta,s)=G(\alpha,t-1/v,v)$, where $\vartheta=(\alpha\cos(\psi),\sin(\psi))$ is a $(n+1)$-dimensional vector whose first n coordinates are equal to the coordinates of the vector $\alpha\cos(\psi)$, and the last coordinate is equal to $\sin(\psi)$; $s=c(t-\tau-1/v)\sin(\psi)$; $\psi=\cot^{-1}(v/c)+\pi/2$ (as can be seen $\pi/2\leq\psi\leq\pi$, but this interval is sufficient since the time cannot take negative values); c and τ are constants of the new coordinate system $z=(x,c(t-\tau))$, chosen for the most convenient representation in it of the concrete object.

Thus formula (1) can be represented in the form

$$g(\vartheta,s) = \int_{\Gamma(s,\vartheta)} f(z)dz. \tag{2}$$

Unfortunately, for $n>1$ the region $\Gamma(s,\vartheta)$ is nonlinear, which substantially complicates the solution. Let us consider the particular case $r\gg r_0$, where r_0 is the radius of the minimal sphere enclosing Ω, itself enclosed within the concentric sphere of sensors. In this case we can set $a=r-x\alpha$, but the region $\Gamma(s,\vartheta)$ is a hyperplane perpendicular to ϑ and lo-

cated a distance s from the origin. Then the problem of determining f from g is the classical Radon problem. Consequently, we can set

$$f(x,t) = f(z) = \mathbf{R}^{-1}g(\vartheta,s), \tag{3}$$

where $\mathbf{R}^{-1}$ is the inverse Radon transform.

At present there are a number of ways of solving the Radon problem, including the case of incomplete data.[1,3,4] Thus formula (3) can be taken as the solution of the problem for $r\gg r_0$ (for $n=1$ without this additional condition).

In the practical realization of the proposed method, due to the fundamental impossibility of having a complete set of velocities the use of standard reconstruction algorithms is impossible in the majority of cases. However, at present an extensive theory of the solution of such problems which makes use of *apriori* information about the object has been developed, which gives reason to hope for success even for an exceedingly small number of angular (in the given case ''velocity'') projections.

Let us consider two concrete examples for a one-dimensional region of spatial integration. Let the function f be the characteristic function of some set ϕ in the space-time plane $z=(x,c(t-\tau))$. We also assume that any ray emanating from the origin intersects the region ϕ precisely once (here we require the optimal set τ). The boundary of the region ϕ in polar coordinates (ρ,φ) is given by the functional relation $\rho=F(\varphi)$. The function F determines the sought-after function f. The Fourier transform of f has the form

$$\begin{aligned}
\hat{f}(\sigma\vartheta) &= \frac{1}{2\pi}\int_{\phi}\exp(-i\sigma\vartheta z)dz \\
&= \frac{1}{2\pi}\int_{0}^{2\pi}\int_{0}^{F(\varphi)}\rho\,\exp(-i\sigma\vartheta\omega)d\rho d\varphi \\
&= \frac{1}{2\pi}\int_{0}^{2\pi}F^2(\varphi)K(\sigma\vartheta\omega F(\varphi))d\varphi,
\end{aligned}$$

where

$$K(u)=\begin{cases} u^{-2}((1+iu)\exp(-iu)-1), & u\neq 0, \\ 1/2, & u=0, \end{cases}$$

$\omega=(\cos(\varphi),\sin(\varphi))$, σ is the spatial frequency and i is the imaginary unit.

It follows from the projection theorem[4] that $\hat{g}(\vartheta,\sigma)=(2\pi)^{1/2}\hat{f}(\sigma\vartheta)$. Hence we have

$$\hat{g}(\vartheta,\sigma)=(2\pi)^{-1/2}\int_0^{2\pi}F^2(\varphi)K(\sigma\vartheta\omega F(\varphi))d\varphi. \tag{4}$$

If the function g is prescribed for p directions, then Eq. (4) generates a system of nonlinear integral equations of the first kind in F (in this system $\vartheta=\vartheta_j$, $j=0,1,\ldots,p$). In Refs. 5 and 6 it was shown that if the function f satisfies the requirements stated above, then for $p=4$ there exists a unique solution $f=\mathbf{R}^{-1}g$, while Ref. 4 presents the results of the solution of a system similar to (3) by the Tikhonov–Phillips method for $p=3$. All this allows us to hope to reconstruct the function f using one high-speed and one low-speed signal, both taken from two sides. This in turn makes it possible to study high-speed processes accompanied by an optical and an acoustic emission or two acoustic (longitudinal and transverse) emissions.

Let us consider one more example. Let the desired function be given in the form of a product $f(x,t)=X(x)T(t)$. We will show that just two projections may be enough to reconstruct it:

$$G_1(t)=\int_{-r}^r X(x)T(t+(r-x)/v_1)dx \tag{5}$$

and

$$G_2(t)=\int_{-r}^r X(x)T(t+(r-x)/v_2)dx, \tag{6}$$

one of which is taken for a very high-speed signal ($v_1\to\infty$). Indeed, as $v_1\to\infty$ formula (5) tends to

$$G_1(t)=\int_{-r}^r X(x)T(t)dx=\left[\int_{-r}^r X(x)dx\right]T(t). \tag{7}$$

Calculating the function $T(t)$ from Eq. (7) and substituting it in Eq. (6), we obtain

$$G_2(t)=\int_{-r}^r X_n(x)G_1(t+(r-x)/v_2)dx, \tag{8}$$

where

$$X_n(x)=X(x)\left[\int_{-r}^r X(x)dx\right]^{-1}. \tag{9}$$

Equation (8) is the Fredholm equation of the first kind in $X_n(x)$, which can be solved by standard methods.[7] Finally we obtain $f(x,t)=X_n(x)G_1(t)$ (although it is most probably $X_n(x)$ that contains exhaustive information about the object).

Obviously, the high-speed signal in this case may be initiating as well as accompanying. Thus the proposed method may be used to investigate the structure of nonlinear crystals by analyzing secondary acoustic radiation initiated by laser pumping.

The case is also interesting in which the function g may be represented in the form of rarefied stochastic fluxes of correlated quanta. Then, to reconstruct the desired function one may use just one signal taken from two sides. This case is considered in Refs. 8 and 9. In the presence of substantial dispersion of the signal being integrated, the possibility is not excluded of using more standard solutions of the Radon problem.

In conclusion we note that for $n>1$ it is also necessary to construct reconstruction algorithms for a severely limited number of projections since such algorithms can find use in the n-dimensional study of processes accompanied by radiation that cannot be treated in the geometrical-optics approximation.

[1] G. G. Levin and G. N. Vishnyakov, *Optical Tomography* [in Russian], Radio Svyaz', Moscow (1989).

[2] A. L. Balandin, N. G. Preobrazenskiĭ, and A. N. Sedel'nikov, Zh. Prikl. Mekh. Tekh. Fiz., No. 6, 34–37 (1989).

[3] V. V. Pikalov and N. G. Preobrazenskiĭ, *Reconstructive Tomography in Gas Dynamics and Plasma Physics* [in Russian], Nauka, Novosibirsk (1987).

[4] F. Natterer, *The Mathematics of Computerized Tomography* (Wiley, New York, 1986) [Russian trans., Mir, Moscow, 1990].

[5] R. J. Gardner and P. McMullen, J. London Math. Soc. **21**, 171 (1980).

[6] K. J. Falconer, Proc. London Math. Soc. **46** 241 (1983).

[7] A. B. Vasil'ev and N. A. Tikhonov, *Integral Equations* [in Russian], Moscow State Univ. Press, Moscow (1989).

[8] A. A. Aliverdiev and M. N. Karimov, ''Using of Stream at Nonlinear Enviroment for Investigation of Inside Structure of Bioobjects,'' LALS-94, Minsk (1994), p. 130.

[9] A. A. Aliverdiev, ''Solution of the Stochastic Reconstruction Problem with the Help of Flow Theory'' [in Russian], Collected Works of Students and Post-Graduates, Natural Sciences (Izd. PoligraficheskiĭTsentr DGU, Makhachkala, 1995), pp. 14–17.

Translated by Paul F. Schippnick

Device for measuring the parameters of an ion beam of circular cross section, permitting estimation of the beam emittance per pulse

L. P. Veresov, O. P. Veresov, and L. P. Skripal'

Sukhumi Physicotechnical Institute, Sukhumi, Georgia

A diagnostic device is described for estimating the beam emittance per pulse (for pulse durations from 1 μs up to a continuous beam) for ion beams of circular cross section with currents from a few milliamperes to hundreds of milliamperes and energies of 10 keV and higher. © *1997 American Institute of Physics.* [S1063-7842(97)03109-7]

In our own laboratory practice, to study ion beams of circular cross section we have used a device that allows us to measure the beam current, beam size, the stepped current density distribution of the beam, and the angular divergence of a beam jet cut out at the center of the device by a slit diaphragm. Taken together, these parameters allow one to estimate instantaneously the beam emittance and to alter the operation of the ion source during the course of an experiment with the aim of varying the emittance.

The device (see Fig. 1) consists of a Faraday cylinder joined to an angular divergence meter. The Faraday cylinder is made up from graphite rings which together form an inner cavity in the shape of a truncated cone of height 110 mm with diameters 10 and 60 mm. From the outside it is a cylinder ($D = 70$ mm, $L = 130$ mm). The rings are spaced at uniform intervals, isolated from one another, and form flow-through gaps. The diameters of the inner openings of each ring vary in the sequence 60/50, 50/40, 40/30, 30/20, 20/10 mm. At the center of the last (sixth) ring of the Faraday cup with a working diameter of 10 mm is located a six-millimeter cylindrical opening. The Faraday cup is situated inside a metal shield. To facilitate evacuation, a large number of small openings are emplaced over the entire surface of the shield. To suppress secondary electron emission, two magnetic plates of ferrite–barium are mounted to the shield, creating a transverse magnetic field ≈ 200 G.

The above design of the Faraday cup solves two problems: 1) in addition to the total beam current, it allows one to measure the current from each ring and to determine stepwise the current density with radius and the beam diameter; 2) it hinders the formation in the collector of near-wall beam plasma influencing the measurement of the beam current.

On the other side of the end-face (sixth) ring is located the slit excising a beam jet with the help of two movable metal plates. This slit is part of the angular divergence meter (for measuring the angular divergence of the beam jet). At a distance of 50 mm from the slit is located the receiver which transforms the electrical signal into a light signal. The receiver is mounted with the help of a waveguide, which in addition to this function shields the beam jet and sensor from external influences. As the beam-jet current sensor we use a luminescent quartz glass in the form of plates stacked in a column with metal spacers (0.1-mm foil) or in the form of

small rods gathered together into a ''fence.'' The resolution of the luminescent quartz glasses of grade KLL-1 at energies from 10 keV to 2 MeV ranged from 3×10^{-6} to $3.6 \times 10^{-9} A/cm^2$ (Refs. 1 and 2). The excised jet current excites a light pulse in the glass which is easily observed visually or recorded by a linear CCD array.

The divergence angle of the beam jet is determined from the number of glasses lighting up. The parameters of the meter are: base 50 mm, slit 0.5 mm, number of plates (rods) 20, spacing 0.5 mm.

At first, the operation of the device proceeded as follows: the divergence angle was determined from the number of illuminated glasses, and the beam diameter from the current from each ring of the Faraday cup, which allowed us to construct a diagram of the emittance in the phase plane. With the help of a summator and a storage oscillograph we determined the total beam current and the beam current density distribution with radius. Later all these parameters were read into an automated system which allowed us to estimate the beam emittance per pulse with the help of a computer. In light of the rapid development of computer technology, we do not recommend our system, but rather recommend the software package ''Ostsillograf'' intended for use with analog-to-digital boards available from the Moscow firm ''L-card.''
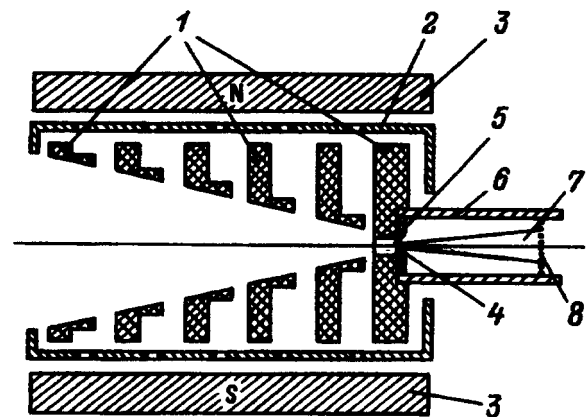

FIG. 1. Diagram of device: *1* — graphite rings, *2* — shield, *3* — magnetic plates, *4* — gap, *5* — movable plates, *6* — waveguide, *7* — beam jet, *8* — receptacle.

It is necessary to mention that the beam emittance estimated by the device somewhat exceeds its actual value since instead of the ellipse in the phase plane its rectangular envelope is used. Many years of experience with the described device have demonstrated its reliability, its ease of use in making beam parameter measurements, its high sensitivity—close to that of the photo method, and the elimination of secondary electron emission effects on the measurements.

The device can be used in both pulsed and continuous modes of the ion source.

[1] E. A. Meleta, N. I. Malykh, I. K. Sidorova et al., Prib. Tekh. Éksp., No. 2, 31 (1991).
[2] V. A. Moskalev and G. I. Sergeev, *Measurement of Parameters of Charged Particle Beams* [in Russian], Énergoizdat, Moscow (1991).

Translated by Paul F. Schippnick

Effect of thermal−field treatment and ionizing radiation on the energy spectrum of interfacial states at the Si−SiO₂ interface of a MOS transistor

A. É. Atamuratov, S. Z. Zaĭnabidinov, A. Yusupov, Kh. S. Daliev, and K. M. Adinaev

Ugrench State University (Al-Khorezmi), 740000 Ugrench, Uzbekistan

The energy distribution of interfacial states at the Si−SiO₂ interface is associated with defects whose nature still remains unclear. The present paper examines the effect of irradiation and thermal−field treatment on the energy spectrum of the interfacial states at the Si−SiO₂ interface of a MOS transistor.

The samples were (100) crystallographic orientation silicon-based (KÉF 7.5) p-channel MOS transistors. The subgate oxide, of thickness 0.2 μm and area 1.87 $\times 10^{-4}$cm², was formed by thermal oxidation of a silicon single crystal in a stream of dry oxygen with a small admixture of hydrogen chloride gas at 1150 °C. The samples were irradiated by high-energy gamma bremsstrahlung with maximum photon energy in the bremsstrahlung spectrum equal to 30 MeV and energy flux density on the sample equal to 0.195 W/cm². Thermal−field treatment was performed in the temperature range 100–250 °C with storage time at each temperature equal to 120 min and electric field intensity in the oxide equal to 5×10^5 V/cm.

The aim of this work is to study the variations of the shape of the energy spectrum of the interfacial states at the Si−SiO₂ interface of the MOS transistor when irradiated by bremsstrahlung and when subjected to thermal−field treatment. The density of interfacial states was determined by the subthreshold current method.[1] This method does not permit one to determine the form of the energy spectrum of the interfacial states over the entire band gap of silicon for each sample, but it does allow one to obtain definite information about its variation over a small energy interval under different conditions. According to Refs. 1 and 2, this technique can give an error in determining the density of interfacial states at the Si−SiO₂ interface for p-channel MOS transistors not greater than 20% in comparison with the traditional high-frequency $C-V$ method and the charge pump method.[3,4] The energy spectrum of interfacial states in the band gap of silicon was estimated in the interval 0.25–0.50 eV above the top of the silicon valence band.

The results show that upon irradiation the integrated density of interfacial states at the Si−SiO₂ interface in the investigated range of the energy spectrum exhibits a monotonic growth with increase of the irradiation time. It is noteworthy that here the largest increase of the differential density of interfacial states is observed with approach to the valence band (around 0.25 eV from the top of the valence band). For long irradiation times (650 min) growth of the density of interfacial states is observed mainly near the lower bound of the investigated region (Fig. 1).

Thermal−field treatment of unirradiated transistors leads to an abrupt change in the shape of the spectrum of interfacial states. Thermal−field treatment at 100, 150, 200 °C does not lead to noticeable changes in the spectrum, but at 250 °C it causes a sudden increase in the differential density of interfacial states around 0.25 eV from the top of the silicon valence band as for irradiation. This apparently indicates the formation of defects which are similar in nature and possibly associated with strained bonds in the region of the Si−SiO₂ interface (Ref. 5). But in the case of irradiation, in contrast to thermal−field treatment, there is also observed an inhomogeneity of the energy spectrum around 0.40 eV from the top of the valence band, which is probably due to the formation in this case of additional defect types. This difference is obviously due to the different character of the processes leading to the formation of defects at the Si−SiO₂ interface. Irradiation leads to a displacement of the subthreshold transmission characteristics toward large negative values, and thermal−field treatment—to small negative values along the gate voltage axis (Fig. 2). This indicates a different variation of posi-
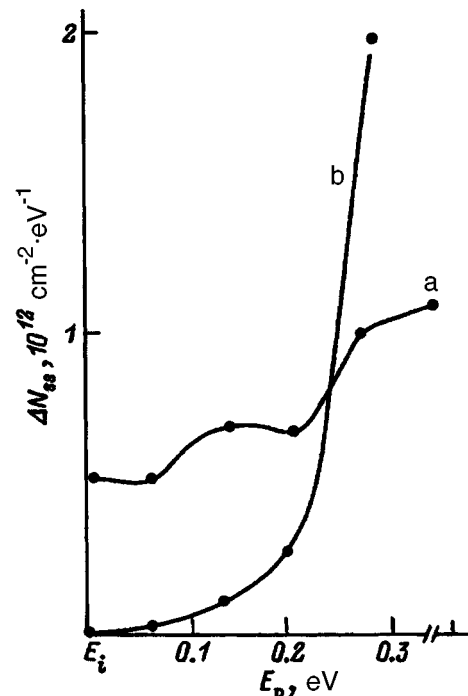

FIG. 1. Surface states energy spectrum near the middle of the silicon band gap of the Si−SiO₂ interface of a MOS transistor subjected to irradiation (a) and to thermal−field treatment (b).
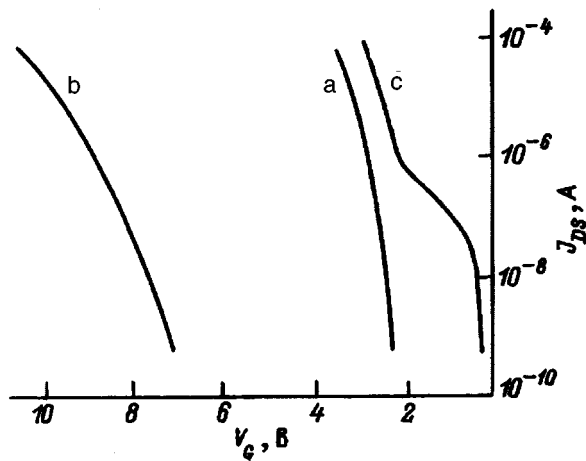
1063-7842/97/091106-02$10.00

FIG. 2. Subthreshold transmission characteristics of a MOS transistor in its initial state (a), after irradiation (b), and after thermal–field treatment (c).

tive charge near the Si–SiO$_2$ interface: irradiation leads to an increase of positive charge while thermal–field treatment

leads to a decrease. The first of these is due mainly to generation of electron–hole pairs,[6] and the second, probably to the untrapping of positively charged formations (holes or hydrogen ions) responsible for the positive charge near the interface and genetically linked with the formation of the MOS transistor structure. The participation of the positively charged formations in the physical reactions near the Si–SiO$_2$ interface during thermal–field treatment apparently leads to an increase in the density of interfacial states.

[1] P. S. Winokur, J. R. Schwank, P. J. McWhorter et al., IEEE Trans. Nucl. Sci. **NS-31**, 1453 (1984).
[2] J. R. Schwank, P. S. Winokur, P. J. McWhorter et al., IEEE Trans. Nucl. Sci. **NS-31**, 1434 (1984).
[3] L. M. Terman, Solid–State Electron., No. 5, 285 (1962).
[4] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. De Keersmaecker, IEEE Trans. Electron Dev. **ED-31**, 42 (1984).
[5] C. T. Sah, IEEE Trans. Nucl. Sci. **NS-23** 1563 (1976).
[6] C. W. Gwyn, J. Appl. Phys. **40**, 4886 (1969).

Translated by Paul F. Schippnick

Spatial orientation of molecules by a heat flux

A. E. Bakarev and A. I. Parkhomenko

*Institute of Automation and Electrometry, Siberian Branch of the Russian Academy of Sciences,
630090 Novosibirsk, Russia*

An effect of spatial orientation of molecules by a heat flux is predicted. Under typical experimental conditions the magnitude of the constant electric field arising due to static polarization of a gas of oriented molecules can reach values $\sim 10^{-4}$ V/cm. © *1997 American Institute of Physics.* [S1063-7842(97)03309-6]

INTRODUCTION

It is well known[1,2] that when polar molecules, i.e., molecules not possessing a symmetry center, drift (or diffuse) relative to a buffer gas they orient along the direction of drift. Since polar molecules possess a dipole moment, their orientation gives rise to a static polarization of the gas and, consequently, a static electrical field. The phenomenon described in Refs. 1 and 2 is physically analogous to the orientation of a weather vane by the wind. The role of the weather vane in the given case is played by the polar molecules, and that of the wind, by the buffer gas relative to which the molecules drift.

In the present paper we want to turn attention to a new mechanism of spatial orientation of molecules located in a buffer medium. This mechanism does not require either drift or diffusion of the particles. It is due to heat transfer and the dependence of the cross section of collisional relaxation of the orientation of the molecules on the relative velocity of the buffer particles flying into them. The molecules orient along the temperature gradient.

GENERAL EXPRESSIONS

Let us consider impurity molecules found in a buffer gas atmosphere. We neglect collisions between the molecules, assuming the buffer gas particle density N to be much greater than that of the molecules ρ ($N \gg \rho$). As is well known, the rotational motion of the molecules in the gas is almost always classical. Therefore we restrict the discussion to a classical description of the orientational degrees of freedom of the molecules. For simplicity we take the molecules to be linear so that the spatial orientation of a molecule is characterized by a single vector $\mathbf{n}$ parallel to the axis of the molecule. We describe the distribution of the molecules over orientations by the equation

$$\frac{\partial}{\partial t}\rho(\mathbf{n}) = S(\mathbf{n}) + R(\mathbf{n}). \tag{1}$$

Here $S(\mathbf{n})$ is the collision integral, the function $R(\mathbf{n})$ allows for the effect of free rotation of the molecules on the distribution $\rho(\mathbf{n})$. For the collision integral we use the model of an isotropic ''approach'' in orientations $\mathbf{n}$

$$S(\mathbf{n}) = -\nu(\mathbf{n})\rho(\mathbf{n}) + S_2, \tag{2}$$

where the term S_2 describes the isotropic ''approach'' in $\mathbf{n}$.

The collision frequency $\nu(\mathbf{n})$ characterizes the collisions of molecules with the buffer particles which establish equilibrium in $\mathbf{n}$.

Let us establish the nature of the behavior of the function $R(\mathbf{n})$ and give it a concrete form. Toward this end, we consider a group of molecules with orientation $\mathbf{n}_0$ at the initial time. These molecules are distributed over rotational levels J and molecules with different J have different angular velocities of rotation ω_J. The values of ω_J are distributed according to some law with a characteristic width $\Delta\omega$. Therefore the molecules, which at the initial time all had the same orientation $\mathbf{n}_0$, very rapidly (in a time $1/\Delta\omega$) become ''well mixed'' in their orientations. In line with what has just been said we model the effect of free rotation of the molecules on the distribution $\rho(\mathbf{n})$ by the function

$$R(\mathbf{n}) = -\Delta\omega\rho(\mathbf{n}) + \Delta\omega\frac{\rho}{4\pi}, \quad \rho = \int \rho(\mathbf{n})d\mathbf{n}. \tag{3}$$

The first (negative) ''drift'' term describes the loss of orientation of the molecules due to their free rotation. The second (positive) term in $R(\mathbf{n})$ describes the rotational ''approach'' of the molecules to the region of orientations $\mathbf{n}$. The function $R(\mathbf{n})$ possesses the obvious property $\int R(\mathbf{n})d\mathbf{n} = 0$, which means that by itself free rotation of the molecules does not change their number.

Taking into account that usually $\Delta\omega \gg \nu(\mathbf{n})$, under stationary conditions we find from Eqs. (1)–(3) that

$$\rho(\mathbf{n}) = \frac{\rho}{4\pi}\left[1 - \frac{\nu(\mathbf{n})}{\Delta\omega}\right]. \tag{4}$$

The spatial orientation vector $\mathbf{q}$ of the gas of molecules is given by

$$\mathbf{q} = \int \mathbf{n}\frac{\rho(\mathbf{n})}{\rho}d\mathbf{n}. \tag{5}$$

Let gradients of the parameters (temperature, density) of the gaseous medium exist in the direction parallel to the z axis. The orientation vector $\mathbf{q}$ can be directed only along this preferred direction. Taking this into account, we find from Eqs. (4) and (5) that

$$\mathbf{q} = \mathbf{l}q, \quad q = -\frac{1}{4\pi\Delta\omega}\int \mathbf{n}\cdot\mathbf{l}\nu(\mathbf{n})d\mathbf{n}, \tag{6}$$

where **l** is the unit vector in the z direction, which we take to be in the direction of the temperature gradient.

SPATIAL ORIENTATION

The problem of finding the orientation **q** reduces to calculating the frequency of collisional relaxation of the orientation of the molecules $\nu(\mathbf{n})$. To calculate $\nu(\mathbf{n})$ we consider the simplest case, in which the mass of the molecules M is large in comparison with the mass of the buffer particles m ($M \gg m$), so that the molecules may be taken to be fixed against the background of the ''fast'' buffer particles.

We characterize collisional relaxation of the orientation of the molecules **n** by the cross section $\sigma(v, \mathbf{n} \cdot \mathbf{k})$, which depends on the speed $v = |\mathbf{v}|$ of the buffer particles flying into the molecule (in general they depend on the relative velocity) and the configuration of the collision defined by the scalar product $\mathbf{n} \cdot \mathbf{k}$, where **k** is the unit vector in the direction of the velocity **v** of the buffer particles. We use the formula[1)]

$$\nu(\mathbf{n}) = \int j(\mathbf{v}) \sigma(v, \mathbf{n} \cdot \mathbf{k}) d\mathbf{v}, \tag{7}$$

to calculate the frequency $\nu(\mathbf{n})$. Here $j(\mathbf{v}) = |\mathbf{j}(\mathbf{v})|$ and $\mathbf{j}(\mathbf{v})$ is the flux density of the buffer particles with prescribed velocity (both magnitude and direction) impinging on the test molecule located at the point **r**

$$\mathbf{j}(\mathbf{v}) = N(\mathbf{r} - \lambda \mathbf{k}) \mathbf{v} W(v, \mathbf{r} - \lambda \mathbf{k}),$$

$$W(v, \mathbf{r} - \lambda \mathbf{k}) = (\sqrt{\pi} \bar{v})^{-3} \exp\left(-\frac{v^2}{\bar{v}^2}\right), \quad \bar{v} = \sqrt{\frac{2k_B T}{m}}. \tag{8}$$

Here $N(\mathbf{r} - \lambda \mathbf{k})$ and $T \equiv T(\mathbf{r} - \lambda \mathbf{k})$ are the density and temperature of the buffer particles at the location of the last collision, i.e., at the point $\mathbf{r} - \lambda \mathbf{k}$, from which they start into the molecule; $\lambda \equiv \lambda(v)$ is the mean free path from the point of the last collision to the test molecule; and k_B is Boltzmann's constant. Equation (8) assumes that the particle velocity distribution is determined by the location of the last collision. Assuming that the factor NW in Eq. (8) varies only slightly over the distance λ, we find from Eq. (7) that

$$\nu(\mathbf{n}) = \bar{\nu} + \Delta \nu(\mathbf{n}), \tag{9}$$

where $\bar{\nu}$ does not depend on **n**, and the term $\Delta \nu(\mathbf{n})$ is given by

$$\Delta \nu(\mathbf{n}) = -\int \mathbf{k} \cdot \mathbf{l} d\mathbf{k} \int_0^\infty v^3 \lambda(v) \sigma(v, \mathbf{n} \cdot \mathbf{k})$$

$$\times \frac{\partial(NW(v))}{\partial z} dv. \tag{10}$$

Substituting expressions (9) and (10) in Eq. (6), we find for the projection q of the orientation vector **q** on the z axis

$$q = \frac{4\pi}{\Delta \omega} \int_0^\infty v^3 \lambda(v) \bar{\sigma}(v) \frac{\partial(NW(v))}{\partial z} dv, \tag{11}$$

where

$$\bar{\sigma}(v) = \int (\mathbf{n} \cdot \mathbf{l})(\mathbf{k} \cdot \mathbf{l}) \sigma(v, \mathbf{n} \cdot \mathbf{k}) \frac{d\mathbf{n}}{4\pi} \frac{d\mathbf{k}}{4\pi}. \tag{12}$$

Formula (11) describes the spatial orientation of the molecules as a result of various transport processes (diffusion, thermal diffusion, energy and heat transfer). If a linear molecule is nonpolar, then as a consequence of symmetry $\sigma(v, \mathbf{n} \cdot \mathbf{k}) = \sigma(v, -\mathbf{n} \cdot \mathbf{k})$ and it follows from Eq. (12) that $\bar{\sigma}(v) = 0$, i.e., the orientation vector **q** = 0, as it should be in this case. It is to be expected that the larger the dipole moment of the molecule, i.e., the more asymmetrical is the molecule, the larger, generally speaking, will be the difference between $\sigma(v, \mathbf{n} \cdot \mathbf{k})$ and $\sigma(v, -\mathbf{n} \cdot \mathbf{k})$ and the larger will be the value of $\bar{\sigma}(v)$ and the orientation vector **q**.

In the case in which the cross section of collisional relaxation of the orientation of a molecule $\sigma(v, \mathbf{n} \cdot \mathbf{k})$ is independent of the velocity ($\bar{\sigma}(v) = \bar{\sigma} = \text{const}$) expression (11) reduces to

$$q = -\frac{3}{\Delta \omega} j_z \bar{\sigma}, \tag{13}$$

where j_z is the integrated flux density of the buffer particles along the z axis,

$$j_z \equiv \int \mathbf{l} \cdot \mathbf{j}(\mathbf{v}) d\mathbf{v} = -\frac{4\pi}{3} \int_0^\infty v^3 \lambda(v) \frac{\partial(NW(v))}{\partial z} dv. \tag{14}$$

Formula (13) describes the ''weather vane'' effect discovered some time ago by Gel'mukhanov and Il'ichev,[1,2] which being the appearance of spatial orientation of the molecules caused by diffusion of the buffer gas relative to them.

In reality, the cross section of collisional relaxation of orientation of a molecule, generally speaking, always depends on the velocity $\bar{\sigma}(v) = \text{const}$. It is precisely this circumstance, as follows from Eqs. (11) and (14), that delivers spatial orientation of the molecules even in the absence of a particle flux (for $j_z = 0$) if only there exists a temperature gradient in the gaseous medium.

For the mean free path we may approximately set

$$\lambda(v) \approx \frac{v}{\nu_b}, \tag{15}$$

where $\nu_b = \text{const}$ is the mean collision frequency of the buffer particles with each other.

For such a dependence $\lambda(v)$ the projection of the orientation vector q (11) is related in a very simple way with the particle flux density j_z (14) and with the heat flux density Q_z along the z axis

$$Q_z = \frac{m}{2} \int \mathbf{l}(\mathbf{v} - \mathbf{u})^3 N(\mathbf{r} - \lambda \mathbf{k}) W(v, \mathbf{r} - \lambda \mathbf{k}) d\mathbf{v}, \tag{16}$$

where $\mathbf{u} = \mathbf{l} j_z / N$ is the directed velocity

$$q = q_j + q_T, \quad q_j = -\frac{3}{\Delta \omega} j_z \sigma(T),$$

$$q_T = -\frac{6}{5\Delta \omega} \frac{Q_z}{k_B} \frac{d\sigma(T)}{dT},$$

A. E. Bakarev and A. I. Parkhomenko   1109

$$j_z = -\frac{d}{dz}(DN), \quad Q_z = -\frac{5}{2}DNk_B\frac{dT}{dz}, \tag{17}$$

where $D = \overline{v}^2/2\nu_b$ is the diffusion coefficient; $\sigma(T)$ is the effective cross section of collisional orientation, defined by the expression

$$\sigma(T) = \frac{8}{3\sqrt{\pi}\,\overline{v}^5}\int_0^\infty v^4 \exp\left(-\frac{v^2}{\overline{v}^2}\right)\overline{\sigma}(v)dv. \tag{18}$$

In formula (17) for q the first term q_i describes the orientation of the molecules by a particle flux (the ''weather vane'' effect[1,2]), and the second term (unknown previously) q_T describes orientation of the molecules by a heat flux.

NUMERICAL ESTIMATES

Let us estimate the relative values of q_j and q_T. Let a be the characteristic scale of variation both of the temperature T and the density N. Then for $d\sigma(T)/dT \sim \sigma(T)/T$ we have

$$|q_j| \sim |q_T| \sim \frac{3DN|\sigma(T)|}{a\Delta\omega}. \tag{19}$$

Thus, a heat flux can cause the same degree of spatial orientation of the molecules as a particle flux.

We will present numerical estimates of the magnitude of the effect based on formula (19). We represent the collision frequency of the buffer particles with each other, which enters into the diffusion coefficient D, in the form $\nu_b = N\sigma_b\overline{v}$, where σ_b is the gas-kinetic collision cross section corresponding to the frequency ν_b. Then estimate (19) for the fraction of heat-flux oriented molecules takes the form

$$|q_T| \sim \frac{3\overline{v}}{2a\Delta\omega}\frac{|\sigma(T)|}{\sigma_b}. \tag{20}$$

Here $\overline{v} \sim 10^5$ cm/s, $\Delta\omega \sim 10^{11}$ s^{-1}, $a \sim 1$ cm, $|\sigma(T)|/\sigma_b \sim 10^{-2}$, and from formula (20) we obtain $|q_T| \sim 1.5\times 10^{-8}$. The constant electric field corresponding to this fraction of oriented molecules is given by the formula

$$\mathbf{E} = -4\pi\rho d\mathbf{q}_T, \tag{21}$$

where $d = |\mathbf{d}|$, $\mathbf{d} = d\mathbf{n}$ is the dipole moment of the molecule, which is assumed to be directed along the axis of the molecule whose direction is given by the unit vector $\mathbf{n}$. For $d \sim 1D$, $\rho \sim 10^{18}$ cm^{-3}, and $|\mathbf{q}_T| \sim 1.5\times 10^{-8}$, formula (21) gives $|\mathbf{E}| \sim 5\times 10^{-5}$ V/cm.

---

[1]Expression (7) implicitly contains the assumption that the relative orientation of the molecule and buffer particle remains fixed during a collision. This is the well-known Mason–Monchick approximation,[3] widely and successfully used in the calculation of the transport characteristics of molecules in gaseous mixtures.

---

[1] F. Kh. Gel'mukhanov and L. V. Il'ichov, Chem. Phys. Lett. **98**, 349 (1983).
[2] F. Kh. Gel'mukhanov and L. V. Il'ichev, Khim. Fiz., No. 5, 590 (1983).
[3] L. Monchick and E. A. Mason, J. Chem. Phys. **35**, 1676 (1961).

Translated by Paul F. Schippnick

Synthesis and study of iron-containing fullerene complexes

G. N. Churilov, O. A. Bayukov, É. A. Petrakovskaya, A. Ya. Korets, V. G Isakova, and Ya. N. Titarenko

L. V. Kirenskiĭ Institute of Physics, Siberian Branch of the Russian Academy of Sciences, 660036 Krasnoyarsk, Russia

At present there is a clear lack of information about the properties of fullerene complexes containing iron despite the existence of publications on this subject.[1,2] We have synthesized iron-containing fullerenes and studied them by infrared and ultraviolet spectroscopy, electron spin resonance, and Mössbauer spectroscopy. To synthesize iron-containing fullerene complexes, we used a plasma-chemical reactor. The operation of this reactor is based on a self-blowing and self-focusing jet of carbon plasma emanating through a conical opening in the outer graphite electrode into a water-cooled copper column (tube). The central electrode was an S-3 grade rod for spectral analysis with an axial opening filled with carbonyl iron powder (TU 6-09-3000-78). Helium was fed into the jet with a flow rate of 3–4 liter/min, and the power supply of the arc was produced by a 66-kHz alternating current with amplitude 300 A. This setup is a modification of the setups described in Refs. 3 and 4. Next, soot, as usual, was poured in together with some nonpolar solvent. Usually we used benzene or toluene. The solution was filtered and boiled down. The dry residue (fullerene mixture) was our object of study.

Spectra of the fullerene mixture in the ultraviolet and visible (200–800 nm) were taken on a Specord UV–vis setup in hexane solutions, while the spectra in the mid-infrared (400–4000 cm^{-1}) were recorded on a Specord IR-75 in pressed pellets of potassium bromide. The electron absorption spectrum of hexane solutions of our fullerene mixture, obtained without filling the central electrode with carbonyl iron (especially in the ultraviolet), clearly reveal the main characteristic absorption bands of the C_{60} molecule: 217, 227, 257, 328 nm, the 404–408 nm doublet, and the long-wavelength band with maximum at 480 nm and shoulders at 544 and 590 nm (Fig. 1a). The maxima of these bands, except for the long-wavelength band, coincide well with the data in the literature, e.g., Refs. 5 and 6. According to the data in the literature, the long-wavelength band of C_{60} has a maximum at 450 nm. The shift of this maximum to 480 nm in our spectrum is explained by the presence in the mixture of a C_{70} impurity. This also explains the presence of the shoulder at 378 nm, which is a characteristic band of C_{70}. The absorption coefficient of the long-wavelength band for C_{70} is approximately seven times larger than for C_{60}. Therefore a small amount of C_{70} mixed in with C_{60} is much more strongly manifested in the long-wavelength band. In the electron absorption spectra of hexane solutions of fullerene extracts of soot obtained from the carbonyl-iron-containing graphite rod (Fig. 1b), the main changes are observed in the

ultraviolet bands at 217 and 257 nm: a redistribution of intensity takes place inside the bands and new shoulder bands appear. The band at 328 nm and the long-wavelength band at 480 nm remain unchanged.

In the infrared absorption spectrum of the mixture obtained from graphite without iron, bands show up with maxima at the frequencies 525, 575, 1190, and 1428 cm^{-1} (Fig. 2a), which are characteristic frequencies of the vibrational states of the C_{60} molecule. The authors of the works cited in Ref. 7 (review) detected these maxima at the frequencies 527, 576, 1183, and 1428 cm^{-1}. The observed discrepancies are within the limits of experimental error. Absorption bands at frequencies characteristic of C_{70} were not observed in the infrared spectrum of the products of synthesis from iron-free graphite. This is because in the infrared the extinctions of the vibrational bands have similar values;
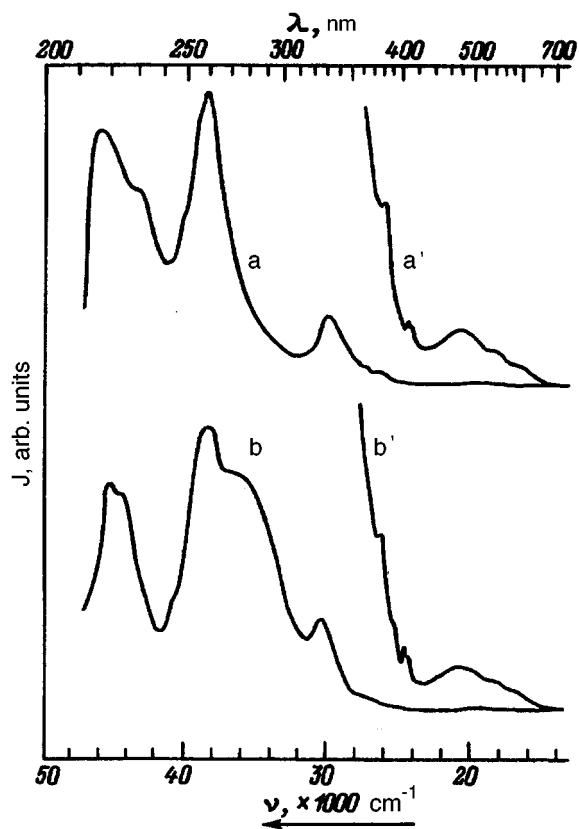

FIG. 1. Electron absorption spectra of hexane solutions of a fullerene mixture obtained from pure graphite (a) and graphite with carbonyl iron (b). Thickness of cuvettes: a,b — 0.1 cm; a′,b′ — 1 cm.
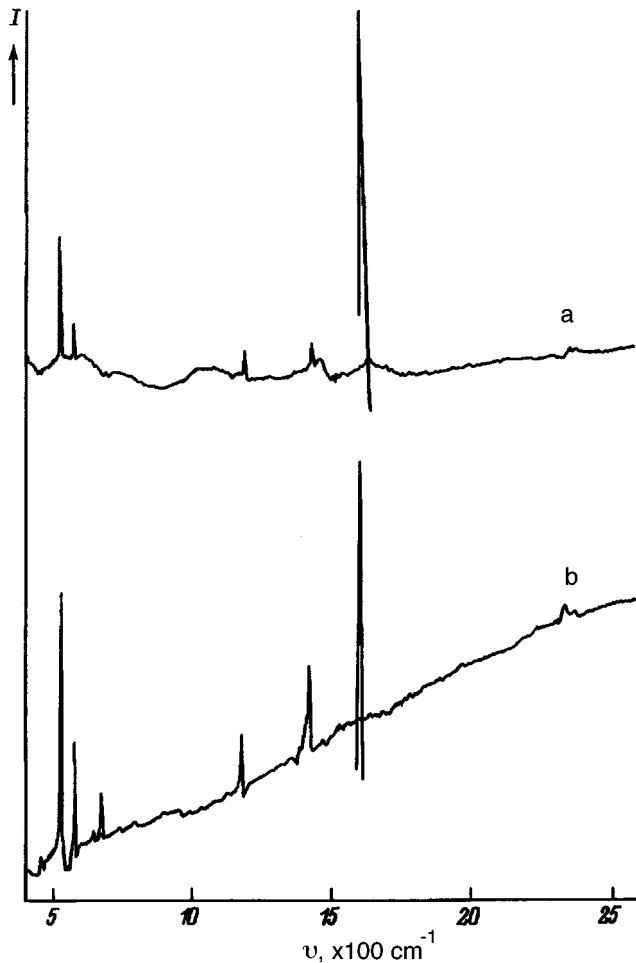
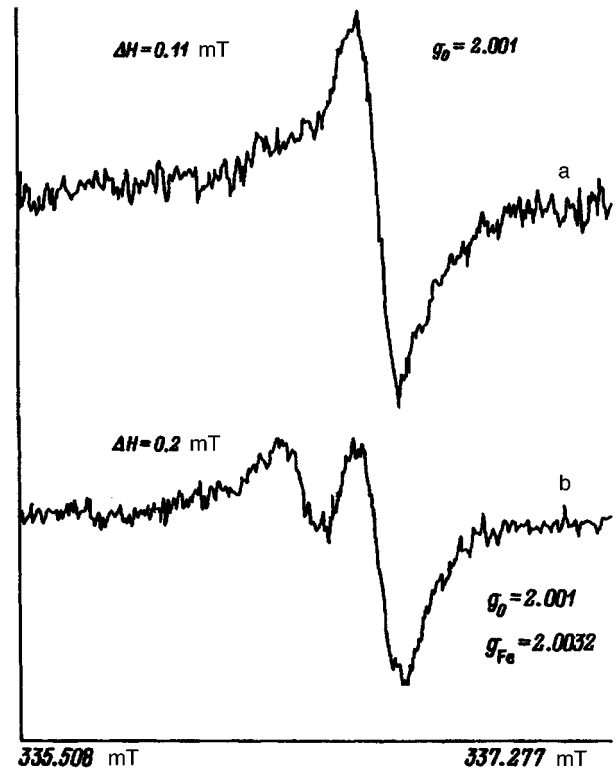FIG. 2. Infrared absorption spectra of a fullerene mixture without iron (a), and with iron (b).

FIG. 3. ESR spectrum of the C_{60} anion-radical (a), and the iron–fullerene complex (b): g_0 is the g factor of C_{60}, g_{Fe} is the g factor of the additional line.

therefore the relative intensities of the absorption lines more adequately reflect the concentration ratios $C_{60} : C_{70}$ than in the ultraviolet and the visible. The presence of a small amount of C_{70} in our fullerene mixture hardly shows up in the infrared spectrum. In the infrared absorption spectrum of the fullerene mixture obtained in an arc with a graphite rod whose axial opening was filled with iron, besides the lines at 528, 577, 1182, and 1430 cm^{-1} with the previous intensity ratios, lines are also absorbed at 673 and 795 cm^{-1} which did not show up earlier (Fig. 2b). These frequencies are close to the frequencies of the iron–carbon stretch vibrations which could be observed in carbides. It is well known that carbides do not dissolve in nonpolar solvents. This means that these bands most likely correspond to a fullerene compound with iron. From the infrared spectrum we determined that carbonides are absent in the mixture, which is important since they are soluble in nonpolar solvents and it would be difficult to distinguish them from fullerenes chemically bound with iron. Thus, our study of this fullerene mixture using optical methods shows that this fullerene mixture contains at a minimum one fullerene compound with iron.

An ESR study of the fullerene mixture was carried out on an x-range spectrometer (SE/X-2544) at temperatures of 80–295 K on solid, polycrystalline samples. The spectra of

fullerenes not containing iron have an absorption line with the parameters $g = 2.001$ and $\Delta H = 0.1$ mT and correspond to the known spectrum of the C_{60} anion-radical[8] (Fig. 3a). For the mixture containing iron–fullerene complexes we have the spectrum shown in Fig. 3b. It contains an additional line with $g = 2.0032$ and $\Delta H = 0.2$ mT. The reaction of the components of the observed doublet to a change in the level of microwave power attests to the presence in this case of two absorption centers with different relaxational characteristics. The ESR spectra of known endohedral complexes[9] show that the ions that have wound up inside the fullerene molecule are found in nonmagnetic states, i.e., the electrons of the atomic shells are paired, but the observed lines are associated with the interaction of the nuclear moment of the ion and the S
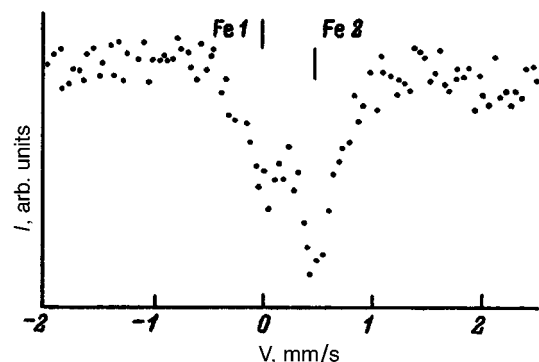

FIG. 4. Mössbauer spectrum of the product of plasma-chemical synthesis of fullerenes with iron.

TABLE I. Mössbauer parameters of the product of plasma-chemical synthesis of complexes of fullerene with iron.

	δ ± 0.02 mm/s	ε ± 0.04 mm/s	Γ ± 0.02 mm/s	S ± 0.05
Fe1	−0.15	0	0.37	0.34
Fe2	0.36	0	0.38	0.66

Note: δ is the isomer shift of α-Fe, ε is the quadrupole splitting, Γ is the spectral linewidth, and S is the fractional population of the indicated iron phase.

electrons weakly polarized by the unpaired electron of the lower molecular orbital of fullerene. The additional line of our spectrum is apparently due to the unpaired electron density arising from the formation of the iron complex. However, the difference of the g factor from 2.001 indicates a change in this case of the fullerene ground state.

With the aim of further study of iron-containing fullerene complexes, we made room-temperature Mössbauer measurements in our fullerene mixture formed by plasma-chemical synthesis, on a NTA-1024 spectrometer with a $Co^{57}(Cr)$ source. The isomer shifts are indicated relative to α-Fe. The spectra were deciphered in the linear approximation by least squares fitting assuming a Lorentzian line shape. The Mössbauer spectrum (Fig. 4) of the product of plasma-chemical fullerene synthesis indicates the presence of two iron phases: Fe1 and Fe2. The Mössbauer parameters of the phases are given in Table I. The Fe1 parameters are characteristic of the strongly covalent state of iron or of four-valent iron found in a highly symmetric local environment.

The fraction of this phase in the sample was 0.34. The isomer shift of Fe2 is characteristic of trivalent iron found in senary symmetric coordination. The fraction of this phase in the sample stood at 0.66. It may be conjectured that iron in this phase is bound with the fullerene surface, but the electron is localized at the iron site. A strong covalent bond or high local symmetry, and the presence of the additional line in the ESR spectrum of iron-containing fullerene complexes are proof that iron is found in the Fe1 state inside the fullerene cage.

[1] T. Pradeep, G. U. Kulkarni, K. R. Kannan et al., J. Am. Chem. Soc., No. 114, 2272 (1992).

[2] L. M. Roth, Y. Huang, J. T. Schwedler et al., J. Am. Chem. Soc., No. 113, 8186 (1991).

[3] G. N. Churilov, International Winter School on Progress in Fullerene Research, Austria (1994), pp. 136–140.

[4] G. N. Churilov, A. Ya. Korets, and Ya. N. Titarenko, Zh. Tekh. Fiz. 66, 191 (1996) [Tech. Phys. 41, 102 (1996)].

[5] H. Yorikowa, M. Itenishi, and S. Muramatsu, J. Phys. Soc. Jpn. 62, No. 10, 3762 (1993).

[6] S. H. Gallagher, R. S. Amstrong, R. A. Lay et al., J. Phys. Chem. 99, 5817 (1995).

[7] H. Kuzmany, R. Winkler, and T. Pichler, J. Phys.: Condens. Matter 7, 6601 (1995).

[8] S. Kukolish and D. Huffman, Chem. Phys. Lett. 182, 263 (1991).

[9] A. Bartl, U. Kirbach, L. Dunsch et al., International Winter School on Progress in Fullerene Research, Austria (1994), pp. 112–115.

Translated by Paul F. Schippnick

New method for solving radiation transfer problems in emitting, absorbing, and scattering media

V. S. Yuferev, M. G. Vasil'ev, and L. B. Proékt

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

The proposed method is based on a novel technique for approximating the angular dependence of the radiated intensity. The entire range of solid angles is divided into N cells, which are symmetric relative to the center of the sphere. In each of the cells the radiation is assigned in the form of the P_1 approximation, and a system of differential equations is obtained to determine the set of local zeroth and first moments. In some special cases the proposed approach can be regarded as a generalization of the discrete-ordinates method, which makes it possible to solve the problem of selecting the weights in the quadrature formulas in a natural manner. The effectiveness of the method is demonstrated in two one-dimensional test cases. It is shown that in these cases fairly high accuracy is achieved in the solution of the problem already for $N=2$. © *1997 American Institute of Physics.* [S1063-7842(97)00109-8]

INTRODUCTION

Radiative heat transfer problems are still among the most tedious from the computational standpoint. Numerous methods have been proposed to solve them (see, for example, the reviews in Refs. 1–3); however, none of them can be regarded as sufficiently universal and suitable for all real cases. Moreover, despite the astounding development of computer technology, there is still no solution for the problem of radiative heat transfer in such a fairly simple geometry as a semitransparent circular cylinder of finite length with specularly reflecting transparent boundaries and a refractive index greater than unity.

The principal methods used to solve radiation transfer problems can be hypothetically divided into two groups, depending on the form of the radiation transfer equation used to obtain the numerical solution, viz., the differential or integral form. The former group includes the spherical-harmonics (P_N) method, the Case eigenfunction method, and the discrete-ordinates (S_N) method. The latter group includes the band method, finite-element analysis, and modifications of those methods. The Monte-Carlo and ray-tracing methods comprise a separate group.

The main problem in solving a radiation transfer equation in the differential form is the choice of the way to approximate the dependence of the radiated intensity on the direction. Two approaches are usually employed. In the first the angular dependence of the radiated intensity is approximated by a set of functions assigned over the entire range of solid angles. This approach can be termed the global approximation. Its practical application is confined to cases in which it is sufficient to use the first terms of the expansion, for example, the P_1 or P_3 approximation in the spherical-harmonics method.[4-6] To improve the accuracy of the P_N method in the presence of discontinuous boundary conditions, Ziering and Schiff[7,8] proposed a modification based on the use of independent expansions[9,10] in each hemisphere in the one-dimensional case and in each quadrant of the sphere in the two-dimensional case.

An alternative approach, which can be termed the local approach, is used in the discrete-ordinates method, where the entire radiation field is divided into a discrete number of fluxes, each of which is associated with a fixed direction in space and a corresponding value for the weight coefficient in the quadrature formulas used to calculate the radiation integrals. As a result, the radiation transfer equation is replaced by a finite system of differential equations, which describe the spatial variation of the radiated intensity in each of these directions. The S_N method was first proposed by Chandrasekhar[11] as a generalization of the Schuster-Schwarzschild method and was subsequently widely employed in the solution of diverse problems in radiative heat transfer (see, for example, Refs. 12–16) and used in commercial software.[17]

The accuracy of the discrete-ordinates method is known to be determined by the design of the quadrature scheme, i.e., by the choice of the weight coefficients. There are presently no rigorous mathematical principles that permit finding the values of the weights, although some fairly interesting results have recently been obtained.[18] This situation lowers the effectiveness of this method significantly. In this paper we propose a new approach to solving radiation transfer problems, which makes it possible to solve this problem in a natural manner.

The approach is based on a new way to approximate the angular dependence of the radiated intensity, which, in a certain sense, combines the spherical-harmonics method and the discrete-ordinates method. In fact, The S_N method can be regarded as a collocation method on a sphere, since the transfer equation is satisfied at a finite set of fixed values of the angles defining the radiation propagation directions. Conversely, in the proposed method the radiation transfer equation is satisfied on the average in each of the elementary solid angles into which the sphere is divided. The angular dependence of the radiation in these cells can be approximated in different ways. In the present work the P_1 approximation is employed for this purpose because of its simplicity and physical clarity. Thus, it can be shown that the proposed

method occupies the same position relative to the S_N method as does finite-element analysis relative to the finite-difference method or the collocation method.

The basic idea of the proposed approach was first formulated in Refs. 19 and 20 as a generalization of the differential approximation and exhibited excellent results in the solution of several test problems. This paper provides a comprehensive description of the method.

FORMULATION OF THE METHOD

Let us divide the entire range of solid angles into N regions (cells) Ω_m in such a manner that $\Omega_m = \Omega_m^+ \cup \Omega_m^-$ and the subregions Ω_m^+ and Ω_m^- are symmetric relative to center of the sphere, i.e., for each direction $\{l_i^{m,+}, i=1,2,3\} \in \Omega_m^+$, where the l_i^m are the direction cosines relative to the coordinate axes x_i, there is a symmetric direction $\{l_i^{m,-}, i=1,2,3\} \in \Omega_m^-$ such that $l_i^{m,-} = -l_i^{m,+}$. We represent the radiated intensity in each of the cells Ω_m in a form similar to the P_1 approximation (to simplify the presentation, the gray approximation is considered)

$$I(\mathbf{r},\Omega) = s_m^{-1}\left(I_0^m(\mathbf{r}) + \sum_{i=1}^{3} l_i^m A_i^m(\mathbf{r})\right), \quad l_i^m \in \Omega_m, \quad (1)$$

where

$$s_m = \int_{\Omega_m} d\Omega.$$

It can easily be shown that because of the symmetry of the angular division into cells, the coefficient I_0^m is equal to the local zeroth moment of the radiated intensity in the region Ω_m, while for the local first and second moments we have

$$q_n^m = \int_{\Omega_m} l_n^m I d\Omega = \sum_{j=1}^{3} p_{nj}^m A_j^m, \quad (2a)$$

$$I_{2,nj}^m = \int_{\Omega_m} l_n^m l_j^m I d\Omega = p_{nj}^m I_0^m, \quad (2b)$$

where

$$p_{nj}^m = s_m^{-1} \int_{\Omega_m} l_n^m l_j^m d\Omega, \quad j,n=1,2,3.$$

It can be seen that Eq. (2b) is equivalent to the closure condition in the P_1 approximation as applied to the cell Ω_m. On the other hand, in contrast to the P_1 approximation, the coefficients A_i^m do not coincide with the components q_i^m of the radiation flux density, but are related to them by Eq. (2a). Expressing the A_j^m from (2a) in terms of the q_n^m and plugging the result into (1), we obtain

$$A_i^m = \sum_{n=1}^{3} \widetilde{p}_{in}^m q_n^m, \quad (3a)$$

$$I = s_m^{-1}\left(I_0^m + \sum_{n=1}^{3} a_n^m q_n^m\right), \quad (3b)$$

where $a_n^m = \sum_{i=1}^{3} \hat{l}_i^m \widetilde{p}_{in}^m$ and the matrix $\{\widetilde{p}_{ij}^m\}$ is the inverse of $\{p_{ij}^m\}$.

The representation (1)–(3) shows that the total moments of the radiated intensity are the sums of the local moments, i.e.,

$$I_0 = \sum_{m=1}^{N} I_0^m, \quad q_i = \sum_{m=1}^{N} q_i^m,$$

$$I_{2,nj} = \sum_{m=1}^{N} I_{2,nj}^m, \quad i,j,n=1,2,3. \quad (4)$$

Thus, all the integral conditions which are used in the discrete-ordinates method to construct the quadrature formulas are satisfied automatically in the present case.

A sphere can be divided in different ways. In the simplest the elementary cells $\Omega_m = \Omega_{pq}^m$ are specified using a spherical coordinate system by the sets

$$(\Theta_{p-1} \leqslant \Theta < \Theta_p) \cup (\pi - \Theta_p < \Theta \leqslant \pi - \Theta_{p-1}),$$

$$(\varphi_{q-1} \leqslant \varphi < \varphi_q) \cup (\varphi_{q-1} + \pi \leqslant \varphi < \varphi_q + \pi),$$

$$1 \leqslant p \leqslant N_1, \quad 1 \leqslant q \leqslant N_2, \quad N = N_1 \cdot N_2. \quad (5)$$

The tensor $\{p_{ij}\}$ corresponding to this division into cells is presented in the Appendix. When one-dimensional and some multidimensional problems are solved, the division (5) can be simplified by setting $0 \leqslant \varphi < 2\pi$, i.e., the division with respect to the angles Θ can be used alone. In the simplest case

$$s_m = 4\pi(\mu_{m-1} - \mu_m), \quad \mu = \cos\Theta,$$

$$p_{ij} = 0, \quad i \neq j, \quad \text{and} \quad A_i^m = \frac{1}{p_{ii}} q_i^m. \quad (6)$$

Another way to construct $\{\Omega_m\}$ involves the symmetric division of the sphere into regions of identical shape and equal area. However, in this case the variety of possible divisions is restricted. The minimum division ($N=3$) is obtained by projecting a cube onto the sphere. Then follow $N=4$, i.e., the projection of an octahedron onto the sphere, $N=6$, i.e., the projection of a rhombic octahedron, etc. In the general case the division into $\{\Omega_m\}$ can depend on the spatial coordinates, and this is another significant advantage of the method under discussion, since it permits utilization of the concrete features of the problem being solved. However, in this paper we shall confine ourselves to a treatment of the situation in which the division into $\{\Omega_m\}$ does not depend on the spatial coordinates.

BASIC EQUATIONS

To obtain the equations with respect to I_0^m and q_i^m we must calculate the local zeroth and first moments of the original radiation transfer equation, which has the form

$$\sum_{j=1}^{3} l_j \frac{\partial I}{\partial x_j} + \beta I = \frac{\omega}{4\pi}\beta$$

$$\times \int_{4\pi} \Phi(\Omega,\Omega')I(\mathbf{r},\Omega)d\Omega' + \beta(1-\omega)I_B, \quad (7)$$

where $\beta = k + \sigma$, $\omega = \sigma/(k+\sigma)$, and k and σ are the absorption and scattering coefficients, respectively.

We integrate Eq. (7) over the region Ω_m. Then, replacing I in the scattering integral by (1) and taking into account the symmetry properties of the scattering indicatrix, we have

$$\sum_{j=1}^{2} \frac{\partial q_j^m}{\partial x_j} = \beta \sum_{n=1}^{N} (\omega r_{mn} - \delta_{mn}) I_0^n + \beta(1-\omega) s_m I_B, \quad (8a)$$

where

$$r_{mn} = \frac{1}{4\pi s_n} \int_{\Omega_n} \int_{\Omega_m} \Phi(\Omega, \Omega') d\Omega d\Omega', \quad (8b)$$

and δ_{mn} is the Kronecker delta.

We next multiply (7) by l_i and integrate again over Ω_m. As a result, with consideration of (1) and (2b) and the symmetry properties of the scattering indicatrix we obtain

$$\sum_{j} p_{ij}^m \frac{\partial I_0^m}{\partial x_j} + \beta q_i^m = \beta\omega \sum_{n=1}^{N} \sum_{j=1}^{3} f_{ij}^{mn} A_j^n, \quad (9a)$$

where

$$f_{ij}^{mn} = \frac{1}{4\pi s_n} \int_{\Omega_m} \int_{\Omega_n} l_i^m l_j^n \Phi(\Omega, \Omega') d\Omega d\Omega'. \quad (9b)$$

The substitution therein of (3a), which relates the coefficients A_j^m to the radiant flux density $\mathbf{q}^m$, gives

$$\sum_{j=1}^{3} p_{ij}^m \frac{\partial I_0^m}{\partial x_j} = \beta \sum_{n=1}^{N} \sum_{l=1}^{3} (\omega h_{il}^{mn} - \delta_{mn}\delta_{il}) q_l^n, \quad (10a)$$

where

$$h_{il}^{mn} = \sum_{j=1}^{3} f_{ij}^{mn} \widetilde{p}_{jl}^n. \quad (10b)$$

The system of first-order differential equations consisting of (8) and (10) specifies the set of local zeroth and first moments I_0^m and q_i^m. It can be reduced in principle to a system of second-order equations with respect to I_0^m or q_i^m. The latter system is more convenient for a numerical solution. Since the reduction procedure is fairly cumbersome in the general case, we shall perform this operation in several concrete examples.

a) Planar layer, axial symmetry. Using the division into cells defined by Eqs. (5) and (6) and assuming that the $0X_3$ axis is perpendicular to the boundaries of the layer, we have

$$f_{ij}^{mn} = \delta_{i3}\delta_{j3} f^{mn}, \quad h_{il}^{mn} = \delta_{i3}\delta_{l3} f^{mn}/p_{33}^n, \quad q^m = p_{33}^m A_3^m.$$

As a result, Eqs. (8) and (10) take the form

$$\frac{1}{\beta} \frac{dq^m}{dx_3} = \sum_{n=1}^{N} (\omega r_{mn} - \delta_{mn}) I_0^n + (1-\omega) s_m I_B, \quad (11a)$$

$$p_{33}^m \frac{dI_0^m}{dx_3} = \beta \sum_{n=1}^{N} (\omega h_{33}^{mn} - \delta_{mn}) q^n. \quad (11b)$$

Differentiating (11a) with respect to x_3 and eliminating I_0, we obtain the desired system of second-order equations with respect to q^m

$$\frac{d^2 q^m}{dx_3^2} = \beta^2 \sum_{n=1}^{N} \sum_{l=1}^{N} \frac{1}{p_{33}^n} (\omega r_{mn} - \delta_{mn})$$

$$\times (\omega h_{33}^{nl} - \delta_{nl}) q^l + \beta(1-\omega) s_m \frac{dI_B}{dx_3}. \quad (12)$$

We note that Eq. (12) holds for any scattering function.

b) Multidimensional case, one-dimensional division. Under certain conditions the one-dimensional division used in the preceding subsection can be sufficiently exact in multidimensional problems as well. Since, as before, $p_{ij} = 0$ for $i \neq j$ in this case, instead of Eq. (12) we have

$$\sum_{j=1}^{3} \frac{\partial^2 q_j^m}{\partial x_j \partial x_i} = \beta^2 \sum_{n=1}^{N} \sum_{l=1}^{N} \sum_{k=1}^{3} \frac{1}{p_{ii}^n} (\omega r_{mn} - \delta_{mn})$$

$$\times (\omega h_{ik}^{nl} - \delta_{nl}\delta_{ik}) q_k^l + \beta(1-\omega) s_m \frac{dI_B}{dx_i}. \quad (13)$$

c) Multidimensional case, division of general form, isotropic scattering. It follows from (8b) and (9b) that $f_{ij}^{mn} = 0$ and $r_{mn} = s_m/(4\pi)$. As a result, eliminating q_i^m from (8) and (9), we have

$$\sum_{i,j=1}^{3} p_{ij}^m \frac{\partial^2 I_0^m}{\partial x_i \partial x_j} = -\beta^2 \sum_{n=1}^{N} \left(\omega \frac{s_m}{4\pi} - \delta_{mn} \right) I_0^n - \beta^2$$

$$\times (1-\omega) s_m I_B. \quad (14)$$

COMPARISON OF THE PROPOSED APPROACH WITH THE DISCRETE-ORDINATES METHOD

As we have already noted, in the discrete-ordinates method Eq. (6) is replaced by a system of equations (to simplify the presentation, the case of isotropic scattering is considered):

$$\sum_{i=1}^{3} l_i^m \frac{\partial I_m}{\partial x_i} + \beta I_m = \frac{\omega}{4\pi} \beta \sum_{n=-N}^{N} \omega_n I_n + \beta(1-\omega) I_B,$$

$$m = \pm 1, \pm 2, \ldots, \pm N, \quad (15)$$

where m labels the corresponding direction in space and the ω_n are the weights in the quadrature formula.

The directions are usually selected in such a manner that $l_i^m = -l_i^{-m}$. Then, introducing the notation $J_0^m = I_{-m} + I_m$ and $q^m = I_m - I_{-m}$, after standard transformations of Eqs. (15), we obtain

$$\sum_{i=1}^{3} l_i^m \frac{\partial J_0^m}{\partial x_i} + \beta q^m = 0,$$

$$\sum_{i=1}^{3} l_i^m \frac{\partial q^m}{\partial x_i} + \beta J_0^m = 2\beta \left(\frac{\omega}{4\pi} \sum_{n=1}^{N} \omega_n J_0^n + (1-\omega) I_B \right). \quad (16)$$

Eliminating q^m from Eq. (16), we have

$$\sum_{i,j=1}^{3} l_i^m l_j^m \frac{\partial^2 J_0^m}{\partial x_i \partial x_j} + \beta J_0^m = -2\beta^2$$

$$\times\left(\frac{\omega}{4\pi}\sum_{n=1}^{N}\omega_n J_0^n+(1-\omega)I_B\right).$$

(17)

A comparison of Eqs. (14) and (17) reveals that they have the same structure. Therefore, the proposed approach can be regarded as a generalization of the discrete-ordinates method that makes it possible to solve the problem of choosing the weights in the quadrature formulas in a natural manner. Moreover, the appearance of the "ray effect" described in Ref. 16 in reference to the S_N method is impossible in the proposed method.

BOUNDARY CONDITIONS

To be specific, let us consider an opaque, emitting, and reflecting boundary. In this case the boundary conditions for the radiated intensity have the form

$$I(\Omega)=\varepsilon I_B(T_\omega)+\rho_s I(\Omega')+\frac{\rho_d}{\pi}\int_{n\cdot\Omega''<0}|\mathbf{n}\cdot\Omega''|I(\Omega'')d\Omega'',$$

$$\mathbf{n}\cdot\Omega>0,$$

(18)

where ε is the emissivity, ρ_s and ρ_d are the specular and diffuse reflection coefficients, $\mathbf{n}$ is the vector of the internal normal to the boundary of the region, and Ω' is related to Ω by the condition that the angle of reflection is equal to the angle of incidence.

Projecting Eq. (18) onto the internal normal and integrating over the region Ω_m or, more precisely, over the directions belonging to Ω_m which satisfy the condition $\Omega\cdot\mathbf{n}>0$, we obtain equations which relate I_0^m and q_i^m on the boundary of the region

$$\int_{\Omega_m,n\cdot\Omega>0}(\mathbf{n}\cdot\Omega)I(\Omega)d\Omega$$

$$=\varepsilon_\omega I_B(T_\omega)\int_{\Omega_m,n\cdot\Omega>0}(\mathbf{n}\cdot\Omega)d\Omega$$

$$+\int_{\Omega_m,n\cdot\Omega>0}(\mathbf{n}\cdot\Omega)\rho_s I(\Omega')d\Omega+\frac{\rho_d}{\pi}\sum_{n=1}^{N}$$

$$\times\int_{\Omega_n,n\cdot\Omega''<0}|\mathbf{n}\cdot\Omega''|I(\Omega'')d\Omega''$$

$$\times\int_{\Omega_m,n\cdot\Omega>0}(\mathbf{n}\cdot\Omega)d\Omega.$$

(19)

The calculation of the integrals appearing in Eq. (19) is the most tedious part of writing a numerical algorithm based on the use of the method under consideration. It is easiest to perform these calculations in the one-dimensional case for a planar layer. Using the division defined by Eqs. (5) and (6) we obtain

$$I_0^m(1-\hat{\rho}_s^m)+2q_0^m\frac{1+\widetilde{\rho}_s^m}{\mu_{m-1}+\mu_m}=\varepsilon I_B s_m-\frac{\rho_d}{\pi}s_m$$

$$\times\sum_{n=1}^{N}\left(\frac{q^n}{2}-I_0^n\frac{\mu_n+\mu_{n-1}}{4}\right),$$

where

$$\hat{\rho}_s^m=\frac{2}{\mu_{m-1}^2-\mu_m^2}\int_{\mu_m}^{\mu_{m-1}}\mu\rho_s d\mu,$$

$$\widetilde{\rho}_s^m=\frac{1}{\mu_{m-1}-\mu_m}\int_{\mu_m}^{\mu_{m-1}}\frac{\mu^2\rho_s(\mu)}{p_{33}}d\mu.$$

NUMERICAL EXAMPLES

Two problems were considered to demonstrate the effectiveness of the proposed method.

1. Passage of radiation through a planar layer of a linearly anisotropic scattering medium: $\Phi(\Omega,\Omega')=1+b\cos\psi$. It was assumed that the refractive index of the medium equals 1 and that the external radiation propagates along a normal to the boundary of the layer. This problem can be treated as a test case, since its solution is known and can be found, for example, in Ref. 4. Uniform division of the range of variation of μ from 0 to 1 was used to solve the problem just described. Figure 1 shows the dependence of the error in the calculation of the reflection coefficient R of the layer on the number of cells N for several values of the albedo ω, the optical thickness τ of the layer, and the scattering anisotropy coefficient b. The error was calculated from the formula

$$\mathcal{E}=\frac{R_n-R_{n+1}}{R_n}100\%.$$

The position of the theoretical points in Fig. 1 between vertical lines $N=$ const indicates that the error exceeds the limits of the figure. It is seen that convergence of the solution is achieved in practice for $N=2-3$.

2. Radiative–conductive heat transfer in a semitransparent plate heated by external radiation. It was assumed that there is no scattering of the radiation and that the absorption coefficient is constant. Similar problems have also previously been considered on numerous occasions (see, for example, Refs. 4 and 5). The goal pursued in this case was to demonstrate the high effectiveness of the method, which is associated with the possibility of taking into account specific features of the problem, particularly total internal reflection. These features were taken into account, because the angle of total internal reflection Θ_B coincided with one of the division angles Θ_m. Figure 2 shows the temperature distribution across the thickness of the plate. It is seen that excellent agreement with the exact solution is achieved when $N=2$.

CONCLUSIONS

The approach proposed in this paper for solving radiation transfer problems has the following advantages.

1. The main problem in the discrete-ordinates method, viz., the correct choice of discrete directions and the corresponding weights in the quadrature formulas, is eliminated.

2. It is now possible to take into account specific features of the problem in choosing the division into the Ω_m.
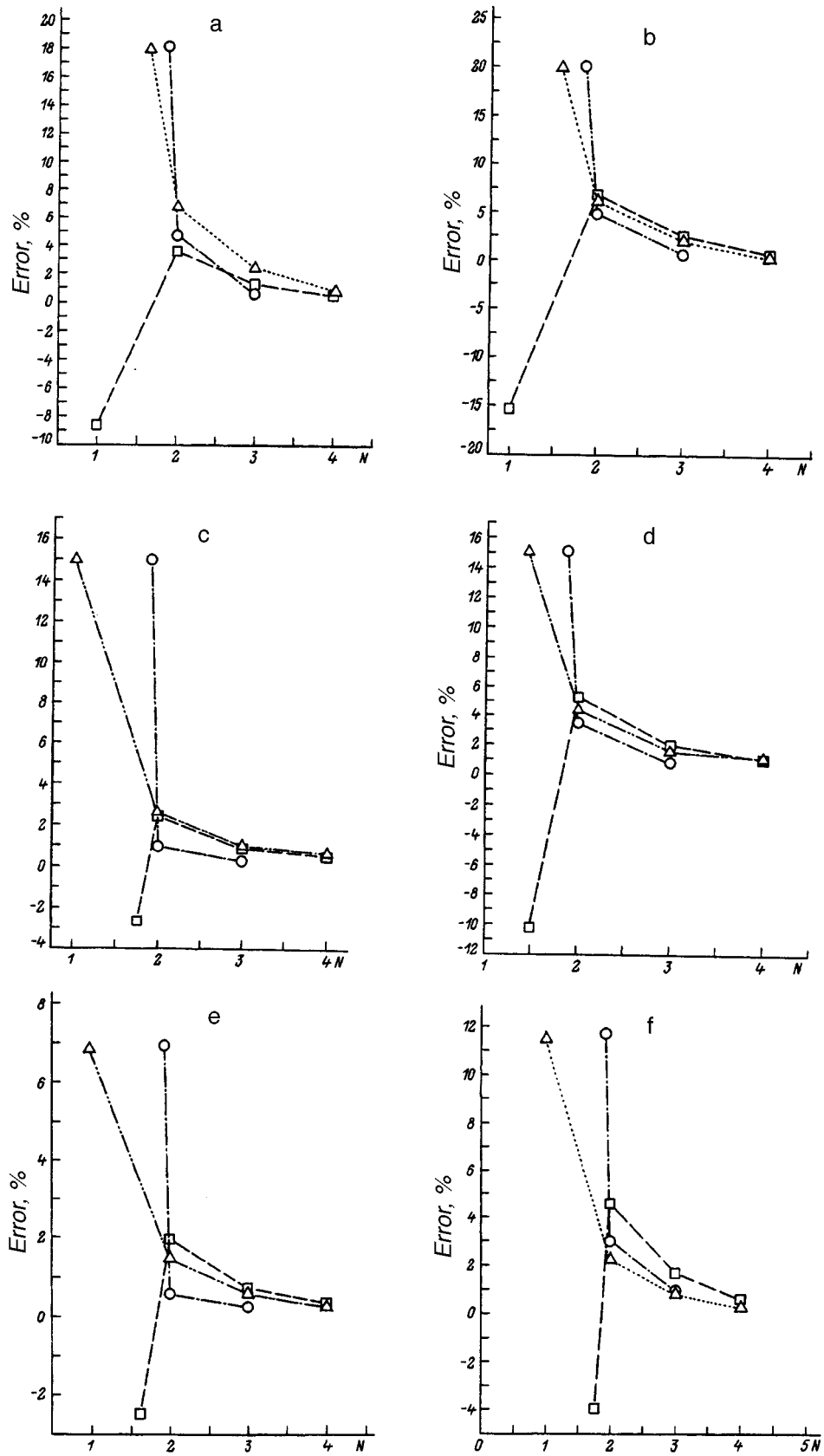
FIG. 1. Dependence of the error in the calculation of the reflection coefficient R on the number of cells N. ω: $\bigcirc$ — 0.1; $\square$ — 0.5; $\triangle$ — 0.9; $\tau=0.1$ (a, b), 1 (c, d), 15 (e, f); $b=0$ (a, c, e), 1 (b, d, f).
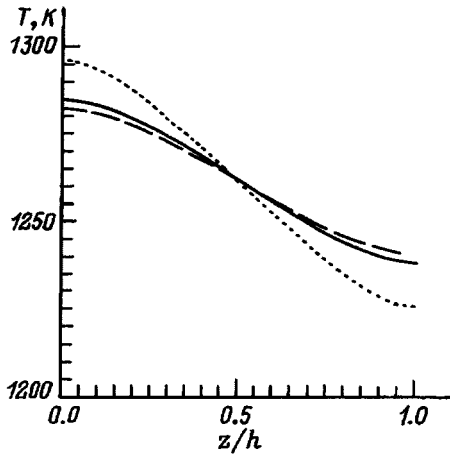
FIG. 2. Distribution of the temperature across the thickness of a semitransparent plate heated on one side by radiation with a temperature $T = 1500$ K. $\tau = 1$; the radiative–conductive parameter equals 0.058; solid curve — exact solution, dotted curve — $N = 1$ (the P_1 approximation), dashed curve — $N = 2$.

3. There is now a possibility, in principle, of correctly approximating the boundary conditions in the case of a boundary of arbitrary form.

4. Since the P_1 approximation itself gives satisfactory results in many problems, it can be expected that satisfactory accuracy can be achieved when a fairly rough division into the Ω_m is used. The results of test calculations presented confirm these expectations. Already $N = 2$, i.e., the first correction to the P_1 approximation, provides a solution that is very close to the exact one.

5. If there is no scattering or the scattering indicatrix is isotropic, the method makes it possible to easily reduce the original problem to the solution of a system of second-order equations, which are solved relative to the former derivatives.

6. The method can be generalized to cases where the division into the Ω_m depends on the spatial coordinates.

The method can be used already in the present stage to solve one-dimensional problems with an arbitrary scattering indicatrix, as well as to solve two-dimensional and three-dimensional problems in a rectangular region, if the medium has isotropic scattering properties.

The main problems associated with the use of this method that require further investigation are as follows: 1) testing the effectiveness of the method in $2D$ and $3D$ problems; 2) constructing symmetric divisions, in which the regions Ω_m have the same shape and area; 3) developing effective methods for calculating the integrals appearing in the boundary conditions for a boundary region of arbitrary form.

APPENDIX A: THE TENSOR p_{ij}

We introduce the notation

$$l_1 = \sin \Theta \cos \varphi = \sqrt{1 - \mu^2} \cos \varphi,$$

$$l_1 = \sin \Theta \sin \varphi = \sqrt{1 - \mu^2} \sin \varphi, \quad l_3 = \cos \Theta = \mu.$$

Then

$$s_m = 2(\mu_{m-1} - \mu_m)(\varphi_m - \varphi_{m-1}),$$

$$p_{11} = \frac{1}{2}(1 - f)(1 + e),$$

$$p_{12} = p_{21} = g \frac{\cos 2\varphi_{m-1} - \cos 2\varphi_m}{4(\varphi_m - \varphi_{m-1})}(1 - f),$$

$$p_{22} = \frac{1}{2}(1 - f)(1 - e),$$

$$p_{13} = p_{31} = g \frac{\sin \varphi_m - \sin \varphi_{m-1}}{\varphi_m - \varphi_{m-1}},$$

$$p_{23} = p_{32} = g \frac{\cos \varphi_{m-1} - \cos \varphi_m}{\varphi_m - \varphi_{m-1}}, \quad p_{33} = f,$$

where

$$f = \frac{1}{3}(\mu_m^2 + \mu_m \mu_{m-1} + \mu_{m-1}^2),$$

$$e = \frac{\sin 2\varphi_m - \sin 2\varphi_{m-1}}{2(\varphi_m - \varphi_{m-1})},$$

$$g = \frac{(1 - \mu_m^2)^{3/2} - (1 - \mu_{m-1}^2)^{3/2}}{3(\mu_{m-1} - \mu_m)}.$$

[1] J. R. Howell, J. Heat Transfer **110**, 1220 (1988).
[2] S. H. Chan, Rev. Numer. Fluid Mech. Heat Transfer **1**, 305 (1987).
[3] R. Viscanta and M. P. Menguc, Prog. Energy Combust. Sci. **13**, 97 (1987).
[4] M. N. Ozisik, *Radiative Transfer and Interactions with Conduction and Convection*, Wiley, New York (1973) [Russian trans., Mir, Moscow, 1976], p. 616.
[5] R. Siegel and J. R. Howell, *Thermal Radiation Heat Transfer*, McGraw-Hill, New York (1972) [Russian trans., Mir, Moscow, 1975], p. 934.
[6] M. P. Menguc and R. Viscanta, J. Quant. Spectrosc. Radiat. Transfer. **33**, 533 (1985).
[7] S. Ziering and D. Schiff, Nuclear Sci. Eng. **3**, 635 (1958).
[8] D. Schiff and S. Ziering, Nuclear Sci. Eng. **7**, 172 (1960).
[9] J. Yvon, J. Nuclear Energy **4**, 305 (1957).
[10] K. M. Case and P. F. Zweifel, *Linear Transport Theory*, Addison-Wesley, Reading, Mass. (1967) [Russian trans., Mir, Moscow, 1972].
[11] S. Chandrasekhar, *Radiative Transfer*, Dover, New York (1960).
[12] B. G. Carlson and K. D. Lathrop, *Computing Methods in Reactor Physics*, H. Greenspan, C. N. Kelber, and D. Okrent, (Eds.), Gordon and Breach, New York (1968).
[13] W. A. Fiveland, J. Heat Transfer **109**, 809 (1987).
[14] J. S. Truelove, J. Heat Transfer **109**, 1048 (1987).
[15] A. Sanchez and T. F. Smith, J. Heat Transfer **114**, 465 (1992).
[16] J. A. Menart, H. S. Lee, and T. K. Kim, J. Heat Transfer **115**, 184 (1993).
[17] M. G. Giridharam, S. Lowry, and A. Krishnan, 30th AIAA Thermophysics Conference, San Diego (1995).
[18] R. Koch, W. Krebs, S. Wittig, and R. Viscanta, J. Quant. Spectrosc. Radiat. Transfer **43**, 353 (1995).
[19] V. S. Yuferev, Teplofiz. Vys. Temp. **33**, 961 (1995).
[20] K. D. Lathrop, Nucl. Sci. Eng. **32**, 357 (1968).

Translated by P. Shelnitz

Instability of a charged layer of a viscous liquid on the surface of a solid spherical core

A. I. Grigor'ev, V. A. Koromyslov, and S. O. Shiryaeva

Yaroslavl State University, 150000 Yaroslavl, Russia

The dispersion relation for the spectrum of capillary waves of a spherical layer of a viscous liquid coating a solid spherical core with a layer of finite thickness is introduced and analyzed. It is shown that the existence of two mechanisms for the viscous dissipation of the energy of the capillary-wave motions of the liquid, viz., damping in the bulk of the layer and on the solid core, leads to restriction of the spectrum of the realizable capillary waves of the liquid on both the high- and low-mode sides. At a fixed value of the system charge which is supercritical for the first several capillary modes, the maximum growth rates in the case of a small solid core are possessed by modes from the middle of the band of unstable modes, while in thin liquid layers the highest of the unstable modes have the largest growth rates. This points out differences in the realization of the instability of the charged surface of the spherical layer for small and large relative sizes of the solid core. © *1997 American Institute of Physics.* [S1063-7842(97)00209-2]

Numerous academic, technical, and technological problems call for dealing with the instability of a charged layer of a viscous liquid of finite depth lying on a curvilinear solid substrate. The instability of the charged liquid surface of a thawing hailstone in a thundercloud, which is accompanied by the emission of a considerable number of charged microdroplets,[1] plays an important role in microscopic charge separation processes and in the ignition of the discharging of streak lightning.[2] Apart from the geophysical ramifications of the electrostatic instability of a charged layer of a viscous liquid, this phenomenon is encountered in liquid mass spectrometry.[3,4] For example, in some types of liquid mass spectrometers ions of nonvolatile and thermally unstable substances are obtained by the emission of microdroplets and clusters from a meniscus at the tip of a metallic capillary tube, through which a solution is supplied to the discharge system, under low-temperature (~ 100 K) vacuum conditions. Because of the low temperature, the solution at the end of the capillary tube freezes, and microdroplets are emitted electrohydrodynamically from the liquid film on the surface of the ice core.[3,5] The existence of the liquid film is ensured by Joule heating from the passage of an electric current through the film. A qualitatively similar situation, i.e., the emission of charged microdroplets from the surface of a thin liquid film in a strong electric field should also be encountered in liquid-metal ion sources.[6]

The problem of the stability of a charged layer of a viscous liquid on the surface of a spherical core was previously considered for the asymptotic case of a low-viscosity liquid in Ref. 7, where a dispersion relation was obtained, but was not analyzed numerically. For this reason, it would be useful to examine the problem under discussion in a more general formulation without confining ourselves to low-viscosity liquids.

1. Let a solid spherical core of radius R_0 be surrounded by a spherical layer of an ideally conducting liquid of external radius R in a spherically symmetric configuration. The liquid is assumed to be incompressible, to have a low viscosity, and to be characterized by a kinematic viscosity ν, a density ρ, and a surface tension σ. The charge Q is distrib-

uted uniformly over the surface of the liquid phase. The velocity field due to the capillary-wave motion and the pressure field are denoted by $\mathbf{U}(\mathbf{r},t)$ and $P(\mathbf{r},t)$, respectively. The distortion $\xi(\mathbf{r},t)$ of the free surface of the liquid appearing because of the capillary-wave motion is assumed to be small together with $\mathbf{U}(\mathbf{r},t)$ and $P(\mathbf{r},t)$. In addition, we assume that the system has axial symmetry.

To simplify the form of the mathematical expressions and the subsequent calculations, we introduce dimensionless variables, in which $R=1$, $\rho=1$, and $\sigma=1$. Then all the remaining quantities (for which we retain the former notation) will be expressed in units of their characteristic values $r_*=R$, $t_*=R^{3/2}\rho^{1/2}\sigma^{-1/2}$, $U_*=R^{-1/2}\rho^{-1/2}\sigma^{1/2}$, $P_*=R^{-1}\sigma$, $Q_*=R^{3/2}\sigma^{1/2}$, and $\nu_*=R^{1/2}\rho^{-1/2}\sigma^{1/2}$.

The system of hydrodynamic equations describing the capillary oscillations of the liquid in such a system consists of the linearized Navier–Stokes equation

$$\frac{\partial \mathbf{U}}{\partial t} = -\frac{1}{\rho}\nabla P^{\text{in}} + \nu \cdot \nabla^2 \mathbf{U}, \tag{1}$$

the condition for an incompressible liquid

$$\nabla \cdot \mathbf{U} = 0, \tag{2}$$

the boundary conditions on the surface of the solid core

$$r=R_0, \quad \text{whence} \quad U_r=0, \quad U_\Theta=0, \quad U_\varphi=0, \tag{3}$$

and the boundary conditions on the free surface of the liquid

$$F(\mathbf{r},t) \equiv r - 1 - \xi(\Theta,t) = 0,$$

whence

$$\frac{dF}{dt} \equiv \frac{\partial F}{\partial t} + \mathbf{U} \cdot \nabla F = 0, \tag{4}$$

$$\boldsymbol{\tau} \cdot (\mathbf{n} \cdot \nabla)\mathbf{U} + \mathbf{n} \cdot (\boldsymbol{\tau} \cdot \nabla)\mathbf{U} = 0, \tag{5}$$

$$-(P^{\text{in}} - P^{\text{ex}}) + 2\rho\nu\mathbf{n} \cdot (\mathbf{n} \cdot \nabla)\mathbf{U} + P_\sigma - P_E = 0. \tag{6}$$

In these expressions $\xi(\Theta,t)$ is a function which describes the perturbation of the equilibrium spherical surface of the drop; $\mathbf{n}$ and $\boldsymbol{\tau}$ are the unit vectors normal and tangen-

tial to the free surface of the liquid; P^{ex} is the pressure of the external medium on the surface of the drop; P^{in} is the pressure within the liquid; $\mathbf{U}$ is the velocity field; P_σ is the Laplace pressure under the spherical surface of the liquid layer distorted by the wave motion[8]

$$P_\sigma = 2 - (2 + \nabla_\Omega^2)\xi(\Theta, t),$$

where ∇_Ω^2 is the angular part of the Laplacian operator in spherical coordinates.

The pressure of the electric field P_E on the surface of a charged, ideally conducting drop is defined by the expression[7]

$$P_E = \frac{Q^2}{8\pi} - \frac{1}{2\pi}Q^2\xi + \frac{1}{4\pi}Q^2$$

$$\times \sum_{m=0}^{\infty} (m+1)P_m(\mu)\int_{-1}^{1} \xi P_m(\mu)d\mu,$$

where the $P_m(\mu)$ are normalized Legendre polynomials, and $\mu \equiv \cos\theta$.

2. To solve the problem defined by Eqs. (1)–(6) we use a scalarization method similar to the one described in Ref. 9. Accordingly, we represent the velocity field in the form of a sum of three orthogonal fields:

$$\mathbf{U}(\mathbf{r},t) = \nabla\Psi_1(\mathbf{r},t) + \nabla\times\mathbf{r}\Psi_2(\mathbf{r},t) + \nabla\times(\nabla\times\mathbf{r})\Psi_3(\mathbf{r},t),$$

where the first term gives the potential part of the velocity field and the second and third the solenoidal parts, viz., the second is the toroidal vortex part and the third the poloidal vortex part.

As a result, the system of vector equations (1) and (2) takes the scalar form

$$\nabla^2\Psi_i(\mathbf{r},t) - \frac{1}{\nu}(1-\delta_{i1})\frac{\partial\Psi_i(\mathbf{r},t)}{\partial t} = 0 \quad (i=1,2,3),$$

$$P^{in}(\mathbf{U},t) = -\frac{\partial}{\partial t}\Psi_1(\mathbf{r},t). \tag{7}$$

Expressing the components of the velocity field $\mathbf{U}(\mathbf{r},t)$ in terms of Ψ_i, we obtain

$$U_r = \frac{\partial\Psi_1}{\partial r} - \frac{1}{r}\nabla_\Omega^2\Psi_3,$$

$$U_\Theta = \frac{1}{r}\frac{\partial\Psi_1}{\partial\Theta} + \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Psi_3}{\partial\Theta}\right),$$

$$U_\varphi = -\frac{\partial\Psi_2}{\partial\Theta}.$$

After some relatively simple mathematical transformations, the boundary conditions (3) on the surface of the solid core for projections of the velocity field can be expressed in terms of the scalar functions Ψ_i in the form

$$\text{for } r = R_0: \quad \frac{\partial\Psi_1}{\partial r} - \frac{1}{r}\nabla_\Omega^2\Psi_3 = 0, \tag{8}$$

$$\frac{1}{r}\Psi_1 + \frac{1}{r}\frac{\partial}{\partial r}(r\Psi_3) = 0, \tag{9}$$

$$\Psi_2 = 0. \tag{10}$$

The boundary conditions (4)–(6) take the form

$$\text{for } r = 1: \quad \frac{\partial\xi(\Theta,t)}{\partial t} = \left[\frac{\partial\Psi_1}{\partial r} - \frac{1}{r}\nabla_\Omega^2\Psi_3\right], \tag{11}$$

$$2\frac{\partial}{\partial r}\left(\frac{\Psi_1}{r}\right) + \frac{\partial^2\Psi_3}{\partial r^2} - \frac{1}{r^2}(2+\nabla_\Omega^2)\Psi_3 = 0, \tag{12}$$

$$r\frac{\partial}{\partial r}\frac{\Psi_2}{r} = 0, \tag{13}$$

$$-P^{in}(\mathbf{U},t) + 2\nu\left[\frac{\partial^2\Psi_1}{\partial r^2} - \nabla_\Omega^2\left(\frac{\partial}{\partial r}\left(\frac{\Psi_3}{r}\right)\right)\right]$$

$$-P_E(\xi) + P_\sigma(\xi) = 0. \tag{14}$$

3. Setting

$$\xi(\mathbf{r},t) = \sum_m Z_m P_m(\mu)\exp(st), \tag{15}$$

we seek solutions of the system of equations (7) in the form

$$\Psi_1(\mathbf{r},t) = \sum_m [C_m^1 r^m + D_m^1 r^{-(m+1)}]P_m(\mu)\exp(st),$$

$$\Psi_j(\mathbf{r},t) = \sum_m \left[C_m^j i_m\left(\sqrt{\frac{S}{\nu}}r\right) + D_m^j k_m\left(\sqrt{\frac{S}{\nu}}r\right)\right]P_m(\mu)$$

$$\times \exp(st) \quad (j=2,3), \tag{16}$$

where C_m^1, C_m^j, D_m^1, and D_m^j are constants.

We note that the problem of determining Ψ_2 is completely autonomous and does not depend on Ψ_1, Ψ_3, and ξ. In other words, the toroidal component of the motion of the liquid, which is described by Ψ_2, does not make a contribution to the dispersion relation of the harmonic poloidal motions of the liquid in the drop and does not influence the surface relief.

Substituting (16) into the boundary conditions (11)–(12) and (14), we obtain

$$mC_m^1 - (m+1)D_m^1 + m(m+1)$$

$$\times\left[C_m^3 i_m\left(\sqrt{\frac{S}{\nu}}\right) + D_m^3 k_m\left(\sqrt{\frac{S}{\nu}}\right)\right] = SZ_m, \tag{17}$$

$$2(m-1)C_m^1 - 2(m+2)D_m^1 + \left\{-2\sqrt{\frac{S}{\nu}}i_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)\right.$$

$$\left.+\left[\frac{S}{\nu}+2(m+1)(m-1)\right]i_m\left(\sqrt{\frac{S}{\nu}}\right)\right\}C_m^3$$

$$+\left\{2\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)\right.$$

$$\left.+\left[\frac{S}{\nu}+2(m+1)(m-1)\right]k_m\left(\sqrt{\frac{S}{\nu}}\right)\right\}D_m^3 = 0, \tag{18}$$

$$[S+2\nu m(m-1)]C_m^1+[S+2\nu(m+1)(m+2)]D_m^1$$

$$+2\nu m(m+1)\left[\sqrt{\frac{S}{\nu}}i_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)\right.$$

$$+(m-1)i_m\left(\sqrt{\frac{S}{\nu}}\right)\right]C_m^3+2\nu m(m+1)$$

$$\times\left[-\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)+(m-1)k_m\left(\sqrt{\frac{S}{\nu}}\right)\right]D_m^3$$

$$+\left[(m-1)(m+2)-\frac{Q^2}{4\pi}(m-1)\right]Z_m=0. \qquad (19)$$

In writing (17)–(19) we used the known[10] relations for cylinder functions

$$\frac{\partial}{\partial x}f_m(x)=f_{m+1}(x)+\frac{m}{x}f_m(x),$$

$$\frac{\partial^2}{\partial x^2}f_m(x)=-\frac{2}{x}f_{m+1}(x)+\left[1+\frac{m(m-1)}{x^2}\right]f_m(x),$$

$$f_m(x)=\begin{cases} i_m(x), \\ (-1)^{m+1}k_m(x). \end{cases}$$

Here the $i_m(x)$ and $k_m(x)$ are spherical cylinder functions of the first and third kinds, respectively. We write out the boundary conditions (8) and (9) for the components of the velocity field on the surface of the solid core with consideration of (16)

$$mR_0^{(m-1)}C_m^1-(m+1)R_0^{-(m+2)}D_m^1+\frac{m(m-1)}{R_0}$$

$$\times\left[C_m^3 i_m\left(\sqrt{\frac{S}{\nu}}R_0\right)+D_m^3 k_m\left(\sqrt{\frac{S}{\nu}}R_0\right)\right]=0, \qquad (20)$$

$$R_0^{(m-1)}C_m^1+R_0^{-(m+2)}D_m^1+\left[\sqrt{\frac{S}{\nu}}\,i_{m+1}\left(\sqrt{\frac{S}{\nu}}R_0\right)\right.$$

$$+\frac{(m+1)}{R_0}i_m\left(\sqrt{\frac{S}{\nu}}R_0\right)\right]C_m^3+\left[-\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}R_0\right)\right.$$

$$+\frac{(m-1)}{R_0}k_m\left(\sqrt{\frac{S}{\nu}}R_0\right)\right]D_m^3=0. \qquad (21)$$

The system of five homogeneous algebraic equations (17)–(21) for the five unknowns C_m^1, D_m^1, C_m^3, D_m^3, and Z_m have a nontrivial solution, only if the determinant composed of its coefficients equals zero:

$$\det A\equiv\begin{vmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} \end{vmatrix}=0, \qquad (22)$$

where $A_{11}=m$, $A_{12}=-(m+1)$, $A_{13}=m(m+1)i_m(\sqrt{S/\nu})$,

$$A_{14}=m(m+1)k_m\left(\sqrt{\frac{S}{\nu}}\right), \quad A_{15}=-S, \quad A_{21}=2(m-1),$$

$$A_{22}=-2(m+2), \quad A_{23}=-2\sqrt{\frac{S}{\nu}}i_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)$$

$$+\left[\frac{S}{\nu}+2(m+1)(m-1)\right]i_m\left(\sqrt{\frac{S}{\nu}}\right),$$

$$A_{24}=2\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)$$

$$+\left[\frac{S}{\nu}+2(m+1)(m-1)\right]k_m\left(\sqrt{\frac{S}{\nu}}\right),$$

$$A_{25}=0, \quad A_{31}=S+2\nu m(m-1),$$

$$A_{32}=S+2\nu(m+1)(m+2),$$

$$A_{33}=2\nu m(m+1)\left[\sqrt{\frac{S}{\nu}}i_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)+(m-1)i_m\left(\sqrt{\frac{S}{\nu}}\right)\right],$$

$$A_{34}=2\nu m(m+1)\left[-\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)\right.$$

$$+(m-1)k_m\left(\sqrt{\frac{S}{\nu}}\right)\right],$$

$$A_{35}=(m-1)(m+2)-\frac{Q^2}{4\pi}(m-1), \quad A_{41}=mR_0^{(m-1)},$$

$$A_{42}=-(m+1)R_0^{-(m+2)}, \quad A_{43}=\frac{m(m+1)}{R_0}i_m\left(\sqrt{\frac{S}{\nu}}R_0\right),$$

$$A_{44}=\frac{m(m+1)}{R_0}k_m\left(\sqrt{\frac{S}{\nu}}R_0\right), \quad A_{45}=0,$$

$$A_{51}=R_0^{(m-1)}, \quad A_{52}=R_0^{-(m+2)},$$

$$A_{53}=\sqrt{\frac{S}{\nu}}i_{m+1}\left(\sqrt{\frac{S}{\mu}}R_0\right)+\frac{(m+1)}{R_0}i_m\left(\sqrt{\frac{S}{\nu}}R_0\right),$$

$$A_{54}=-\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}R_0\right)+\frac{(m+1)}{R_0}k_m\left(\sqrt{\frac{S}{\nu}}R_0\right),$$

$$A_{55}=0.$$

Algebraic expression (22) relates the frequencies and the mode numbers to one another, i.e., is the dispersion relation defining the spectrum of possible harmonic poloidal and purely poloidal vortex motions in a liquid layer on the surface of solid spherical core.

4. Figures 1, 2, 3, 4, 5, and 6 present the results of numerical calculations based on (22) in the form of plots of the real and imaginary components of the dimensionless complex frequency as functions of the dimensionless radius of the core for various modes and various values of two dimensionless physical parameters, viz., the viscosity ν and the charge W.

In Figs. 1 and 2 the curves corresponding to $W=0$ and $\nu=0.03$ are presented for the modes with $m=2$ (the fundamental mode) and 6, respectively. In Fig. 3 the dependence
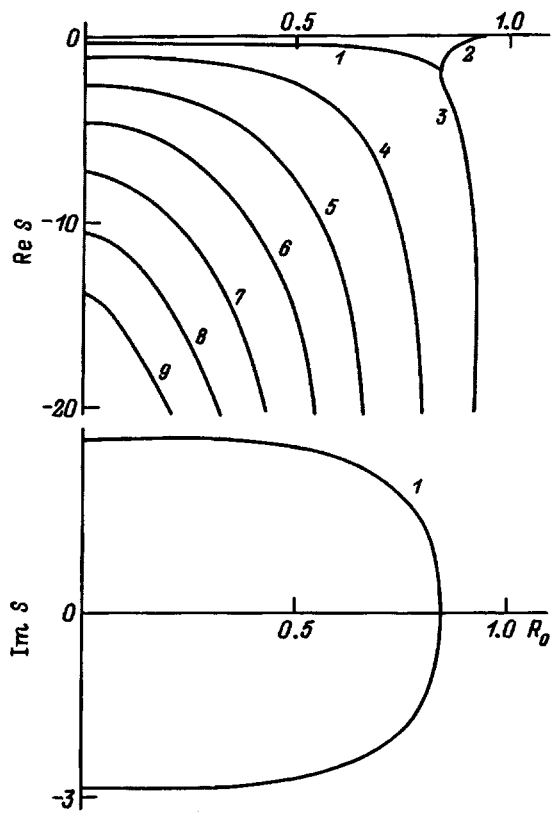
FIG. 1. Dependence of the real and imaginary components of the frequency S of the actually occurring capillary poloidal motions of the liquid on the core radius R_0.
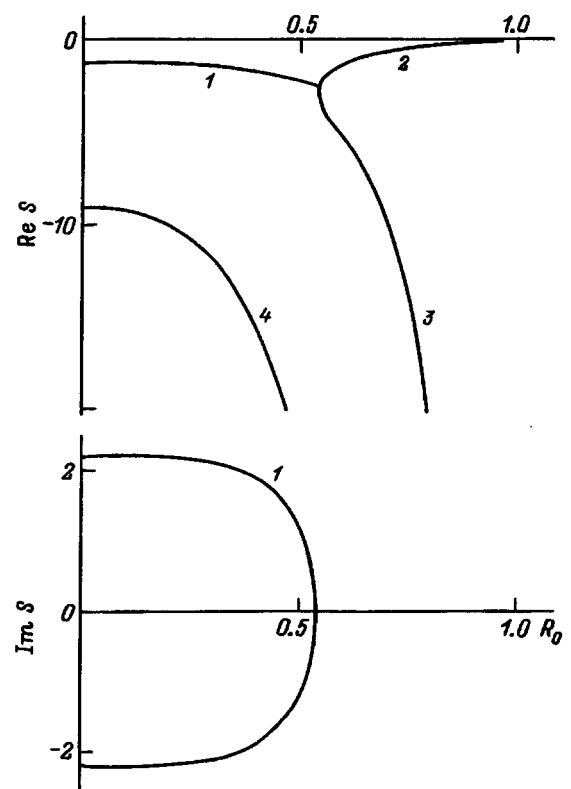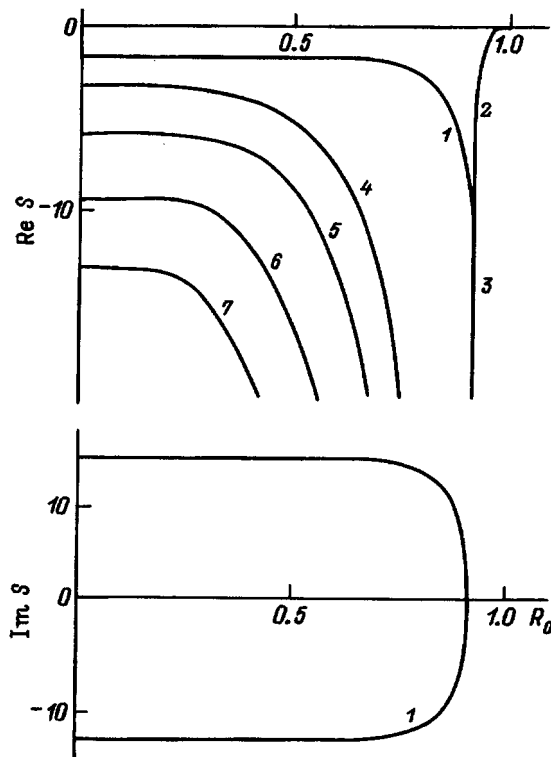

FIG. 2. Same as in Fig. 1, but for $m = 6$.

FIG. 3. Same as in Fig. 1, but for $\nu = 0.3$.

corresponding to a viscosity ten times larger ($\nu = 0.3$) and $W = 0$ is presented for the fundamental mode. It can be seen from these figures that both the damping in the bulk of the liquid layer and the damping on the core have a significant influence on the spectrum of actually occurring motions. Moreover, it is seen from Figs. 1 and 2 that the influence of damping on the core is more significant for the mode with $m = 2$ than for the mode with $m = 6$. It is also seen that at the viscosity value taken the range of values of the core radius R_0 in which capillary oscillations exist for the sixth mode is broader than the range for the fundamental mode, pointing out the decisive role of the damping of motions on the solid core for thin liquid layers. This phenomenon should be manifested (as a result of restriction of the spectrum of actually
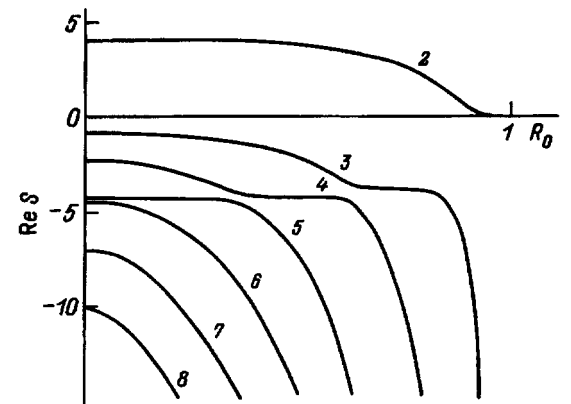

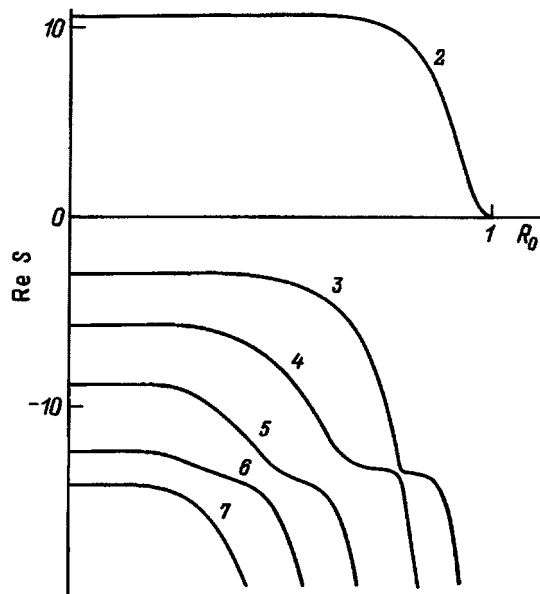FIG. 4. Same as in Fig. 1, but only the real component for $W = 13$.

FIG. 5. Same as in Fig. 4, but for $m = 6$.

occurring capillary waves) in the course of the final stage of instability of the charged surface of the liquid, when emission bulges form on it and emission of the excess charge begins.[1] In particular, in sufficiently thin layers of a viscous liquid, where damping on the core suppresses the first several modes of capillary oscillations of the liquid layer, instability will be realized on account of the preferential development of the first higher mode not suppressed by the viscosity with the number $m \equiv m_*$. This reduces to the formation of m_* emission bulges on the surface of a liquid layer that is unstable with respect to its self-charge, rather than the two such bulges observed for a drop without a core.[1] Just such a situation was apparently recorded on the photograph on p. 2437 in Ref. 11. It is seen from Figs. 1 and 3 that the frequencies of the capillary oscillations of an individual mode decrease with increasing viscosity.

The plots of $\mathrm{Re}S = \mathrm{Re}S(R_0)$ for a significantly supercritical charge in the system ($W = 13$) and $\nu = 0.03$ are presented for $m = 2$ in Fig. 4 and for $m = 6$ in Fig. 5 (we note that the modes with $m < 11$ are unstable when $W = 13$). The
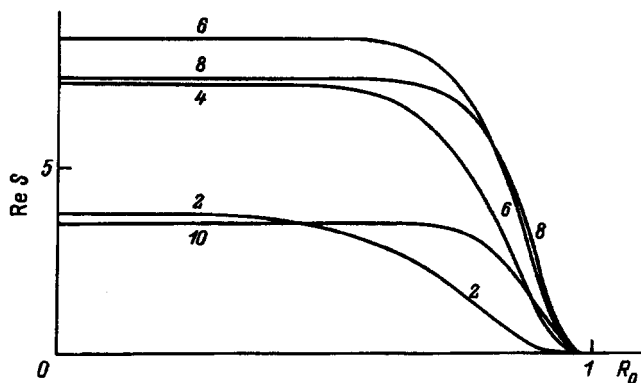

FIG. 6. Dependence of the dimensionless growth rates of several unstable modes on the dimensionless radius core R_0.

curves labeled *2* in these figures describe the instability growth rates of the respective modes. The curves with larger numbers describe the damping decrements of the poloidal vortex motions. It is not difficult to see that thinning of the liquid layer (an increase in R_0) leads to lowering of the growth rates of the unstable motions and to an increase in the decrements of the damped vortex motions of the liquid.

Figure 6 presents the calculated values of the growth rates of several unstable modes, whose numbers are indicated near the curves, for $W = 13$ and $\nu = 0.1$. It is clear from the physical meaning of the problem that the relationship between the growth rates and the mode numbers is determined by the degree of supercriticality of the charge and the influence of viscous damping in the bulk and on the surface of the solid core. We note that the mth mode becomes unstable when $W = 2 + m$. This means that at the value of W taken the charge is maximally supercritical for the fundamental mode with $m = 2$ and minimally supercritical for $m = 10$. Nevertheless, the growth rate of the fundamental mode in the system described here is not very great because of the significant influence of the viscous dissipation. Another significant point is that the ratio between the values of the growth rates of different modes depends on the thickness of the liquid layer: in thin layers the growth rates of the higher modes are greater than the growth rates of the modes with small numbers. The strong dependence of the growth rates on the thickness of the liquid layer at small thicknesses stands out.

In Figs. 1, 2, and 3 curves *1* describe the capillary oscillations of the liquid layer. In Figs. 1, 2, 3, 4, and 5 curves *2* and *3* describe aperiodically damped, harmonic poloidal motions. The curves with numbers greater than *3* correspond to aperiodically damped, purely poloidal vortex motions. It is easily seen from the curves presented that thinning the liquid layer, increasing its viscosity, and increasing the mode number lead to a rapid increase in the damping decrements of the motions of this type.

5. Let us turn to the problem of determining the toroidal vortex component of the velocity field associated with the scalar function Ψ_2, which is defined by the system of equations (10) and (12). Substituting (16) into (10) and (12), we find

$$C_m^2 i_m\left(\sqrt{\frac{S}{\nu}}R_0\right) + D_m^2 k_m\left(\sqrt{\frac{S}{\nu}}R_0\right) = 0, \qquad (23)$$

$$\left[\sqrt{\frac{S}{\nu}}i_{m+1}\left(\sqrt{\frac{S}{\nu}}\right) + (m-1)i_m\left(\sqrt{\frac{S}{\nu}}\right)\right]C_m^2$$

$$+ \left[-\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}\right) + (m-1)k_m\left(\sqrt{\frac{S}{\nu}}\right)\right]D_m^2 = 0. \qquad (24)$$

The system of algebraic equations (23) and (24) has a nontrivial solution when the determinant composed of the coefficients in front of C_m^2 and D_m^2 is equal to zero. This condition also defines the dispersion relation describing the spectrum of toroidal vortex motions in a layer of a viscous liquid on the surface of a solid spherical core, which has the form
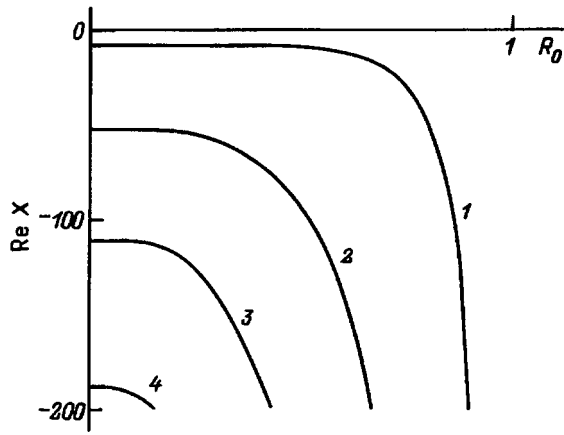
FIG. 7. Dependence of the damping decrements of the toroidal vortex motions associated with the fundamental mode ($m=2$) on the dimensionless core radius.

6. Summarizing the foregoing material, we note that the frequencies of the capillary oscillations of a liquid layer on the surface of a solid core, the decrements of the actually occurring poloidal and toroidal vortex motions, and the growth rates of the unstable modes depend very strongly on the presence of a solid core in the drop when the ratio of the core radius to the drop radius approaches unity. The existence of two mechanisms of viscous dissipation, viz., damping in the layer and on the solid core, restricts the spectrum of modes taking part in the formation of the emission bulges on the highly charged surface of the liquid layer, assigning the decisive role to the higher modes. This can be reduced for thin layers of viscous liquids to the formation of a large number of emission bulges, in contrast to a drop without a core, where there are only two such bulges.

$$i_m\left(\sqrt{\frac{S}{\nu}}R_0\right)\left[-\sqrt{\frac{S}{\nu}}k_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)+(m-1)k_m\left(\sqrt{\frac{S}{\nu}}\right)\right]$$

$$-k_m\left(\sqrt{\frac{S}{\nu}}R_0\right)\left[\sqrt{\frac{S}{\nu}}i_{m+1}\left(\sqrt{\frac{S}{\nu}}\right)\right.$$

$$\left.+(m-1)i_m\left(\sqrt{\frac{S}{\nu}}\right)\right]=0. \tag{25}$$

Figure 7 presents the dependence of the damping decrements of the toroidal vortex motions associated with the fundamental mode ($m=2$) on the thickness of the layer of the viscous liquid calculated from Eq. (25) in the form of plots of Re $x=$ Re $x(R_0)$, where $x\equiv S/\nu$. It is seen from Fig. 7 that thinning of the liquid layer leads to a rapid increase in the damping decrements of the toroidal motions. Numerical calculations also show that the decrements of the toroidal motions associated with a specific mode increase rapidly as the mode number increases.

[1] A. I. Grigor'ev and S. O. Shiryaeva, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza (3), 3 (1994).

[2] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **59** (5), 6 (1989) [Sov. Phys. Tech. Phys. **34**, 502 (1989)].

[3] N. B. Zolotoĭ, G. V. Karpov, and V. E. Skurat, Zh. Tekh. Fiz. **58** (2), 315 (1988) [Sov. Phys. Tech. Phys. **33**, 193 (1988)].

[4] S. O. Shiryaeva and A. I. Grigor'ev, Zh. Tekh. Fiz. **63** (8), 162 (1993) [Tech. Phys. **38**, 715 (1993)].

[5] A. I. Grigor'ev, M. I. Munichev, and S. O. Shiryaeva, J. Colloid Interface Sci. **166**, 267 (1994).

[6] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **62** (12), 9 (1992) [Sov. Phys. Tech. Phys. **37**, 1136 (1992)].

[7] A. I. Grigor'ev and A. É. Lazaryants, Zh. Tekh. Fiz. **60** (6), 29 (1990) [Sov. Phys. Tech. Phys. **35**, 651 (1990)].

[8] V. G. Levich, *Physicochemical Hydrodynamics*, Prentice-Hall, Englewood Cliffs, N.J. (1962).

[9] S. O. Shiryaeva, A. É. Lazaryants *et al..*, Preprint No. 27, Institute of Microelectronics, Russian Academy of Sciences, Yaroslavl, 1994.

[10] *The Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun (eds.), Dover, New York (1976) [Russian trans., Nauka, Moscow (1969)].

[11] G. A. Ostroumov, *Interaction of Electric and Hydrodynamic Fields* [in Russian], Nauka, Moscow (1979).

Translated by P. Shelnitz

Shaping of the electron distribution function in a striated solution

Yu. B. Golubovskiĭ, V. O. Nekuchaev, N. S. Ponomarev, and I. A. Porokhova

Scientific-Research Institute of Physics, St. Petersburg State University, 198904 St. Petersburg, Russia

Numerous papers have been devoted to the investigation of striations in inert gases at low pressures ($p \leqslant 2$ Torr) and small currents ($i < 100$ mA) [A. V. Nedospasov, Sov. Phys. Usp. **11**, 174 (1968); L. Pekarek, Sov. Phys. Usp. **11**, 188 (1968); N. L. Oleson and A. W. Cooper, Adv. Electron. Electron Phys. **24**, 155 (1968); P. S. Landa, N. A. Miskinova, and Yu. V. Ponomarev, Sov. Phys. Usp. **23**, 813 (1980)]. Since the nature of striations is determined under these conditions by the nonlocal kinetics of the electrons in spatially periodic fields [L. D. Tsendin, Sov. J. Plasma Phys. **8**, 228 (1982)], an investigation of the electron distribution function in space and time would be very interesting. The purpose of the present work is to experimentally investigate the potential profiles and distribution functions in S and P striations and to analyze the mechanism which shapes the distribution functions for striations of these types. © *1997 American Institute of Physics.* [S1063-7842(97)00309-7]

SPATIAL POTENTIAL PROFILES IN STRIATIONS

The potential profiles and distribution functions in different phases of striations have been measured in several studies.[6–8] As a rule, the temporal variation of the potential was measured using a stationary probe and then transformed into the spatial variation by replacing the time by $x = vt$, where v is the phase velocity of the striation. Such calculations frequently led to a significantly nonmonotonic dependence of the potential on the longitudinal coordinate, which was attributed to the presence of potential wells and, accordingly, reversed fields. Such a procedure for reconstructing the spatial potential profile can be incorrect, if along with wave propagation there are oscillations of the plasma potential as a whole with a frequency equal to the frequency of the striations, for example, relative to the anode. These oscillations can be caused by self-modulation of the striations, which leads to oscillations of the voltage on the tube electrodes. The amplitude of the voltage oscillations depends on the ballast resistance. It was shown in Ref. 9 that under the conditions of a nonlocal mechanism for shaping the electron distribution function and a discharge gap of finite length, voltage oscillations appear in the near-anode region due to the passage of striations through this region.

In the present work systematic measurements of the electron distribution function were performed for a neon discharge in a tube of radius $R = 1.4$ cm at pressures $p = 1 - 2$ Torr and currents $i = 10 - 20$ mA using a mobile probe at different distances from the anode on the discharge axis over three wavelengths with a spacing of $L/10$ for P striations and $L/15$ for S striations (L is the striation thickness). Measurements of the electron distribution function were performed with a temporal resolution of 10 μs in 10–12 phases of a striation period at each fixed position of the probe. The measurements were taken by the standard method from the second derivative of the probe current, and the plasma potential relative to the anode was determined from the point where the second derivative vanishes.

Figures 1a–1d present the variable components of the oscillations of the plasma potential with time for various points along the tube axis within one wavelength. It is seen that the oscillations of the potential with time at different points differ significantly, but the characteristic minimum of each oscillation (which is marked by an arrow in the figure) is achieved at the same time. When the probe is displaced by one wavelength, the picture is repeated (the curves in Figs. 1a and 1d coincide). Such a spatiotemporal picture of the behavior of the potential indicates that there is superposition of the oscillations of the plasma potential as a whole with time and the variation of the potential associated with wave propagation. An attempt to reconstruct the spatial potential profile from the temporal profile by replacing t by $x = vt$ (for example, for case *a*) leads, with consideration of the constant field E_0 in the plasma, to the profile shown in Fig. 1a. One characteristic feature of the profiles thus obtained is the apparent presence of pronounced potential wells and reversed fields. A correct procedure requires elimination of the temporal fluctuations, which can be accomplished by direct measurements of the plasma potential along the axis with a fixed time delay relative to the voltage oscillations on the electrodes. The results of such measurements are presented in Fig. 2 for S and P striations at two different moments in time, which correspond to the zero (curves *1s* and *1p*) and maximum values (curves *2s* and *2p*) of the potential oscillations as a whole. It is seen from the figure that there are, in fact, no apparent potential wells of great depth. Similar results were obtained in Ref. 10, where the cathode and anode were displaced simultaneously relative to a stationary probe. The measured potential profile at distances exceeding one wavelength from the anode at different moments in time is similar. Oscillations develop near the anode at distances of the order of a few millimeters. It is noteworthy that the amplitude of these oscillations coincides approximately with the potential drop on an S or P striation (Fig. 2) and amounts to ~ 10 V (a P striation) or 17 V (an S striation). The anode fall for different striation phases is generally positive and varies with time from zero to the oscillation amplitude. However, at some moments in time the plasma potential at distances of
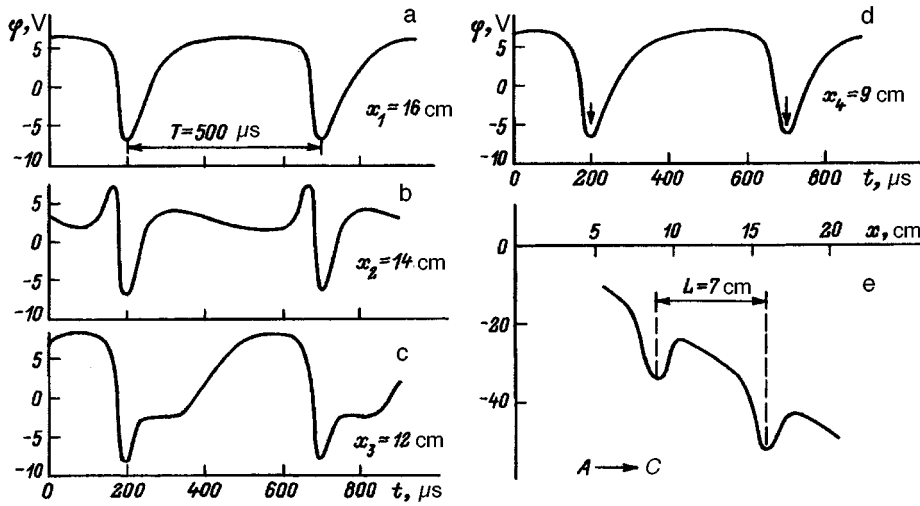
FIG. 1. Variable component of the plasma potential oscillations with time at various points on the discharge axis (a–d) and spatial potential profile (e) (A — anode, C — cathode).

~ 2 cm can be positive relative to the anode, and a potential well with a small depth of ~ 1 V forms. An example of such a potential curve is presented in Fig. 2 for a P striation (curve $1p$, hatched area). The presence of the potential well clearly correlates with the measured distribution functions. The distribution functions in a real potential well have a pronounced maximum in the slow-electron region, while the distribution functions measured in other striation phases for a potential without a well have the usual form. Experiments devised to measure the distribution function in the region of the apparent potential wells far from the anode do not reveal a characteristic peak in the slow-electron region, indicating once again the absence of these potential wells.

MECHANISM FOR SHAPING THE DISTRIBUTION FUNCTION IN S AND P STRIATIONS

According to numerous experimental data, the main difference between the S and P striations appearing under similar discharge conditions is the twofold difference between the potential drops over the striation thickness. For example, the characteristic potential drops in S striations amount to 17–21 V, while those in P striations equal 9–10.5 V for $p = 1-5$ Torr and $i = 1-100$ mA. In addition, the wavelengths in S striations are twice as great as the wavelengths in P striations at equal values of the period-averaged electric field E_0. According to the model proposed in Ref. 5, a wave which travels through space and along the distribution function forms in a spatially periodic field with a period $\varepsilon_L / e E_0$ due to the bunching of electrons according to their energy. The kinetic equation for the distribution function in the variables $\varepsilon = w + e \varphi(x)$ and x [ε is the total energy, $e \varphi(x)$ is the potential energy, and w is the kinetic energy] for energies $0 < w < \varepsilon_1$ under the conditions considered here can be written in the form

$$\frac{\partial}{\partial x} \frac{v^3}{3\nu} \frac{\partial f_0(\varepsilon,x)}{\partial x} + \frac{\partial}{\partial \varepsilon} \frac{m^2}{M} \nu v^3 f_0(\varepsilon,x) = 0. \tag{1}$$

It is assumed that the energy balance is dominated by inelastic collisions and that interelectronic collisions are negligibly small. If the drop of the electron distribution function in the inelastic region is fairly steep, a zero boundary condi-

tion can be imposed in an approximation for the electron distribution function at the excitation threshold:

$$f_0(\varepsilon,x)|_{w=\varepsilon_1} = 0. \tag{2}$$

Equation (1) describes the diffusion of the electron distribution function along the coordinate x and its drift with respect to the energy ε due to elastic collisions in the x, ε phase plane. We approximate the potential profile by functions of the form
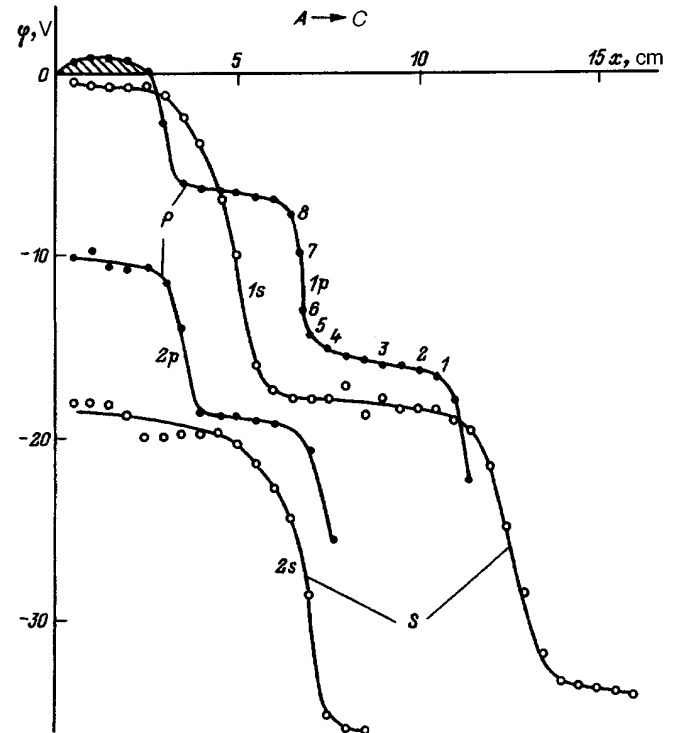

FIG. 2. Spatial potential profile measured by a mobile probe for S striations (curves $1s$ and $2s$) and P striations (curves $1p$ and $2p$) at the moments in time corresponding to zero ($1s$ and $1p$) and maximum ($2s$ and $2p$) oscillations of the plasma potential. Points 1–8, at which the measured electron density distributions were compared with the calculated distributions, are marked on curve $1p$.
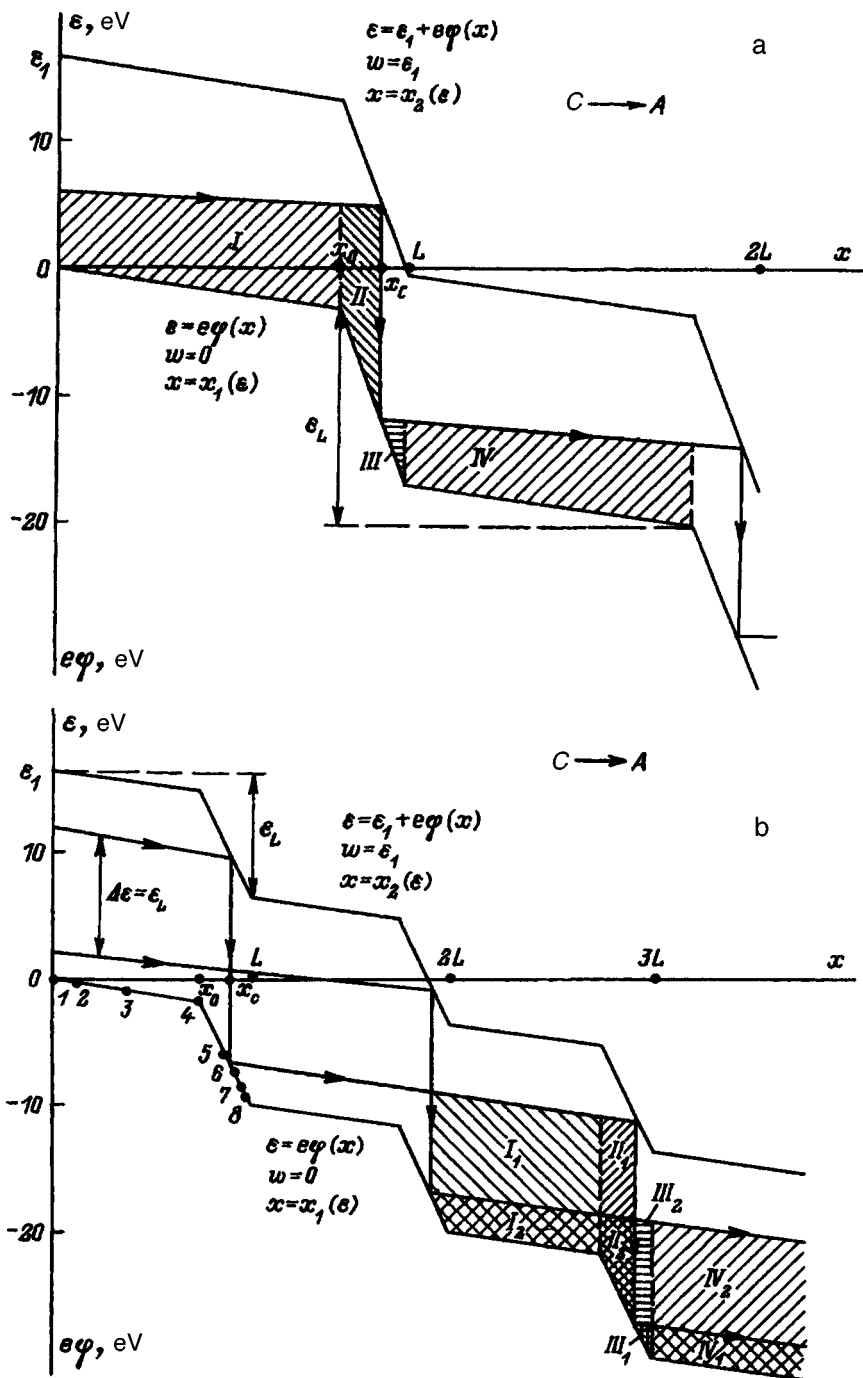
FIG. 3. Approximation of the potential fields in S striations (a) and P striations (b), in which kinetic equation (1) was solved with boundary condition (2). Curves in $\varepsilon = e\varphi(x)$, $x = x_1(\varepsilon)$ coordinates — the potential profile; curves in $\delta = \varepsilon_1 + e\varphi(x)$, $x = x_2(\varepsilon)$ — the potential profile shifted on the total energy scale by ε_1. The resonant trajectories are marked by arrows, and ε_L is the potential drop on a striation. The calculated and measured electron distribution functions were compared at points 1–8. The kinetic energy regions for the resonant trajectories in the weak-field (I and IV) and strong-field (II and III) regions are hatched. The maximum of $\Phi(w, x_c)$ undergoes a jump from $w = \varepsilon_1$ to $w = 0$ at x_c.

$$e\varphi(x) = \begin{cases} -eE_1 x, & 0 < x < x_0, \\ -eE_1 x_0 - eE_2(x - x_0), & x_0 < x < L, \end{cases}$$

$$e\varphi(x + nL) = e\varphi(x) - n\varepsilon_L, \tag{3}$$

where E_1 and E_2 are the values of the fields in regions of weak $0 < x < x_0$ and strong $x_0 < x < L$ variation of the potential, and ε_L is the potential drop on the striation.

Figure 3 presents the potential curves on the ε, x phase plane for S and P striations. On the $x = x_1(\varepsilon)$ curve the kinetic energy is equal to zero, while the $x = x_2(\varepsilon)$ curve is

shifted upward by ε_2, and $w|_{x = x_2(\varepsilon)} = \varepsilon_1$. It would be interesting to analyze the character of the motion of electrons in this phase plane.

Let us consider the case of an S striation (Fig. 3a). If the energy losses in elastic collisions are neglected, electrons will move with conservation of the total energy ($\varepsilon = $ const) along horizontal straight lines from the $x_1(\varepsilon)$ curve to the $x_2(\varepsilon)$ curve, and then, after losing the energy ε_1 in an inelastic collision, they jump abruptly long a vertical to the $x_1(\varepsilon)$ curve. The gain in kinetic energy on any of the trajectories and the potential drop on a striation should be equal to ε_1. If we take into account the small energy losses in the
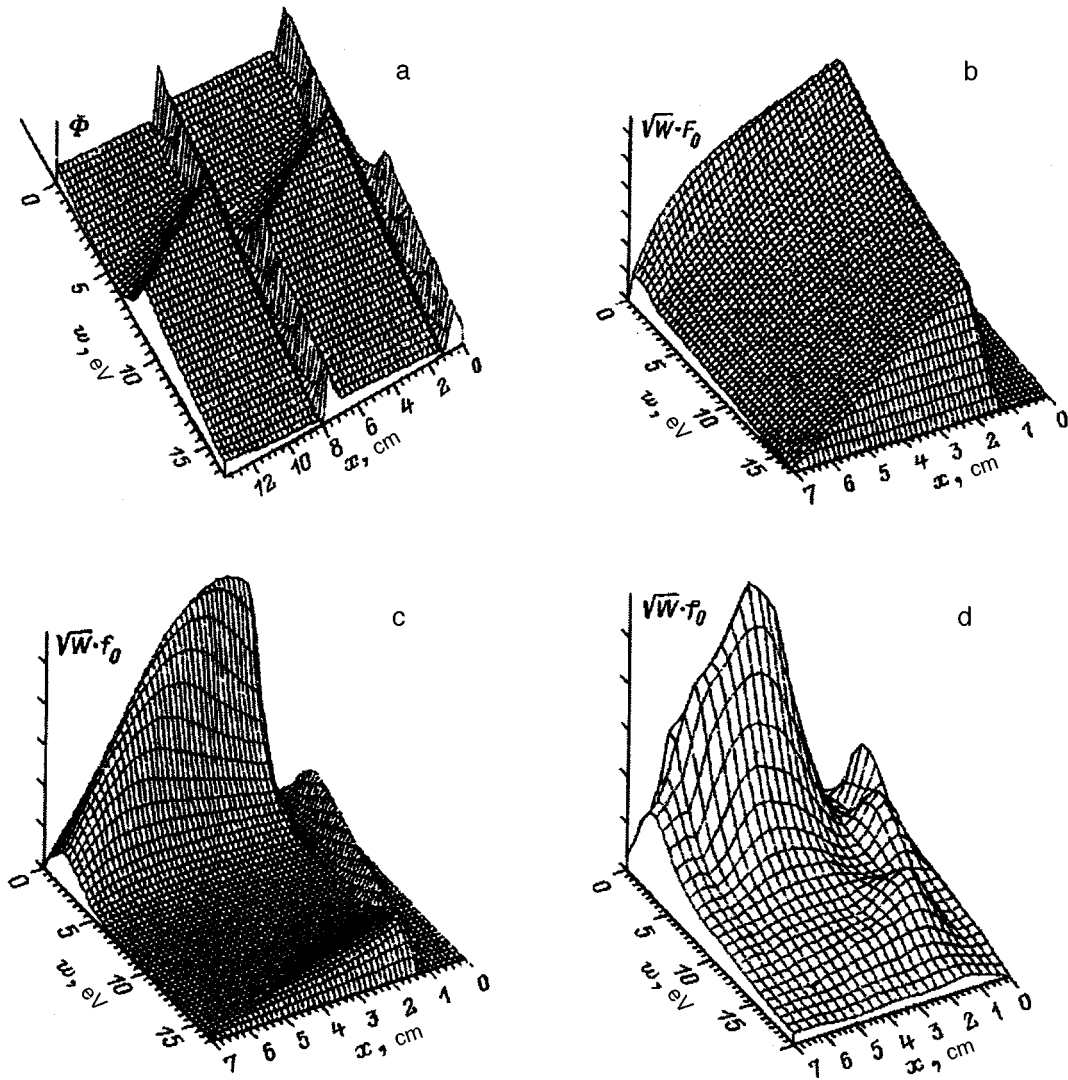
FIG. 4. S striations: a — movement of the maximum of $\Phi(\delta)$ in the w, x plane along the resonant trajectory in a two-wavelength interval; b — the distribution function $\sqrt{w}F_0(w,x)$; c — the resultant distribution function $\sqrt{w}f_0(w,x)$; d — the experimentally measured distribution function $\sqrt{w}f_0(w,x)$.

elastic collisions, the trajectory bends, and the energy gain in the potential field $e\varphi(x)$ on any trajectory from $x_1(\varepsilon)$ to $x_2(\varepsilon)$ is equal to the energy losses in the elastic and inelastic collisions. The potential drop on the striation thickness ε_L exceeds the excitation potential ε_1. The trajectory on which an electron loses the energy $\Delta\varepsilon = \varepsilon_L - \varepsilon_1$ in elastic collisions during one period is preferred. All the other trajectories converge toward the preferred trajectory (the resonant trajectory) after several periods, because the energy losses in elastic collisions are proportional to the kinetic energy, i.e., because of the bunching effect described in Ref. 5.

In the case of P striations (Fig. 3b) the potential drop on a striation is approximately two times smaller than for S striations and is smaller than the excitation threshold ε_1. For this reason, electrons should traverse two spatial periods of variation of the potential as they move from the $x_1(\varepsilon)$ curve to the $x_2(\varepsilon)$ curve in a potential field. This case corresponds to the two resonant trajectories in Fig. 3b. Therefore, an appreciable difference should be expected between the distribution functions in S and P striations.

The solution of Eqs. (1) and (2) in a spatially periodic

field was discussed in Ref. 5. In Ref. 11 the electron distribution function was calculated from an experimentally measured potential profile for S striations near the lower current limit for its existence. The solution of Eqs. (1) and (2) in the case of small energy losses in elastic collisions can be represented in the form

$$f_0(\varepsilon,x) = \Phi(\varepsilon) \int_x^{x_2(\varepsilon)} \frac{\nu(\varepsilon,x')}{v^3(\varepsilon,x')}\, dx' = \Phi(\varepsilon)F_0(\varepsilon,x),$$

$$(4)$$

where $\Phi(\varepsilon)$ is the amplitude of the distribution function, and $F_0(\varepsilon,x)$ corresponds to the electron distribution function formed in a field without energy losses in elastic collisions.

Taking into account small energy losses by expanding the electron distribution function in the small parameter $\Theta = 3m^2\varepsilon_1 v^2(\varepsilon_1)/Me^2E_0^2$ with consideration of the terms that are quadratic with respect to Θ leads to an equation for $\Phi(\varepsilon)$, which has the following form in the dimensionless variables $\tilde{\varepsilon} = \varepsilon/\varepsilon_1$, $\Psi_1 = (\nu(v)/v^3)/(\nu(v_1)/v_1^3)$, and $\Psi_2 = \nu(v)v^3/(\nu(v_1)v_1^3)$:
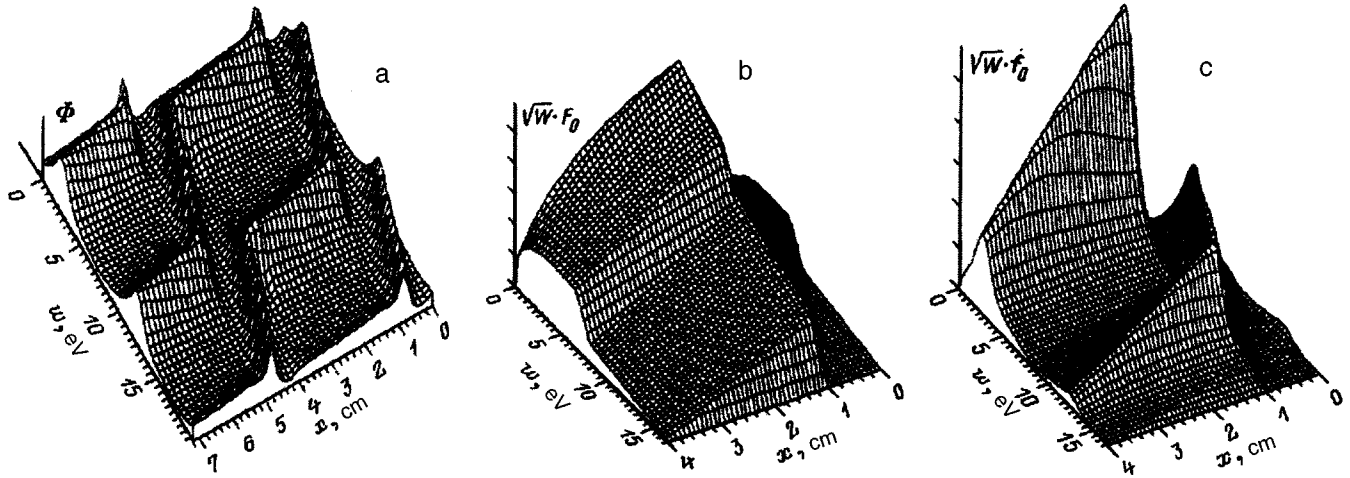
FIG. 5. P striations: a — movement of the two maxima of $\Phi(\varepsilon)$ in the w, x plane along the resonant trajectories in a two-wavelength interval; b — the distribution function $\sqrt{w}F_0(w,x)$; c — the resultant distribution function $\sqrt{w}f_0(w,x)$.

$$\Phi(\bar{\varepsilon}-1)=\Phi(\bar{\varepsilon})+\Theta\frac{\partial}{\partial\bar{\varepsilon}}\Phi(\bar{\varepsilon})\Psi(\bar{\varepsilon})$$

$$+\Theta^2\frac{\partial^2}{\partial\bar{\varepsilon}^2}\Phi(\bar{\varepsilon})C(\bar{\varepsilon}), \qquad (5)$$

$$\Psi(\tilde{\varepsilon})=\int_{x_1(\bar{\varepsilon})}^{x_2(\bar{\varepsilon})}\Psi_2F_0(\tilde{\varepsilon},x)dx, \qquad (6)$$

$$C(\tilde{\varepsilon})=\int_{x_1(\bar{\varepsilon})}^{x_2(\bar{\varepsilon})}\Psi_2dx\int_x^{x_2(\bar{\varepsilon})}\frac{dx'}{\Psi_1}\int_{x_1(\bar{\varepsilon})}^{x'}\Psi_2F_0(\tilde{\varepsilon},x'')dx''. \qquad (7)$$

The quantity $\Theta\Psi(\tilde{\varepsilon})$ has the physical meaning of the energy losses of an electron with an energy ε in elastic collisions as it moves from moves from the $x_1(\varepsilon)$ curve to the $x_2(\varepsilon)$ curve (the energy drift). The quantity $\Theta^2C(\tilde{\varepsilon})$ is equivalent to the energy diffusion coefficient due to the gradients $\partial\Phi/\partial\tilde{\varepsilon}$ that appear. A periodic structure with an energy period equal to ε_1 should form for $\Theta=0$. When $\Theta\neq0$, the energy period for $\Phi(\varepsilon)$ increases due to the energy losses in elastic collisions. In fact, in a spatially periodic field $\Psi(\varepsilon)$ can be represented in the form

$$\Psi(\varepsilon)=A+\beta(\varepsilon), \quad \frac{1}{\varepsilon_L}\int_0^{\varepsilon_L}\beta(\varepsilon)d\varepsilon=0,$$

$$A=\frac{1}{\varepsilon_L}\int_0^{\varepsilon_L}\Psi(\varepsilon)d\varepsilon, \qquad (8)$$

where $\beta(\varepsilon)$ is a periodic function with respect to the energy.

The quantity ΘA has the physical meaning of the mean energy loss in elastic collisions in one spatial period. Shifting the argument by ΘA in Eq. (5) and expanding into a series in this small parameter with consideration of the quadratic terms, we obtain the equation

$$\Phi(\tilde{\varepsilon}-1-\Theta A)-\Phi(\tilde{\varepsilon})$$

$$=\frac{\partial}{\partial\tilde{\varepsilon}}\left[\Theta\Phi(\tilde{\varepsilon})\beta(\tilde{\varepsilon})+\Theta^2\frac{\partial}{\partial\tilde{\varepsilon}}\Phi(\tilde{\varepsilon})k(\tilde{\varepsilon})\right], \qquad (9)$$

where $k(\tilde{\varepsilon})=C(\tilde{\varepsilon})-(A^2/2)$.

We shall henceforth assume that $k(\tilde{\varepsilon})$ depends weakly on the energy.[5] The steady-state solution (9) can be obtained by setting the right-hand side equal to zero, which corresponds to an equality between the energy diffusion and drift.[5] Equality of the left-hand side to zero corresponds to the energy periodicity of $\Phi(\varepsilon)$ with a period equal to $\varepsilon_L=\varepsilon_1(1+\Theta A)$. Integrating the right-hand side of Eq. (9) and assuming that the energy flux in (9) is equal to zero in the steady state, we obtain

$$\frac{\partial\Phi(\tilde{\varepsilon})}{\partial\tilde{\varepsilon}}+\frac{\beta(\tilde{\varepsilon})}{\Theta k}\Phi(\tilde{\varepsilon})=0. \qquad (10)$$

The solution of Eq. (10) has the form

$$\Phi(\tilde{\varepsilon})=D\exp\left\{-\int_0^{\tilde{\varepsilon}}\frac{\beta(\tilde{\varepsilon})}{\Theta k}d\tilde{\varepsilon}\right\}, \qquad (11)$$

which permits construction of the distribution function from Eq. (4). The constant D can be found from the normalization condition on the current density in terms of the directed part of the electron distribution function.

The results of the calculations for S striations in a potential field (Fig. 3a) replotted in the plane of the kinetic energy and the spatial coordinate are presented in Figs. 4a–c. Figure 4a shows the transformation of the amplitude of the electron distribution function $\Phi(w,x)$ in a two-wavelength interval. This amplitude has a characteristic maximum, which slowly moves along the w coordinate as x increases in the weak-field region ($0<x<x_0$) in accordance with Fig. 3a, where the kinetic energy varies weakly along the resonant trajectory (region I in Figs. 3a and 4a). Upon passage into the strong-field region (region II), the maximum of the ampli-

tude of the electron distribution function moves abruptly along the kinetic energy coordinate to the value ε_1, which is achieved at x_c. As it passes through x_c, the maximum abruptly shifts to zero kinetic energy. In the strong-field region (region *III* in Fig. 3a) the maximum moves rapidly along the kinetic energy coordinate to an energy value corresponding to a weak field. Then the maximum enters the

weak-field region (region *IV* in Fig. 3a), and this sequence of events repeats periodically with a period L (the striation thickness). Figure 4b presents the energy distribution function which would form in the initial field (Fig. 3a) with neglect of the elastic losses $[\sqrt{w}F_0(w,x)$ (4)]. Figure 4c presents the final electron distribution function $\sqrt{w}f_0(w,x)$
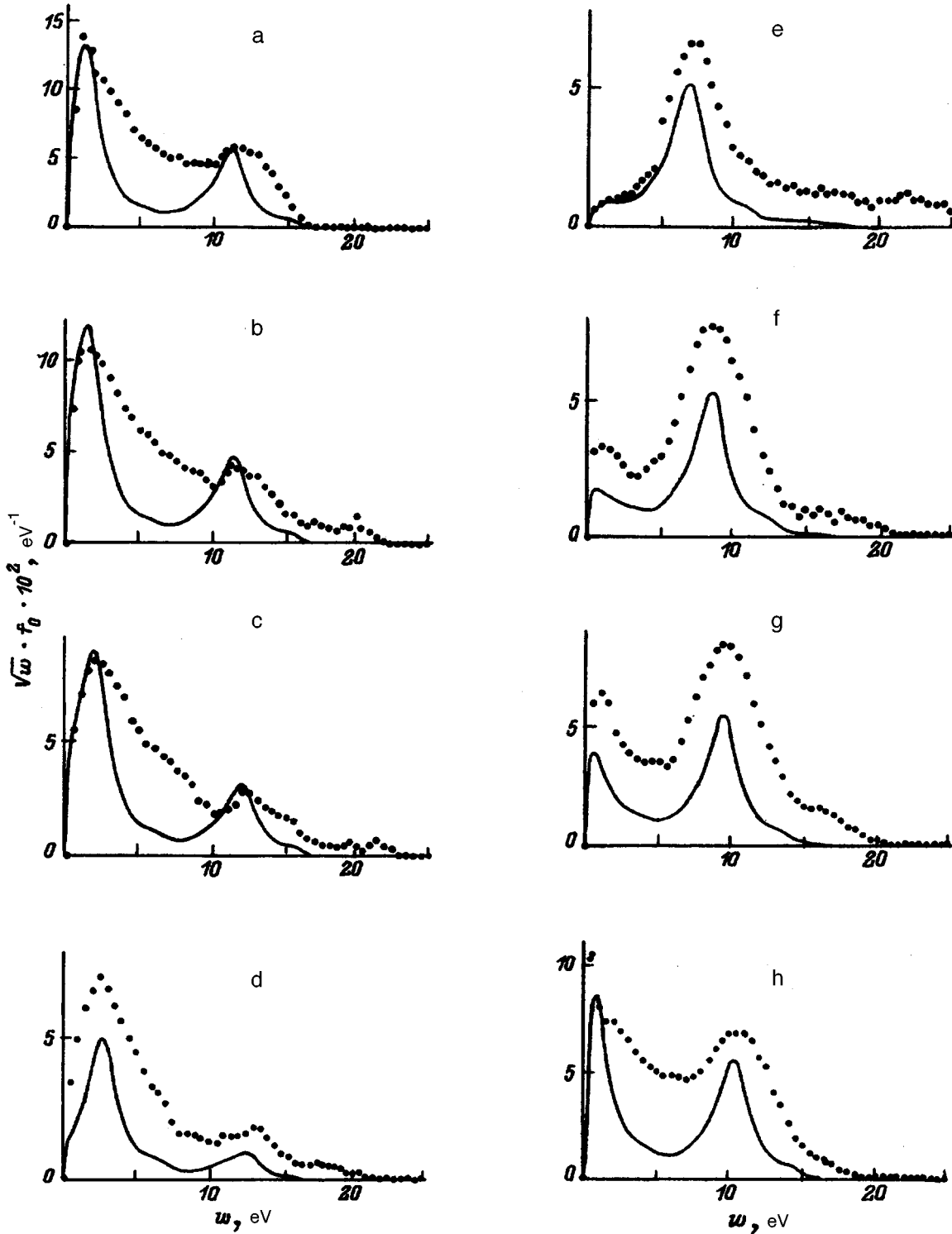

FIG. 6. Comparison of the calculated distribution functions $\sqrt{w}f_0(w,x)$ (solid curves) with the measured distribution functions (points) for a P striation in the eight phases corresponding to points *1–8* in Fig. 3b (calculation) and in Fig. 2 (experiment): a–c — weak-field region; d — transition from a weak field to a strong field; e–g — strong-field region; h — transition from a strong field to a weak field.

calculated from (4) and (11). The figure shows how the perturbation wave caused by electron bunching along the resonant trajectory propagates along the electron distribution function. This effect was observed in the experiments performed in Ref. 11, where the results of calculations and measurements of the electron distribution function in different phases of S striations were compared. The measurements of the electron distribution function performed in the present work using a mobile probe in 15 phases of S striations permit the construction of the three-dimensional picture of $\sqrt{w}f_0(w,x)$ presented in Fig. 4a, which clearly exhibits a correlation with the calculated picture (Fig. 4c).

The shaping of the electron distribution function in P striations is a very interesting, but scarcely studied question. An approximation of the measured potential profile in accordance with (3) is presented in Fig. 3b. The results of calculations of the amplitude of the electron distribution function $\Phi(w,x)$ in this potential field according to the method described above are presented in Fig. 5a in a two-wavelength interval. Since for P striations the potential drop on a striation ε_L is smaller than the excitation threshold ε_1, there are two maxima of $\Phi(\varepsilon)$ separated by ε_L in the kinetic energy range $0 < w < \varepsilon_1$. These two maxima move along two resonant trajectories (Fig. 3b), in accordance with Fig. 5a, in the w, x plane. The function $\sqrt{w}F_0(w,x)$ which would form in the potential field of a P striation without consideration of the energy losses in elastic collisions is shown in Fig. 5b. A distinctive feature (in comparison to an S striation) is the presence of two characteristic discontinuities in the weak-field region, which are caused by the structure of the potential profile ($\varepsilon_L < \varepsilon_1$). Figure 5c presents the resultant electron distribution function $\sqrt{w}f_0(w,x)$ (4) which forms in a P striation with consideration of the energy losses in elastic collisions. As we see, in the weak-field region the P striation has an additional maximum (in comparison to an S striation) in the vicinity of 12 eV, which moves in accordance with the second resonant trajectory.

The results of the calculations and experiments at the eight points along a P striation marked in Figs. 2 and 3b are compared in Fig. 6. The theory and experiment show that in the weak-field region (Figs. 6a–c) there are two maxima, which scarcely move along the energy coordinate and are separated by a distance corresponding to a potential drop of ~ 10 V on a P striation. The absolute values of the electron distribution function decrease in accordance with the behavior of $\sqrt{w}F_0(w,x)$. The transition to a strong field (Fig. 6d) is accompanied by displacement of the maxima toward higher energies. In the strong-field region (Figs. 6e–g) the second maximum passes through the excitation threshold and returns to the slow-electron region, while the first maximum

gradually moves along the energy coordinate to $w \sim 12$ eV, achieving this position on the boundary between strong and weak fields. On this boundary the second maximum corresponds to $w \sim 2$ eV, and the picture repeats periodically.

The theory faithfully describes the experimental data with respect to the form of the distribution function in different striation phases. The modulation depth of the electron distribution function and of the density is somewhat greater in the theory than in the experiment.

CONCLUSIONS

When ionization waves in a plasma are investigated, the possible oscillations of the plasma potential as a whole relative to the anode must be taken into account to reproduce the potential profile. These oscillations can be caused by the passage of striations through the near-anode region. The use of a mobile probe in the potential measurements can lead to errors associated with the appearance of apparent potential wells on the reconstructed spatial potential profile.

The mechanisms which shape the electron distribution function in S and P striations in measured potential fields under the conditions of nonlocal electron kinetics have been considered on the basis of Ref. 5. It has been shown that for P striations (as opposed to S striations) there are two resonant trajectories due to electron bunching. The movement of two characteristic maxima of the amplitude of the distribution function along these resonant trajectories, which is superimposed on the unperturbed electron distribution function formed in an assigned potential field without consideration of the energy losses in elastic collisions, describes the experimentally observed features of the behavior of the electron distribution function in P striations.

[1] A. V. Nedospasov, Usp. Fiz. Nauk **94**, 439 (1968) [Sov. Phys. Usp. **11**, 174 (1968)].
[2] L. Pekarek, Usp. Fiz. Nauk **94**, 463 (1968) [Sov. Phys. Usp. **11**, 188 (1968)].
[3] N. L. Oleson and A. W. Cooper, Adv. Electron. Electron Phys. **24**, 155 (1968).
[4] P. S. Landa, N. A. Miskinova, and Yu. V. Ponomarev, Usp. Fiz. Nauk **132**, 601 (1980) [Sov. Phys. Usp. **23**, 813 (1980)].
[5] L. D. Tsendin, Fiz. Plazmy **8**, 400 (1982) [Sov. J. Plasma Phys. **8**, 228 (1982)].
[6] N. D. Twiddy and S. W. Rayment, J. Phys. D: Appl. Phys. **2**, 1747 (1969).
[7] Yu. M. Kagan, N. B. Kolokolov, T. A. Krylova, and V. M. Milenin, Zh. Tekh. Fiz. **41**(1), 120 (1971) [Sov. Phys. Tech. Phys. **16**, 88 (1971)].
[8] K. F. Bessonova, D. M. Oreshak, E. P. Ostapchenko, and V. A. Stepanov, Zh. Tekh. Fiz. **41**(5), 979 (1971) [Sov. Phys. Tech. Phys. **16**, 770 (1971)].
[9] Yu. B. Golubovskiǐ, V. O. Nekuchaev, and I. É. Suleǐmanov, Zh. Tekh. Fiz. **63**(3), 194 (1993) [Tech. Phys. **38**, 257 (1993)].
[10] A. B. Stewart, J. Appl. Phys. **27**, 911 (1956).
[11] Yu. B. Golubovskiǐ and S. U. Nisimov, Zh. Tekh. Fiz. **66**(7), 20 (1996) [Tech. Phys. **41**, 645 (1996)].

Translated by P. Shelnitz